

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
FACULDADE DE BIBLIOTECONOMIA E COMUNICAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM COMUNICAÇÃO  
NÍVEL DOUTORADO

Luiza Carolina dos Santos

**MÁQUINAS QUE FALAM (E ESCUTAM):**  
**as formas de agência e de interação das/com as assistentes pessoais digitais**

Porto Alegre

2020

LUIZA CAROLINA DOS SANTOS

MÁQUINAS QUE FALAM (E ESCUTAM):

as formas de agência e de interação das/com as assistentes pessoais digitais

Tese de doutorado apresentada à banca examinadora do Programa de Pós-Graduação em Comunicação da Universidade Federal do Rio Grande do Sul (PPGCOM-UFRGS) como requisito parcial para obtenção do título de doutor em Comunicação.

Orientadora: Profa. Dra. Suely Dadalti Fragoso

Linha de Pesquisa: Redes Sociais, Interações e Sociabilidades.

PORTO ALEGRE

2020

Luiza Carolina dos Santos

**MÁQUINAS QUE FALAM (E ESCUTAM):  
as formas de agência e de interação das/com as assistentes pessoais digitais**

Tese de doutorado apresentada à banca examinadora do Programa de Pós-Graduação em Comunicação da Universidade Federal do Rio Grande do Sul (PPGCOM-UFRGS) como requisito parcial para obtenção do título de doutor em Comunicação.

---

Prof. Dr. Suely Dadalti Fragoso – UFRGS (orientador)

---

Prof. Dr. Eduardo Campos Pellanda – PUCRS

---

Prof. Dr. Gustavo Daudt Fischer – UNISINOS

---

Prof. Dr. Tarcísio de Sá Cardoso – UFBA

---

Prof. Dr. Willian Fernandes Araújo – UNISC

---

Prof. Dr. Marcelo Ruschel Träsel – UFRGS (suplente)

## AGRADECIMENTOS

O texto final desta tese carrega consigo a marca de muitos encontros – alguns acadêmicos, outros não. Todos eles contribuíram, de formas distintas, para as tantas linhas aqui escritas. Correndo sempre o risco de cometer uma injustiça, agradeço:

À Profa. Suely Fragoso, pela orientação cuidadosa neste trabalho e pelos ensinamentos, da vida e da academia, durante os anos de convívio. E, por mais do que neste projeto, acreditar (sempre) em mim.

Aos colegas do LAD pelo espaço de pesquisa, de troca, de risadas, de compreensão e de carinho: André, Dennis, Ana, Alexandre, Breno, Mariana, Mayara, Natan, Fabiana e Samyr. As colegas do PPGCOM que foram minhas companheiras nesta jornada e tornaram este caminho um pouquinho mais leve: Marília, Camila, Nathalia, Laura, Ludimila M., Anelise Dias, Anelise De Carli e Ludmila.

Aos professores do Programa de Pós-Graduação da Universidade Federal do Rio Grande do Sul por fornecerem um ambiente sólido para o meu desenvolvimento e pelas tantas oportunidades de diálogo e de aprendizagem. Aos funcionários do PPGCOM/UFRGS, pela paciência e diligência. Aos alunos que tive o prazer de ensinar e que fizeram de mim uma docente melhor.

Ao Prof. Florian Sprenger e aos colegas do grupo *Konfigurations of Film*, por me receberem como um dos seus durante meu período de estágio doutoral na Goethe Universität, Alemanha, e pelas oportunidades de troca que me forneceram. Às minhas *roomates* alemãs, Sarah e Sara, pelo aconchego do lar em terras estrangeiras; à Juliana e Borissa pela amizade, leveza e apoio.

À Profa. Beatrice Fazi e aos colegas do Sussex Humanities Lab, pela recepção calorosa e pelos diálogos intensos e produtivos durante meu período como aluna visitante na Universidade de Sussex, Inglaterra. Ao Daniel e Bárbara, pelo lar temporário; à Manoela e Elif pelo companheirismo.

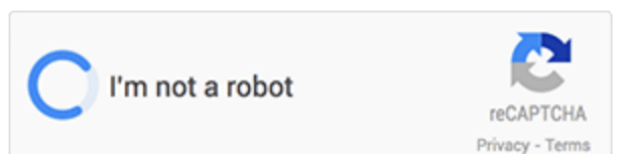
Às minhas sempre amigas, que fazem parte da minha formação humana muito antes deste doutoramento, mesmo com a infeliz distância geográfica: Anna, Tatiana, Camyla, Carmen, Bárbara, Luciele, Letícia e Amanda. À Nicole, pelos encontros no Alcides sempre que preciso; Gabrielli Tiburi pelos drinks e torcida ilimitada; e a Giulia, pelos áudios intermináveis e pela presença irrestrita nos momentos de crise – e por nunca deixar que a solidão se instalasse.

Aos professores que compuseram as bancas de qualificação e de defesa: Nísia M. Rosário, Eduardo Pellanda, Gustavo Fischer, Willian Araújo e Tarcísio Cardoso. E a Gabrielli Tiburi, Marília Gehrke e Marcio Telles, pelas preciosas observações sobre o formulário quantitativo desta pesquisa.

Aos meus pais, Dóris e Francisco: pela solidez e suavidade que possibilitaram todo o meu desenvolvimento. Ao meu irmão, também Francisco, por ter dividido comigo muito mais do que a infância. Aos meus sogros, Vera e Marco Aurélio, pelo carinho constante.

Ao Marcio Telles, por sempre me escutar com o coração. Por compreender as ausências e impossibilidades, pelos chás, cafés e massagens, pela tolerância e disponibilidade afetiva durante os últimos quatro anos - e especialmente nos meses que antecederam a conclusão deste trabalho. E por topar construir um futuro conjunto.

Por fim, agradeço à CAPES pela bolsa concedida durante os quatro anos deste doutorado, possibilitando minha dedicação integral a esta pesquisa, e pela bolsa de Doutorado Sanduíche – PDSE, que viabilizou minha produtiva estadia na Alemanha e na Inglaterra.



*(Invisible reCaptcha, do Google)*

## RESUMO

Essa tese se foca na relação entre o design e a forma de funcionamento de artefatos baseados em inteligência artificial e as formas de interação que são estabelecidas pelos usuários. Através do caso específico das assistentes pessoais digitais, buscamos compreender as possíveis decorrências sociais, culturais e políticas da conformação específica destes objetos, das formas de agência emergentes e das interações e percepções dos usuários. A base teórica utilizada parte da materialidades da comunicação (GUMBRECHT, 2010; KITTLER, 2017), especificidades das tecnologias digitais (RIEDER, 2018, WEIZENBAUM, 1966, FAZI E PARISI, 2014) formas de agência dos objetos (LAGNDOM, 1980; LATOUR, 1992) e efeitos das tecnologias no mundo e nas formas de interação que os usuários estabelecem (TURKLE, 2005; SWEENEY, 2013; WALSH, 2015, entre outros) O problema de pesquisa que esta tese responde é: Como podemos compreender as formas de agência dos sistemas de inteligência artificial baseado em interação por linguagem natural? Para responder esta pergunta, utiliza-se três camadas analíticas: uma camada técnica, uma científica e outra social. A camada técnica se preocupa com a forma como os objetos são programados e de que forma funcionam; a científica propõe perceber como os discursos e os fazeres científicos do âmbito da inteligência artificial e do design determinam certos direcionamentos para os artefatos desenvolvidos; e a camada social analisa elementos e estruturas culturais que são replicados nas tecnologias, assim como formas de relação e interação entre pessoas e estes dispositivos. Os passos metodológicos adotados para operacionalizar essa investigação são: a) pesquisa bibliográfica no âmbito da ciência da computação, design e ciências humanas e sociais; b) leitura e análise de material documental sobre o uso e recepção das assistentes pessoais digitais; c) auto-observação da utilização das assistentes pessoais digitais, inspirada em métodos etnográficos; d) coleta de dados quantitativos através de um formulário online em torno dos usos e percepções dos usuários; e) coleta de dados qualitativos através de entrevistas semi-estruturadas em profundidade sobre usos e percepções dos sujeitos. Os dados obtidos em todas as etapas mencionadas foram analisados a partir de cinco fatores: a lógica computacional; a interface e as possibilidades de interação; o antropomorfismo dos artefatos; a voz e a fala como forma primordial de interação; o gênero dos objetos e as relações sociais. Destacamos como resultados obtidos: as formas de relacionamento entre humanos e máquinas humanoides ainda está sendo negociada e depende das condições específicas e dos contextos nas quais é utilizada. As assistentes se constituem como artefatos ficcionais, que propõe uma aproximação e um afastamento do humano, e, com isso, geram duas formas de interação particulares com os usuários: a brincadeira ficcional e o teste sistemático das possibilidades do sistema. A utilização da fala como forma de interação passa por técnicas e fórmulas dos usuários para se fazerem compreender adequadamente, não sendo compreendida por estes como mais natural e sendo fonte do sentimento de estranhamento. As relações com as assistentes são cortadas por características de gênero e personalidade, com percepções variadas dos usuários sobre os papéis que estes fatores operam na interação.

**Palavras-Chave:** Inteligência Artificial; Assistentes pessoais digitais; Agência; Interação por voz; Gênero.



## ABSTRACT

This thesis focuses on the relationship between design and how artefacts based on artificial intelligence work and the forms of interaction that are established by users. Through the specific case of digital assistants, we seek to understand the possible social, cultural and political consequences of the specific conformation of these objects, the forms of agency and the interactions and perceptions of users. Theoretical approach used starts from the materialities of communication (GUMBRECHT, 2010; KITTLER, 2017), specificities of digital technologies (RIEDER, 2018, WEIZENBAUM, 1966, FAZI AND PARISI, 2014) forms of agency of objects (LAGNDOM, 1980; LATOUR, 1992) and effects of technologies in the world (TURKLE, 2005; SWEENEY, 2013; WALSH, 2015, among others) The research problem is: How can we understand the forms of agency of artificial intelligence systems based on natural language interaction? To answer this question, we used three analytical layers: a technical, a scientific and a social layer. The technical layer is concerned with how objects are programmed and how they work; the scientific proposes to perceive how the speeches and scientific actions in the realm of artificial intelligence and design determine specific directions for the developed artefacts; and the social layer analyzes the replication of elements and cultural structures in technologies, as well as forms of relationship and interaction between people and these devices. The methodological steps adopted are: a) bibliographic research in the scope of computer science, design and human and social sciences; b) reading and analyzing documentary material on the use and reception of personal digital assistants; c) self-observation of use of personal digital assistants, inspired by ethnographic methods; d) collection of quantitative data through an online form about the uses and perceptions of users; e) collection of qualitative data through in-depth semi-structured interviews about subjects' uses and perceptions. The data obtained in all the mentioned steps were analyzed based on five factors: computational logic; the interface and the possibilities for interaction; the anthropomorphism of artefacts; voice and speech as a primary form of interaction; the gender of objects and social relations. We highlight as results obtained: the forms of relationship between humans and humanoid machines are still being negotiated and depend on the specific conditions and contexts. The assistants its use are constituted as fictional artefacts, which proposes an approach and a departure from the human, and, with this, generate two forms of a particular interaction with users: fictional play and the systematic test of the system's possibilities. The use of speech as a form of interaction involves techniques and formulas of users to make themselves understood correctly, not being understood by them as more natural and is the source of the feeling of strangeness. Relationships with assistants are severed by gender and personality characteristics, with varying perceptions from users about the roles that these factors play in the interaction.

**Keywords:** Artificial Intelligence. Virtual Assistants. Agency. Voice interaction. Gender.



## LISTA DE FIGURAS

Figura 1 O que é um carro para o deep learning?.....	66
Figura 2 Linha do tempo das assistentes pessoais digitais .....	74
Figura 3 Imagem da Siri no iPhone .....	75
Figura 4 Imagem da Cortana no Celular .....	77
Figura 5 Imagem da Cortana no Celular .....	78
Figura 6 Imagem da Google Home .....	79
Figura 7 Imagem da Echo.....	80
Figura 8 Gráfico Idade e Gênero dos Participantes.....	90
Figura 9 Gráfico Grau de Escolaridade dos Participantes.....	91
Figura 10 Gráfico Áreas do Conhecimento de Formação dos Participantes.....	92
Figura 11 Gráfico Faixa etária e gênero dos entrevistados .....	100
Figura 12 Gráfico Grau de Escolaridade dos entrevistados .....	101
Figura 13 Gráfico Área de Formação dos entrevistados .....	102
Figura 14 Gráfico Assistentes utilizadas pelos entrevistados.....	103
Figura 15 Gráfico Número de assistentes utilizadas .....	104
Figura 16 Gráfico Frequência de uso dos entrevistados.....	104
Figura 17 Gráfico Assistente de preferência dos entrevistados.....	105
Figura 18 Imagem da tela de conversação do autor com versão da ELIZA disponível online .....	108
Figura 19 Gráfico Pensa em sua assistente como um computador?.....	114
Figura 20 Gráfico Você se preocupa com o modo de funcionamento?.....	115
Figura 21 Gráfico Você se preocupa com privacidade?.....	130
Figura 22 Jibo, o robô social .....	154
Figura 23 Gráfico Assistentes utilizadas pelos Participantes .....	158
Figura 24 Gráfico Quantidade de Assistentes utilizadas .....	159
Figura 25 Gráfico Frequência de uso dos participantes .....	160
Figura 26 Gráfico Comparação da frequência de uso .....	161
Figura 27 Gráfico Assistentes de preferência dos participantes.....	162
Figura 28 Gráfico Hardwares utilizados para acesso .....	163
Figura 29 Gráfico Locais de utilização das assistentes .....	164
Figura 30 Gráfico Locais de utilização das assistentes .....	165
Figura 31 Gráfico Locais de utilização das assistentes .....	165
Figura 32 Gráfico Atributos que motivam o uso.....	166

Figura 33 Gráfico Percepção sobre interface de voz.....	167
Figura 34 Gráfico Assistentes utilizadas pelos entrevistados.....	168
Figura 35 Gráfico Número de assistentes utilizadas .....	169
Figura 36 Gráfico Frequência de uso dos entrevistados.....	170
Figura 37 Gráfico Assistente de preferência dos entrevistados.....	171
Figura 38 Gráfico Atividades realizadas pelos entrevistados.....	172
Figura 39 Filmes associados com o futuro da inteligência artificial .....	203
Figura 40 Printscreen Sentimentos da Google Assistente .....	210
Figura 41 Printscreen Google Assistente Personalidade .....	211
Figura 42 Nuvem de Palavras da Personalidade da Cortana.....	216
Figura 43 Gráfico Você já teve algum sentimento? .....	236
Figura 44 Gráfico Sentimentos associados às Assistentes .....	237
Figura 45 Gráfico Sentimentos Positivos X Assistentes .....	239
Figura 46 Gráfico Sentimentos Negativos X Assistentes.....	240
Figura 47 Gráfico Você desenvolveu uma relação?.....	241
Figura 48 Gráfico Relação X Frequência de Uso.....	242
Figura 49 Gráfico Relação X Frequência de Uso.....	243
Figura 50 Gráfico Sentimentos de quem não desenvolveu uma relação.....	244
Figura 51 Gráfico Sentimentos de quem desenvolveu uma relação.....	245
Figura 52 Gráfico Assistente já ligou sem querer?.....	257
Figura 53 Anúncio da Bebê Cochichinho.....	272
Figura 54 Gráfico Sensação ao falar com um objeto.....	292
Figura 55 Gráfico Baixa frequência X sensação ao falar .....	294
Figura 56 Gráfico Média frequência X sensação ao falar .....	295
Figura 57 Gráfico Alta frequência X sensação ao falar.....	295
Figura 58 Gráfico Fala diferente com as assistentes?.....	303
Figura 59 Gráfico Como fala com as assistentes?.....	304
Figura 62 Gráfico de preferência do gênero da voz .....	337
Figura 63 Gráfico de preferência de voz dividido pelo gênero dos respondentes.....	338
Figura 64 Gráfico Motivos para preferência da voz feminina.....	340
Figura 65 Gráfico Motivos para preferência de voz neutra.....	342
Figura 66 Gráfico Motivos para preferência de voz neutra.....	343
Figura 67 Gráfico Relação entre o gênero da voz e a interação .....	344
Figura 68 Gráfico Gênero da voz X Interação – Homens e Mulheres .....	345

## LISTA DE TABELAS

Tabela 1 Distribuição dos Participantes por áreas e subáreas.....	92
Tabela 2 Locais e Formato das Entrevistas.....	96
Tabela 3 Quadro comparativo de respostas das assistentes pessoais digitais.....	206
Tabela 4 Assistentes pessoais digitais e o gênero da voz.....	318
Tabela 5 Gráfico de respostas das assistentes pessoais virtuais para insultos verbais.....	322

## SUMÁRIO

INTRODUÇÃO.....	15
1. CONSTRUÇÃO DO QUADRO TEÓRICO.....	22
1.1 MATERIALIDADES DA COMUNICAÇÃO .....	22
1.2 NÃO EXISTE SOFTWARE.....	25
1.3 O SOFTWARE COMO MATERIALIDADE.....	32
1.4 ALGORITMOS COMPUTACIONAIS.....	38
1.5 BANCO DE DADOS E O ARQUIVO.....	47
1.6 PODEM AS MÁQUINAS PENSAR?.....	51
1.7 INTELIGÊNCIA ARTIFICIAL: PROBLEMATIZAÇÃO E CONCEITOS.....	57
1.8 AS POLÍTICAS DOS ARTEFATOS.....	66
2. DESENHO METODOLÓGICO .....	73
2.1 CONHECENDO AS ASSISTENTES PESSOAIS DIGITAIS .....	73
2.1.1 E AÍ, SIRI.....	74
2.1.2 EI, CORTANA.....	76
2.1.3 OK, GOOGLE.....	78
2.1.4 ALEXA .....	80
2.2 DAS ESCOLHAS METODOLÓGICAS .....	81
2.2.1 SOBRE O FORMULÁRIO PARA OBTENÇÃO DE DADOS QUANTITATIVOS .....	86
2.2.2 DO PERFIL DOS RESPONDENTES DO FORMULÁRIO ONLINE.....	89
2.2.3 SOBRE A ENTREVISTA EM PROFUNDIDADE .....	95
2.2.4 DO PERFIL DOS ENTREVISTADOS .....	99
3 AS LÓGICAS DA COMPUTABILIDADE.....	106
3.1 COMO FUNCIONAM AS MÁQUINAS PARA CONVERSAÇÃO.....	107
3.1.1 ANÁLISE DOS DADOS QUANTITATIVOS SOBRE O MODO DE FUNCIONAMENTO DAS ASSISTENTES .....	114

3.2 LIMITES E CRÍTICAS DA INTELIGÊNCIA ARTIFICIAL .....	116
3.3 UMA INTELIGÊNCIA NÃO-HUMANA? .....	121
3.4 QUESTÕES ÉTICAS NA RELAÇÃO HUMANO-MÁQUINA.....	126
4 OBJETOS INTERACIONAIS .....	133
4.1 INTERFACE.....	133
4.2 INTERAÇÃO .....	139
4.3 OBJETOS INTERACIONAIS.....	142
4.4 A MÁQUINA COMO OUTRO?.....	151
4.5 USOS, PERCEPÇÕES E PREFERÊNCIAS DOS RESPONDENTES DO FORMULÁRIO ONLINE .....	157
4.6 TODO DIA ELA FAZ TUDO SEMPRE IGUAL? .....	171
4.7 PERCEPÇÃO DA INTERFACE DE VOZ.....	174
4.8 PROBLEMAS NA INTERAÇÃO COM AS ASSISTENTES.....	182
4.9 FORMAS DE INTERAÇÃO COM AS ASSISTENTES.....	185
4.9.1 JOGANDO COM AS ASSISTENTES: ENTRE A BRINCADEIRA E O TESTE .....	186
4.9.2 O CARÁTER DE SOCIABILIDADE NA BRINCADEIRA COM AS ASSISTENTES .....	192
4.10 SOMOS EDUCADOS COM NOSSOS ARTEFATOS DIGITAIS? .....	195
5 QUASE-HUMANAS? OS SENTIMENTOS PARA COM O ESTRANHO FAMILIAR .	199
5.1 AGENTES E ROBÔS (DE FIGURAÇÃO HUMANOIDE).....	200
5.2 AS ASSISTENTES PESSOAIS DIGITAIS ENQUANTO PERSONAGENS FICCIONAIS .....	206
5.3 UNCANNY VALLEY.....	223
5.4 QUEM SABE O QUE É UMA ASSISTENTE PESSOAL DIGITAL?.....	229
5.5 ROBÔ É SÓ AQUELE DO DESENHO DOS JETSONS?.....	232
5.6 UNCANNY FEELINGS.....	235
5.7 ASSISTENTES VIRTUAIS TAMBÉM TÊM SENTIMENTOS?.....	246

5.8 QUEM É A SUA ASSISTENTE PESSOAL?.....	250
5.9 O ACIONAMENTO ACIDENTAL: ENTRE O ESTRANHAMENTO E O HUMOR .....	256
5.10 ANTROPOMORFIZANDO AS ASSISTENTES.....	260
6 QUEM [O QUE] FALA? .....	264
6.1. A VOZ, A FALA E A LINGUAGEM ENQUANTO CONSTITUTIVA DOS SUJEITOS .....	266
6.2 AS SONORIDADES DA VOZ E DA FALA.....	276
6.3 PODEM AS MÁQUINAS FALAR?.....	283
6.4 FALANDO COM “UM PEDAÇO DE SILÍCIO”.....	291
6.5 RELAÇÕES ENTRE PÚBLICO E PRIVADO (OU, AS ASSISTENTES ESTÃO NO ARMÁRIO?).....	299
6.6 COMO VOCÊ FALA COM A SUA ASSISTENTE?.....	302
6.7 A VOZ E A FALA DAS ASSISTENTES.....	310
7 OBJETOS DIGITAIS TÊM GÊNERO? .....	316
7.1 ‘EU SOU DO [GÊNERO] FEMININO EM PERSONALIDADE’ .....	318
7.2 (AUSÊNCIA DE) MULHERES NO ÂMBITO DAS TECNOLOGIAS DE INFORMAÇÃO E COMUNICAÇÃO .....	326
7.3 ‘LUGAR DE MULHER É NA ...’: SOBRE A DIVISÃO SEXUAL DO TRABALHO .....	329
7.4 QUAL É A VOZ DA SUA ASSISTENTE?.....	336
7.4.1 PREFERÊNCIAS E PERCEPÇÕES DOS RESPONDENTES DO FORMULÁRIO ONLINE SOBRE O GÊNERO DA VOZ.....	336
7.4.2 PREFERÊNCIAS E PERCEPÇÕES DOS ENTREVISTADOS SOBRE O GÊNERO DA VOZ.....	345
7.5 ASSISTENTES PESSOAIS DIGITAIS TÊM... SEXUALIDADE?.....	357
7.6 POR QUE AS ASSISTENTES PESSOAIS DIGITAIS SÃO MULHERES?.....	362
CONSIDERAÇÕES FINAIS .....	366
REFERÊNCIAS .....	372

ANEXO 1 – FORMULÁRIO QUANTITATIVO.....	389
ANEXO 2 – TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO DO FORMULÁRIO.....	399
ANEXO 3 – DADOS BRUTOS DO FORMULÁRIO .....	401
ANEXO 4 – MODELO DE CONVITE POR E-MAIL E/OU REDES SOCIAIS DIGITAIS .....	402
ANEXO 5 – TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO DA ENTREVISTA .....	403
ANEXO 6 – ROTEIRO SEMI-ESTRUTURADO PARA A ENTREVISTA .....	406
ANEXO 7 – MODELO DA FICHA DE DECUPAGEM.....	409

## INTRODUÇÃO

O exercício imaginativo do futuro pós-*virada do milênio* falhou em nos fornecer uma ideia adequada da realidade que nos aguardava: não temos carros voadores, exercícios físicos sem esforço ou uma empregada doméstica em forma robótica, como antecipava o desenho animado “Os Jetsons”<sup>1</sup>. Nossas tecnologias contemporâneas seguiram um rumo mais discreto, alterando o tecido social de forma gradativa – e, talvez, com modos de funcionamentos menos evidentes. Seguimos nos locomovendo pelas ruas e não pelos ares, mas nossos motoristas podem ser máquinas – a própria máquina do carro, com os carros autônomos. Nossos telefones podem falar conosco, nossas bonecas podem falar com nossas crianças. Pouco a pouco, estamos dotando de animação não seres que reconheceríamos imediatamente como robôs, de forma figurativa humana: animamos nossos objetos em si, sem a necessidade de um terceiro ente autônomo animado que manipule esses objetos – são os próprios objetos que podem se manipular. Essa não foi a narrativa prevista na ficção científica tradicional, ainda que objetos dotados de vida sejam um conto – e um pequeno medo – tão antigo quanto o tempo.

É nesse campo curioso, onde surgem objetos digitais capazes de formas de ação autônomas, que essa investigação se instala. São objetos pensados, planejados, desenhados, programados e desenvolvidos, posteriormente postos no mundo para serem utilizados – de formas previstas e imprevisas. As assistentes pessoais digitais, dispositivos de inteligência artificial baseada em linguagem natural falada, podem estar presentes em nossos computadores, dispositivos móveis e casas, como a *Siri* (da Apple), a *Google Assistente* (do Google), a *Cortana* (da Microsoft) e a *Alexa* (da Amazon). São interfaces comunicacionais baseadas na interação por voz, capazes de executar tarefas associadas a nossas contas nestas empresas, tocar música, realizar buscas e conceder informações. Para além da realização de tarefas, as assistentes pessoais também se configuram como objetos com os quais dialogamos, algumas vezes sem uma funcionalidade necessária: elas incorporam personagens, nos respondem de formas imprevisas, engraçadas ou ousadas; Os próprios diálogos estabelecidos propõem um modo de ocupar o mundo e de se relacionar com humanos: não importa o que ouvimos, “assistentes virtuais também tem sentimentos”, informa *Siri*. Mas, no que implica dizer que objetos tem sentimentos?

Estas assistentes configuram nesse trabalho o objeto empírico de investigação, a forma

---

<sup>1</sup> Série animada de televisão criada durante a década de 60, que apresentava uma família vivendo no contexto tecnológico do futuro.



pela qual se torna possível a materialização de nossas inquietações. Tais objetos operam tanto como mediadores em processos comunicacionais (fazendo ligações, buscas na internet etc), quanto como objeto de interação direta com humanos (estabelecendo diálogos, por exemplo). Ao mesmo tempo em que partem de uma programação específica, também se alteram na interação com os humanos, através da coleta de dados que serão, com o passar do tempo, transformados em padrões (de comportamento, preferências etc) – desde a perspectiva específica do *machine learning*. Como sistema, seus modos de funcionamento são opacos para os usuários, mas sua constituição traz em si operações específicas, de resultados nem sempre previsíveis, sendo compostas por camadas que são tanto políticas quanto sociais. As nuances destas questões é que são, muitas vezes invisíveis, outras, incompreensíveis. Não deixam de ser, entretanto, objetos autônomos que habitam o mundo entre seres biológicos.

Para a comunicação, o campo da inteligência artificial representa um entrelaçamento interessante através, principalmente, dos robôs de conversação, dos quais derivam as assistentes pessoais digitais. Esse entrelaçamento entre comunicação e inteligência artificial está posto desde o princípio do campo (GUNKEL, 2012; 2017), quando Alan Turing associa a possibilidade de existência de uma máquina inteligente à simulação do comportamento humano, a ser avaliada através do que ele denomina o jogo da imitação (TURING, 1950), cuja questão crucial para a comunicação está no fato de que tal teste deve se dar a partir da interação conversacional entre um juiz, um agente humano e um agente de Inteligência Artificial, pela avaliação das habilidade comunicacionais dos concorrentes.

Observamos, ainda, um avanço crescente nos desenvolvimentos do campo da Inteligência e da Robótica contemporaneamente, relacionados com quatro mudanças exponenciais: poder de processamento; quantidades de dados disponíveis online; métodos algorítmicos como o deep learning; financiamento. O investimento em startups de Inteligência Artificial tem duplicado a cada dois anos (nos últimos cinco anos, aproximadamente), e as grandes empresas do Vale do Silício (*Apple, Microsoft, Google, Amazon*) também vem aumentando a compra de patentes no setor (WALSH, 2015).

Como avanços consideráveis no âmbito da Inteligência Artificial podemos citar os seguintes setores: a transcrição, a capacidade computacional de transcrição do mandarim falado é atualmente mais precisa que a humana; questões de guerra, com computadores capazes de pilotar veículos aéreos para combate mais precisamente que humanos; e diagnóstico médico, com programas capazes de diagnosticar doenças pulmonares com maior precisão que médicos (WALSH, 2015). Ainda que precisão seja a palavra-chave aqui, vale ressaltar que, especialmente nos últimos dois casos, erros de ordem maquínica acarretariam efeitos

potencialmente catastróficos dado seu campo de atuação.

Os avanços na área e a capacidade de tornar parte desses avanços em produtos lucrativos, traz uma preocupação crescente com as posições ocupadas (ou que podem vir a ser ocupadas) por objetos de inteligência artificial no âmbito social. A preocupação com questões relacionadas à ética e responsabilização podem ser verificadas a partir de propostas legislativas sobre o tema como o relatório para regulação da criação e uso de robôs e inteligência artificial europeu<sup>2</sup> (em fevereiro de 2017), o Tratado de Toronto (maio de 2018)<sup>3</sup> e o General Data Protection Regulation<sup>4</sup> (em maio de 2018). Essas iniciativas vêm sendo capitaneadas por países desenvolvidos, principalmente da Europa e da América do Norte, que tanto produzem quanto utilizam artefatos baseados em inteligência artificial.

Em uma pesquisa realizada mundialmente<sup>5</sup> pela Mozilla Foundation sobre a percepção dos sujeitos em torno da inteligência artificial, sentimentos diferentes aparecem sobre a questão para indivíduos de regiões geográficas com uma maior quantidade de países com níveis mais baixos de IDH em comparação com regiões geográficas com níveis mais elevados de IDH (MOZILLA FOUNDATION, 2019). Ao serem perguntados sobre como se sentiam em relação à inteligência artificial, respondentes da América do Sul, Ásia e África escolheram em sua maioria dois sentimentos positivos (*excitado*, *curioso* ou *esperançoso*). Em contrapartida, Europa, Oceania e América do Norte escolheram em sua maioria um sentimento negativo (*preocupação*) e um positivo (*esperançoso* ou *curioso*<sup>6</sup>). Essa diferença aponta para a necessidade de estudos específicos em regiões diversas do globo capazes de entender qualitativamente as formas de percepção e utilização dos usuários de sistemas inteligentes, assim como os efeitos nas estruturas sociais. No caso específico da América Latina, interessa o desenvolvimento de pesquisas que possam avaliar o impacto e as particularidades da adesão de artefatos de inteligência artificial em um local historicamente marcada por desigualdades sociais tanto internamente quanto no desenho geopolítico mundial.

Além disso, muitas das tecnologias que envolvem inteligência artificial são produzidos em um contexto local e cultural, sendo posteriormente utilizadas em outros (geográficos,

---

<sup>2</sup> Primeira tentativa de legislação sobre o assunto, elaborado por Mady Delvaux.

<sup>3</sup> Tratado que propõe parâmetros para responsabilização, inclusão e não-discriminação em sistemas autônomos. Disponível em: <https://www.accessnow.org/the-toronto-declaration-protecting-the-rights-to-equality-and-non-discrimination-in-machine-learning-systems/>

<sup>4</sup> Lei europeia que determina a transparência no processamento de dados pessoais e o direito do indivíduo de não ter decisões tomadas sobre si de forma totalmente mecanizada, caso assim desejar.

<sup>5</sup> Devemos notar, entretanto, que ainda que tenha sido aplicada e recebido retorno de moradores de todos os continentes, apenas 7% dos respondentes eram sul americanos, comparado a um retorno de 42% de europeus e 39% de moradores da América do Norte.

<sup>6</sup> Note aqui a ausência do retorno “excitado” como um dos principais para esses países.

culturais, políticos), como é o caso das assistentes pessoais digitais. Em pesquisa de mercado realizada em 2017 pela agência Ovum<sup>7</sup>, dos Estados Unidos, estima-se que a partir de 2021 existirão mais assistentes pessoais digitais como a *Alexa* do que humanos no planeta. Tal proporção, claro, não será igualmente distribuída entre as nações mundiais, nem entre as diferentes classes sociais de cada uma delas. Esse tipo de dispositivo também vem sendo utilizado como tecnologia assistiva para usuários com deficiência visual<sup>8</sup>, como ferramenta de facilitação da socialização de crianças autistas<sup>9</sup> e como forma de acesso facilitada para idosos e pessoas que não dominam a língua escrita. Todos esses elementos apontam para um cenário comunicacional contemporâneo cada vez mais complexo, que demanda a reflexão em torno das formas interacionais destes sistemas no cotidiano social.

Essa investigação se foca na relação entre o design e a forma de funcionamento de artefatos baseados em inteligência artificial e as formas de interação que são estabelecidas pelos usuários. Através do caso específico das assistentes pessoais digitais, buscamos compreender as possíveis decorrências sociais, culturais e políticas da conformação específica destes objetos, das formas de agência emergentes e das interações e percepções dos usuários. O processo investigativo se baseia, por um lado, na análise das técnicas, lógicas e decisões de design envolvidas no funcionamento desses sistemas; por outro, na observação e discussão dos efeitos interacionais e sociais destas tecnologias.

Para o desenvolvimento deste trabalho, partimos dos seguintes pressupostos teóricos: a) objetos digitais com base em inteligência artificial interagem em nosso tecido social de formas cada vez mais cotidianas (BASSET, 2018), gerando um imbricamento entre as técnicas das tecnologias e as práticas sociais (RIEDER, 2018); b) a forma de funcionamento e a materialidade de nossos objetos digitais compõem formas de sentido não-hermenêuticas (GUMBRECHT, 2010; KITTLER, 2017); c) objetos digitais baseados em Inteligência Artificial geram efeitos materiais no mundo (FAZI E PARISI, 2014), através de seus modo de funcionamento (SPRENGER, 2015; KITTLER, 2017) e de interação com humanos (TURKLE, 2005); d) a Inteligência Artificial pode ser pensada fora de um paradigma simulativo, não atrelada à características humanas (FAZI, 2018); e) as micro-decisões que envolvem estes objetos digitais (SPRENGER, 2015), daquelas que surgem de seus modos de programação e

---

<sup>7</sup> GREEN, Penelope. *Alexa, where have you been all my life?* Publicado no New York Times em julho de 2017. Disponível em: <https://www.nytimes.com/2017/07/11/style/alexa-amazon-echo.html>.

<sup>8</sup> Idem

<sup>9</sup> CHIN, Louie. *Para Siri, com amor.* Publicado na Zero Hora em 28 de outubro de 2014.

Disponível em: <https://gauchazh.clicrbs.com.br/saude/vida/noticia/2014/10/Como-Siri-assistente-virtual-da-Apple-se-tornou-a-melhor-amiga-de-uma-crianca-autista-4630628.html>.

funcionamento do sistema até aquelas ligadas à construção de um objeto-personagem dotado de Inteligência Artificial, possuem desdobramentos sociais e políticos (SWEENEY, 2013; WALSH, 2015);

A partir destes pressupostos, desenhamos o problema de pesquisa que esta tese visa responder da seguinte maneira: **Como podemos compreender as formas de agência dos sistemas de inteligência artificial baseado em interação por linguagem natural?**

Para responder esta pergunta, propomos investigar o objeto de pesquisa partindo de três camadas analíticas: uma camada técnica, uma científica e outra social. A camada técnica foca na programação e funcionamento dos objetos e nas implicações de tais funcionalidades. A científica analisa os discursos e os fazeres científicos do âmbito da inteligência artificial e do design como determinantes para o direcionamento dos artefatos desenvolvidos, resultando em características específicas das assistentes pessoais digitais hoje. E a camada social analisa elementos e estruturas culturais que são replicados nas tecnologias, assim como formas de relação e interação entre pessoas e estes dispositivos. Ou seja, uma camada que pensa o social na tecnologia e a tecnologia no social.

As três camadas – técnica, científica e social – dizem respeito tanto às condições de emergência de determinado objeto quanto a sua própria circulação no tecido social. Ou seja, estas três camadas são tanto um emaranhado de relações complexas que fazem com que um objeto emerja de determinada maneira quanto vão sendo ressignificadas ao atuarem no mundo. Essas três camadas se sobrepõem e entrecruzam em momento distintos, uma vez que a separação completa do objeto de análise é inviável por características inerentes ao mesmo.

Enquanto esse cruzamento representa um desafio metodológico, também acreditamos que é justamente no entrecruzamento analítico entre estas camadas que reside o potencial de compreensão do objeto de pesquisa. Não apenas tal desenho nos permite um olhar múltiplo sobre nosso próprio objeto, **como também é objetivo desta proposta teórico-metodológica fazer emergir uma análise da agência destes artefatos que seja capaz de integrar os âmbitos técnicos, científicos e sociais de nossas tecnologias.**

A partir do problema de pesquisa e das camadas analíticas propostas, os passos metodológicos adotados para operacionalizar essa investigação são: a) pesquisa bibliográfica no âmbito da ciência da computação, design e ciências humanas e sociais; b) leitura e análise de material documental sobre o uso e recepção das assistentes pessoais digitais; c) auto-observação da utilização das assistentes pessoais digitais, inspirada em métodos etnográficos; d) coleta de dados quantitativos através de um formulário online em torno dos usos e percepções dos usuários; e) coleta de dados qualitativos através de entrevistas semiestruturadas em

profundidade sobre usos e percepções dos sujeitos; f) análise dos materiais coletados tendo em vista o referencial teórico.

Os dados obtidos em todas as etapas mencionadas foram analisados a partir de cinco fatores considerados centrais na forma como as assistentes agem e fazem agir e nas interações que os usuários estabelecem com as mesmas. Estes fatores são: a lógica computacional; a interface e as possibilidades de interação; o antropomorfismo dos artefatos; a voz e a fala como forma primordial de interação; o gênero dos objetos e as relações sociais.

Os capítulos deste trabalho dividem-se da seguinte forma. No primeiro capítulo, o leitor encontra uma exposição e debate das correntes teóricas e conceitos que constituem a base da construção desta tese, criando um chão comum para as análises que se seguirão. O segundo capítulo apresenta os objetos empíricos de pesquisa, *Cortana*, *Siri*, *Google Assistente* e *Alexa*, para posteriormente detalhar os passos metodológicos realizados, ponderando suas limitações. Finalizamos apresentando o perfil dos respondentes do formulário quantitativo e dos entrevistados na etapa qualitativa, criando uma baliza que deve ser medida em relação às análises seguintes.

Nos capítulos três, quatro, cinco, seis e sete apresentamos as análises a partir dos elementos percebidos como centrais para compreender a agência e as formas de interação das/com as assistentes pessoais digitais. O capítulo três, *As lógicas da computabilidade*, apresenta, a partir do caso ELIZA, as formas de funcionamento implicadas em artefatos computacionais que interagem a partir da linguagem natural, para posteriormente adentrarmos nas especificidades das técnicas envolvidas nas assistentes pessoais digitais. Discutimos questões éticas na relação humano-inteligência artificial e analisamos como os usuários percebem questões como a privacidade, o modo de funcionamento e a computabilidade das assistentes. A lógica computacional, foco deste capítulo, volta a aparecer em todos os capítulos subsequentes, justamente por se constituir como uma camada básica sobre a qual as demais camadas dessas assistentes são colocadas: por mais humanas que possam (ou não) parecer, esses artefatos são máquinas de computar e, por isso, possuem potenciais e limitações específicos.

No capítulo quatro, *O que são objetos interacionais?*, discutimos a ideia de transparência das interfaces gráficas e como a possibilidade de fazer a percepção da mediação desaparecer resulta, por um lado, em desenvolvimentos que utilizam a fala como forma de interação, partindo do pressuposto de que é mais natural e, portanto, menos perceptível. Seguimos para a discussão das formas de interação online e da interatividade, adotando o último termo como forma de nos referirmos aos potenciais específicos de nosso objeto empírico. Na sequência, mostramos como a interatividade estabelecida é de uma forma específica, que coloca

objetos em um local que antes pertencia ao humano: o do diálogo. Nesse capítulo, propomos compreender as assistentes enquanto objetos interacionais, por utilizarem a linguagem natural falada como forma de interagir, colocando, portanto, o humano, em uma posição específica em relação a elas. Por fim, analisamos dados empíricos sobre frequência, local e forma de utilização das assistentes pessoais digitais.

O capítulo cinco, *Quase-humanas? Os sentimentos para com o estranho familiar*, analisa as assistentes enquanto objetos de forma antropomórfica, para as quais são criadas personalidades: ou seja, elas habitam mundos ficcionais. Analisamos as particularidades das personalidades de *Siri*, *Alexa*, *Cortana* e *Google Assistente* e as formas de interação que os usuários estabelecem, em especial as brincadeiras e testes. Seguimos para pensar de que forma essa proximidade e afastamento com o humano que as assistentes possuem podem gerar o sentimento de *Umheimliche* (FREUD, 1919; 2019), examinando os sentimentos apontados pelos usuários nas etapas quantitativa e qualitativa da coleta de dados.

No capítulo seis, *Quem [o que] fala?*, tomamos a fala e a linguagem como elementos constitutivos do ser humano, propondo que a integração de outras coisas que falam, não-humanas, pode gerar efeitos particulares. Desconstruímos a ideia, presente no design de interface e na ciência da computação, de que a fala é natural para os seres humanos e que se fala com uma assistente digital da mesma forma que se fala com um humano. Para isso, mostramos as técnicas e as fórmulas utilizadas pelos usuários para realizar uma interação efetiva com as assistentes, uma vez que, por suas lógicas computacionais e questões de design, elas respondem apenas a formas de falar específicas.

O capítulo sete, *Objetos digitais têm gênero?*, a adoção de uma voz reconhecida como do gênero feminino como exclusiva ou padrão para as assistentes pessoais digitais. Nesta análise aparecem elementos relacionados aos baixos índices de presença de mulheres na indústria da tecnologia, ao machismo estrutural que permeia o campo da inteligência cultural e também a cultura de forma mais ampla e o conceito de cuidado como categoria de trabalho feminino. Por fim, analisamos como as empresas apresentam as *personalidades femininas* destas assistentes e como os usuários percebem essa questão e como interagem com o gênero da voz nestes artefatos.

## 1. CONSTRUÇÃO DO QUADRO TEÓRICO

### 1.1 MATERIALIDADES DA COMUNICAÇÃO

Indagações a partir do campo da ciência da informação a cerca da ontologia da informação (GONZÁLEZ DE GÓMEZ, 2009) são também importantes para a Comunicação. Ao questionar-se sobre a estrutura fundamental da informação, a autora se depara com o dilema de pressupor teoricamente uma origem material ou imaterial da informação – assumir uma ou outra posição ontológica resulta na construção de saberes específicos distintos para a ciência da informação. Tal colocação não é menos verdadeira para a Comunicação, ou campos correlatos de investigação que possam vir a se deparar com questão semelhante.

A constante referência à forma na qual a informação está contida, o meio no qual a informação é veiculada, parece pressupor que forma e conteúdo são não apenas coisas distintas como também dissociáveis entre si. Essa noção se intensifica na abordagem do conceito de documento, parte essencial do trabalho do cientista da informação (Gómez, 2009). Nesse caso, documento é o objeto no qual a informação esta contida e tal informação parece ser sempre algo que se fala sobre os fatos e objetos – me parece aqui, que fica subjacente a ideia de que informação é algo que se extrai e, posteriormente, se cataloga, resume, organiza, segmenta. O conteúdo – a informação – é, novamente, parte, tanto do objeto no qual estaria contida quanto do objeto do qual foi extraída: os objetos, em si, não constituem sentidos, apenas os contêm.

Perspectivas teóricas recentes do que se caracteriza como a virada material, propõem abordagens que possibilitam uma compreensão distinta de nossos objetos comunicacionais, focando em suas características intrínsecas, modos de funcionamento específicos ou capacidade de agir como atores em uma rede sociotécnica (LATOURET, 2012), por exemplo, tornando o olhar sobre os aparatos comunicacionais mais complexo. Ou seja, um computador não apenas media interações humanas ou um livro não apenas contém informações sobre uma história – as particularidades desses objetos podem ser componentes de sentido, possuir implicações políticas, impactar e modular a interação humana ou agir materialmente no mundo que habitamos.

As Teorias da Mídia Alemã, as Materialidades da Comunicação, os Novos Materialismos, o Realismo Especulativo e a Teoria Ator-Rede, abordagens contemporâneas de entrada relativamente recente no campo da comunicação, compõem, para Felinto (2013) uma

nova ambiência dos estudos das Ciências Humanas e Sociais. Tais ideias trariam consigo, como ponto em comum, uma busca por perspectivas “menos antropocêntricas, antitecnológicas e menos transcendentais”, semeando, portanto, transformações estruturais para as Ciências Humanas e Sociais, ainda que tais perspectivas possam ser conflitantes entre si.

A vertente que caracterizamos como virada material é composta por uma gama variada de teóricos e perspectivas de análise, tais como as Teorias das Mídias Alemãs (GUMBRECHT, 2010; KITTLER, 1996; 1999; 2013; 2016; 2017; SPRENGER, 2015; 2016), Arqueologia da Mídia (PARIKKA, 2012) e a Ontologia Orientada ao Objeto (HARMAN, 2016), por exemplo. No Brasil, a pesquisa em comunicação também vem desenvolvendo pesquisas que podem ser localizadas no contexto geral da virada material - tanto na crítica do campo e na especulação de rumos possíveis de investigação que levem em conta os meios e suas materialidades (como, por exemplo, FELINTO, 2001; 2004; 2010; 2011, GIRARDI JÚNIOR, 2016 e TELLES, 2016), quanto em experimentações empíricas que visam análises mais materialistas de nossos objetos e modos comunicacionais contemporâneos (como, por exemplo, MAIA, 2013; FALCÃO, 2014; LEMOS; PASTOR, 2014; FRAGOSO, 2016 e SÁ; POLIVANOV, 2012).

No contexto deste trabalho, nos interessa, inicialmente, os desenvolvimentos que chamamos de Teorias Alemãs das Mídias, ainda que tal rótulo seja, como apontado por Whintrop-Young (2006) uma delimitação anglo-americana e que o conjunto de teorias desenvolvidos no país em torno da mídia não possuam uma coesão que possa ser caracterizada como unidade. Nosso foco aqui se caracteriza principalmente pelo interesse em conceitos específicos da obra de Kittler (1996; 1999; 2013; 2016; 2017) e pelos desenvolvimentos recentes de Sprenger (2015; 2016).

As teorias das materialidades nos estudos alemães compõem uma vertente que, em linhas genéricas, aponta para a importância de estudos na área que levem em conta os próprios meios nos quais os conteúdos são veiculadas, considerando os efeitos e modulações gerados a partir dessa materialidade na qual o sentido está inscrito. Hans Ulrich Gumbrecht é considerado o teórico fundador desta vertente, com bases iniciais já em 1988, ano de publicação de uma coletânea de artigos na Alemanha intitulada “*Materialität der Kommunikation*”, vindo a compor um grupo que agrega outros nomes como Jeffrey Schnapp, Niklas Luhman, Friedrich Kittler e David Weebery.

Gumbrecht (2010) detecta e questiona-se acerca da supremacia do significado sobre o significante a partir de um desenvolvimento histórico desenrolado depois do século XV, que situa o sujeito como centro de uma noção hermenêutica que compreende o material apenas como meio de expressão de um sentido descolado do corpo no qual está inscrito. Tal noção se



fundamenta na ideia de que a cisão entre corpo e espírito é possível e que o sentido é exclusividade do espírito, e, conseqüentemente, do sujeito, jamais residindo no objeto.

O início do pensamento de Gumbrecht acerca das materialidades da comunicação gira em torno da perspectiva literária – entretanto, Kittler é o autor que irá levar as premissas dessa forma de pensamento para os objetos técnicos de mídia, tornando sua obra relevante para este trabalho por nos auxiliar mais adequadamente na tentativa de compreender a relação homem-máquina no âmbito da comunicação contemporânea. Para o autor:

As materialidades da comunicação são um enigma moderno - talvez o maior. Investiga-las só tem sentido desde que saibamos de duas coisas: não existe sentido - como aquele que os filósofos e hermenêuticos sempre têm procurado nas entrelinhas - sem que haja um portador físico. Tampouco existem, porém, meras materialidades como informações ou geradores de comunicação (KITTLER, 2017, p. 279).

Se não existe sentido sem materialidade, nem materialidades dotadas de uma capacidade de gerar comunicação e sentido por si mesmas, a questão que as materialidades da comunicação e da informação nos colocam é justamente de perceber o físico enquanto parte não autônoma na produção de sentido. Partimos, nesta tese, da perspectiva das materialidades da comunicação como inspiração teórica que permite fundamentar uma investigação que busca relacionar os modos de funcionamento de objetos digitais baseados em inteligência artificial - como são as assistentes pessoais digitais - com as ações que estes operam na realidade material e com as formas de interação estabelecidas entre estes e os seres humanos.

Parte do trabalho investigativo a partir de perspectivas materialistas da comunicação e da informação já se desenvolve no âmbito das tecnologias digitais, como forma de resposta à teorizações da área que apontam para uma desmaterialização destes meios. Tais perspectivas de investigação podem ser divididas (DOURISH, 2017, p. 4) em cinco frentes distintas: a) “materialidade dos bens culturais digitais”, exame específico de produtos digitais particulares; b) “materialidade transformadora das redes digitais”, trabalha na interseção entre as tecnologias de comunicação e informação e o espaço geográfico; c) “condições materiais para a produção de tecnologias da informação”, de viés Marxista de tópicos como infraestrutura, condições de trabalho e economia; d) “consequências materiais de metáforas informacionais”, centro de análise no discurso público e nas condições culturais; e) “materialidade da representação informacional”, na qual o próprio autor se localiza, na tentativa de compreender as consequências de formas particulares de práticas representacionais digitais.

No que diz respeito ao presente trabalho, localizamos nossa proposta em uma interseção com a perspectiva a) e e), uma vez que tanto propõe um estudo focado em uma

tecnologia específica, quanto busca compreender as decorrências de uma formação específica de prática representacional. Uma vez que compreendemos que este trabalho se inspira nas materialidades digitais para desenvolver seu problema de pesquisa, sem se filiar extensivamente a uma perspectiva, não nos propomos aqui a exaurir os autores e linhas existentes, apenas em delimitar os contornos dos conceitos que são relevantes para o desenvolvimento deste trabalho.

## 1.2 NÃO EXISTE SOFTWARE

Para o desenvolvimento deste trabalho, adotaremos, a princípio, as seguintes premissas da obra Kittleriana (KITTLER, 1996; 1999; 2013; 2016; 2017; HANSEN, 2013; HOLL, 2017; entre outros) para esta pesquisa: a) noção de mídia como característica circunstancial que aponta para uma determinada configuração capaz de gerar um sistema de arquivamento, transmissão e acesso de informação, não sendo, portanto, uma definição essencialista do objeto; b) o modo como nossos objetos comunicacionais foram projetados e com qual finalidade resulta em vieses específicos; c) a corporalidade: o real sempre se expressa em um corpo, seja ele orgânico ou inorgânico, ou seja, existe um nível sensorial para todos os eventos; d) a opacidade dos modos de funcionamento de nossos objetos técnicos e os processos de apagamento de nossos hardwares; e) a base matemática do computador pessoal, e, por extensão, das tecnologias digitais em geral. Focaremos nossos esforços, ao longo dessa seção, em explicitar os conceitos a serem utilizados ao longo deste trabalho, como parte integrante da obra de Kittler, sem a pretensão de realizar uma explicação exaustiva da profícua produção do autor.

Ao analisar os sistemas computacionais o teórico alemão (KITTLER, 2013; 2017) aponta para os mecanismos de encobrimento do hardware através do software, ou, de forma mais ampla, da materialidade a partir da imaterialidade. Viveríamos hoje uma implosão do hardware por uma explosão do software – o hardware se torna invisível, e o software funciona tanto como forma de mediação entre hardware e usuário quanto forma de limitação do próprio hardware. Para Kittler, é ponto central que existe uma base sensorial para todos os eventos, que podem ser humanos ou não - desta forma, podemos compreender a afirmação do autor de que “não existe software” (KITTLER, 2013; 2017), como forma de apontar que o real é inscrito, necessariamente, na corporalidade.

A centralidade do corpo como aparato no qual o real se inscreve decorre da apropriação específica feita por Kittler de autores franceses do período, como Foucault e Lacan. Assim, um

passo importante para compreender a perspectiva de Kittler é a partir da própria ideia da tríada Lacaniana (Hansen, 2012), constituída por três diferentes níveis: o real, o simbólico e o imaginário. O âmbito do Real se expressa a partir de experiências que podem ser vividas pelo ser humano, ao mesmo tempo em que tornam-se insuportáveis para este ser: corresponde aquilo que está além do que pode ser representado pelo simbólico. É da ordem do indizível, do chocante, do excesso. O simbólico é o campo no qual a vida é estruturada, onde existem as leis, as proibições, os códigos - é da ordem do significante e onde se encontra a linguagem. Já o imaginário é o âmbito no qual o ser humano projeta situações na psique, âmbito no qual, para Lacan, está situado o eu, assim como outras instâncias, como amor e ódio.

A partir da noção de Real para Lacan, como aquilo que excede, Kittler situa o corpo como lugar de excesso de inscrição sobre o significado – aqui compreendendo o corpo não apenas como o corpo humano, mas também o corpo inorgânico, o corpo máquina (Hansen, 2012). Decorre, parcialmente, desta formulação a centralidade dos processos materiais para Kittler, uma vez que o autor "construiu em cima do pensamento importado da França ao modificar um a priori histórico para um a priori tecnológico e ao se focar nas fundações materiais – e não apenas discursivas - do conhecimento, da cultura e da humanidade" (SPRENGER, 2016, p. 74, tradução nossa<sup>10</sup>).

Em sua teoria de diferenciação de mídia, Kittler (1999) situa o filme e o gramofone enquanto sistemas simbólicos distintos que capturam dados no âmbito do Real Lacaniano, considerando que um determinado domínio sensorial se refere à materialidade específica do corpo de formas distintas. É possível que essa captura através dos meios técnicos exceda aquilo que o próprio corpo humano é capaz de sentir e registrar, fazendo com que o corpo máquina seja capaz de expandir a sensibilidade através de suas próprias operações – o que não implica, de forma alguma, em uma substituição dos próprios sentidos humanos.

Tecnologias de gravação, tais como o filme e o gramofone, realizam registros que provem diretamente dos fluxos do real, sem passar por qualquer processamento simbólico humano, ou seja, funcionam de forma independente de operações de consciência ou de captura e incorporação pelo corpo humano. Assim, as máquinas, ao registrarem no campo do real coisas que nós, enquanto seres humanos, seríamos incapazes de captar, acabam por expandir o próprio domínio do sensível, tanto para os humanos quanto para as próprias máquinas. A temporalidade atrelada à existência humana, que tem lugar, assim como a própria fala, no fluxo irreversível

---

<sup>10</sup> Original: "Kittler built upon thinking imported from France by shifting from the historical a priori to the technical a priori and by focusing on the material—and not merely discursive—foundations of knowledge, culture, and humanity" (2016, p. 74).

do tempo, realiza, de certa forma, um bloqueio informacional: isso é dizer que existe um limite temporal para aquilo que podemos registrar a partir de nossa sensibilidade. Justamente esta sensibilidade humana é que pode ser expandida através dos registros maquínicos.

Vale apontar que o autor não se detém em traçar distinções entre corpo orgânico ou inorgânico, estabelecendo que humanos e máquinas são elementos que não se substituem e atuam de forma autônoma um em relação ao outro, funcionando como um conjunto que integra um sistema sensorial mais amplo, uma vez que as operações computacionais são formas de simbolização que não estão atreladas a uma cognição humana (Hansen, 2012).

Considerando o domínio computacional, nosso foco neste trabalho, o que ocorre não é uma alteração do simbólico para o real, como propõe Kittler (1999) ao falar dos registros realizados através do filme, do gramofone ou da máquina de escrever, mas sim uma alteração de um simbólico para outro. Assim, o computador realiza uma simbolização através dos números, em oposição a uma simbolização através da linguagem natural realizada pelos seres humanos, fazendo a passagem de um simbólico centrado no humano para um simbólico que é centrado na máquina. Ou seja, computadores “calculam e representam estruturas simbólicas como codificações do real.” (KITTLER, 2016, p. 328). São dois corpos, duas linguagens e duas simbolizações distintas e independentes – voltaremos a questão da possibilidade de modo de ação independente (ou ainda, da impossibilidade de modos de operação de mesma ordem) entre sistemas computacionais com técnicas de inteligência artificial e seres humanos nas seções seguintes deste capítulo.

A questão aqui é que, justamente por compreender, assim como Lacan, que a linguagem está no âmbito do simbólico, e que os próprios processos computacionais são um tipo de linguagem que não está atrelada ao humano, Kittler situa o corpo máquina não apenas como local de inscrição do real, mas também local de simbolização, de forma distinta e independente de nós seres humanos (Hansen, 2012). A autonomia da máquina computacional contemporânea esta, para Kittler (2017), justamente nessa possibilidade simbólica, não atribuída aos computadores desenvolvidos no período da Segunda Guerra Mundial, uma vez que “não possuíam quadros de referência para comunicar-se com o simbólico das línguas naturais, não conheciam os procedimentos de reconhecimento de padrões para compreender o imaginário de formas e imagens” (2017, p. 303).

Os processos realizados por computadores duram um período de tempo determinado e ocupam um espaço determinado, possuindo também limitações temporais, de forma correlata a própria experiência humana. Humanos e máquinas seriam, portanto integrantes de um registro mais amplo com experiências limitadas ao tempo, existindo de formas que não anulam um ao

outro, mas coexistem (Hansen, 2012). Assim, nosso futuro enquanto seres técnicos para Kittler não estaria centrado em uma noção de ciborgue – menos ainda em modo de funcionamento que pensa e utiliza a tecnologia como ferramenta – mas sim em uma cooperação com o tempo de inscrição computacional.

A radicalidade da materialidade proposta por Kittler está em pensar um sistema de redes discursivas (1990) emergentes em um determinado contexto históricos a partir do uso de tecnologias específicas e reside na ideia de atrelar o destino humano, cada vez mais, à materialidade dos corpos – sejam eles orgânicos ou não: "as condições discursivas e técnicas de um determinado presente influenciam o que se pode dizer de um dado passado, e essas relações de poder, que liberam o destino de uma pessoa da essência da alma, ligam o seu destino ainda mais fortemente ao corpo"<sup>11</sup> (SPRENGER, 2016, p. 72). Ou seja, as condições de possibilidade de emergência de um sistema são dadas por condicionantes tanto de ordem discursiva quanto tecnológica, ponto onde Kittler (1990) se difere, por exemplo, de sua própria base de teoria foucaultiana.

O entrelaçamento entre os desenvolvimentos de tecnologias de comunicação e informação e a própria história militar é, para Kittler (2017), questão de ordem central nos estudos das tecnologias digitais, uma vez que “a palavra e o objeto da cibernética coincidem com controle; e o controle, com o fluxo das ordens militares” (KITTLER, 2017, p. 305). Se a história das tecnologias de comunicação também é a história militar, - e, ainda, se na obra de Kittler existe uma “narrativa da história técnica como história da cultura” (GUMBRECHT, 2017) - existe um viés intrínseco de relação de domínio e poder, embutidos nos desenvolvimentos comunicacionais: informação e desinformação são armas de guerras, igualmente potentes. Da mesma forma, o próprio processo de apagamento dos modos de funcionamento das tecnologias disponíveis hoje, seja de produção privada ou governamental, exhibe certa correspondência às suas potencialidades bélicas. Os desenvolvimentos em inteligência artificial são hoje o principal campo de disputa geopolítica (HUI, 2018).

Outro ponto importante para pensarmos nosso cotidiano comunicacional, que é não apenas marcado pelo uso constate de tecnologias digitais, mas extremamente dependente de tais dispositivos, é o nível de complexidade que nossas máquinas possuem hoje e o fato de deixarmos de perceber que, na realidade, grande parte dos assuntos que nos empenhamos em debater enquanto pesquisadores – como a internet e as plataformas de mídias sociais – na

---

<sup>11</sup> No original: "The discursive and technical conditions of a given present influence what it is able to say about a given past, and these power relations, which liberate a person's destiny from the essence of the soul, bind this destiny all the more strongly to the body" (SPRENGER, 2016, p. 72).

verdade operam a partir de uma ideologia de participação que ofusca a realidade das operações maquímicas (HANSEN, 2012). Para Kittler (KITTLER e ARMITAGE, 2003), seguimos acreditando que a conexão estabelecida através da internet é uma conexão entre pessoas, enquanto na realidade é uma conexão global entre computadores.

Ainda que a colocação de Kittler figure entre a caricatura e o exagero, a decorrência deste pensamento é considerar que os modos de operação de nossas tecnologias importam não apenas para os seus desenvolvedores, mas, talvez principalmente para os seus usuários. Uma abordagem técnica das mídias e suas operações, tal como as análises do autor, levam, em seu trabalho, a pensar sobre a necessidade de “expulsar o espírito das ciências do espírito” (KITTLER, 2013), ou seja, das ciências humanas. Outro caminho possível, entretanto, a partir das mesmas constatações seria a inversão dessa lógica de pensamento, apontando para a entrada do espírito nas chamadas ‘ciências duras’. Ou seja, o domínio das técnicas de engenharia e ciências da computação apenas fazem sentido, no âmbito dos estudos de mídia e da comunicação, se voltadas a colocar questões que apenas as ‘ciências do espírito’ (em alemão traduz exatamente desta forma, sendo o termo original para as ‘ciências humanas’ na língua ‘*Geisteswissenschaften*’) seriam capazes. Propomos, portanto, nesta tese, o caminho inverso ao Kittleriano: as análises das técnicas, lógicas e modos operacionais dos sistemas comunicacionais nos auxiliam a colocar questões do âmbito da relação com os seres humanos e das implicações materiais no mundo. Não se trata, portanto, de expulsar o espírito das ciências humanas: mas sim, de aproximar as ciências humanas da ciência da computação, a fim de pensar de que forma o modo de funcionamento técnico dos sistemas transpassa questionamentos de nossa área de conhecimento.

Outra questão que perpassa este trabalho é a definição particular daquilo que caracteriza o conceito de ‘mídia’ na obra de Kittler. A definição elementar do que é uma mídia, para Kittler (2017), é tudo aquilo que é capaz de gerar um sistema de arquivamento, transmissão e processamento de informações. Tal delimitação, um tanto quanto mais ampla do que usualmente compreendemos por mídia no contexto da comunicação no Brasil – frequentemente associada aos aparatos técnicos produzidos a partir da mídia de massa -, faz com que elementos bastante distinto entre si possam ser compreendidos enquanto mídia, em um dado momento, uma vez que mídia não é um status permanente, mas justamente este estado específico do sistema.

Nesse sentido, uma cidade, o livro, a moeda, o sistema de correios e o computador podem ser compreendidos enquanto mídia. Entretanto, em termos de implementação técnica das qualidades midiáticas, Kittler (2017), pontua que:

Mas quem implementou tecnicamente essa definição pela primeira vez (ou como seu fim) na história foi a arquitetura computacional de von Neumann. Um microprocessador contém exclusivamente unidades de cálculo, de arquivo e *buses*. A unidade de cálculo executa ordens lógicas ou aritméticas em conformidade com a memória do programa; os buses transmitem ordens, endereços e dados em conformidade com a unidade de cálculo e sua última ordem; a memória, por fim, permite selecionar ou inscrever ordens ou dados entre endereços claramente identificados. (p. 244)

O computador seria, portanto, a mídia que torna esse sistema completamente técnico, por ser “essa rede de processamento, transmissão e arquivamento - isto é, de ordens, endereços e dados - é suficiente para calcular tudo que é calculável” (2017, p. 245), ou seja, todos os problemas algorítmicos que podem ser computados. A questão da computabilidade, ponto central para pensarmos as tecnologias contemporâneas e a centralidade concedida aos algoritmos, será explorada ao longo deste capítulo. Importa ressaltar, entretanto, que embora os desenvolvimentos do autor forneçam uma perspectiva analítica interessante para esta pesquisa, dado seu contexto histórico e material, sua produção não abarca teoricamente todos os nossos pontos de interesse uma vez que "para citar apenas um exemplo, os computadores descritos por Kittler não são conectados e, portanto, ficam excluídas as formas sociais em seu trabalho, falhando em fazer justiça ao entrelaçamento irreversível entre tecnologia e sociedade na mídia conectada contemporânea" (SPRENGER, 2016, p. 85, tradução nossa<sup>12</sup>).

Além das noções Kittlerianas com base nas materialidades da comunicação, desenvolvimentos mais recentes das Teorias das Mídias Alemãs, tal como a ideia de micro-decisões de ordem maquina (SPRENGER, 2015), serão importantes no desenvolvimento da perspectiva investigativa deste trabalho. O conceito de micro-decisões (SPRENGER, 2015) parte da análise da infraestrutura da internet, com foco nos *nós da rede*, para compreender como as interrupções da comunicação nesse sistema, em uma velocidade imperceptível para os humanos, resulta em decisões que adquirem um caráter legislativo, uma vez que é na construção de um sistema que define suas leis de operação, e, portanto, possuem uma política inerente. A circulação de informações pelo meio digital está associada a uma série de decisões que não são feitas por indivíduos, mas são automatizadas, funcionando a partir de um protocolo, ou seja, um conjunto fixo de regras. Também na obra de Sprenger (2015) é na materialidade das operações do sistema que residem ações que passam despercebidas, mas que marcam o campo

---

<sup>12</sup> Original: "To name just one example, the computers described by him are not connected, and thus the exclusion of the social from his work necessarily fails to do justice to the irreversible entanglement of technology and sociality in today's networked media." (SPRENGER, 2016, p. 85).

não-hermenêutico da produção de sentido.

A partir de micro-decisões que são imediatas, eficientes e determinadas, capazes de alocar tarefas e de controlar o fluxo de informação de forma automática, discute-se o contexto contemporâneo da lógica da interação entre humanos e não-humanos. Essa perspectiva não apenas recoloca o questionamento sobre o humano, mas alerta para a importância de voltar a atenção para os modos de conformação e funcionamento de objetos técnicos que são capazes de se tornar agentes tanto cotidianos quanto políticos. Esta perspectiva se mostra adequada para pensar desenvolvimentos de inteligência artificial recentes, que também carregam em si micro-decisões automatizadas com possíveis desdobramentos políticos, sociais e culturais – especialmente pela utilização de machine learning a partir de grandes bancos de dados.

Ao olhar para o desenvolvimento histórico destes dispositivos e para seu modo de funcionamento localizado, poderemos compreender melhor seus efeitos. Para além das micro-decisões de ordem mecânica, ou seja, de funcionamento dos próprios sistemas, outras escolhas podem balizar e direcionar questionamentos que, tanto refletem um modo de pensar marcado no contexto social – e, portanto, com implicações políticas – quanto agem nesse mesmo tecido social. Dentre estas, podemos marcar como exemplo de interesse desta pesquisa, a opção tanto cultural quanto empresarial de fabricar assistentes pessoais virtuais com características socialmente compreendidas como femininas. Se essa decisão não decorre de um modo de operar do sistema computacional em si, não quer dizer que não seja reflexo de um modo de operar do sistema social contemporâneo, ou mesmo desse setor da indústria em específico, conforme veremos no capítulo sete, o que expande o conceito de micro-decisões para fora do âmbito computacional.

Outros elementos também podem ser configurados como micro-decisões na construção das assistentes pessoais: escolhas que - derivadas de potenciais e limitações técnicas, de conhecimentos estabelecidos no âmbito da computação e da inteligência artificial e de questões relacionadas às configurações históricas, sociais e culturais do nosso tempo – somam-se para por fim configurar as assistentes tal como as temos hoje. Nesse sentido, os capítulos de análise (três, quatro, cinco, seis e sete), que abordam cada um dos fatores de análise, são guiados na tentativa de esmiuçar tanto como esses dispositivos se conformam no que diz respeito à interface, ao antropomorfismo, à voz e ao gênero, quanto as implicações que decorrem em formas interacionais específicas a partir destas configurações. As assistentes são um somatório de micro-decisões do sistema técnicos, do campo de produção do conhecimento científico e da cultura e sociedade em si, a partir dos quais agenciam formas de ação possíveis.



### 1.3 O SOFTWARE COMO MATERIALIDADE

Ao inscrever-se no espaço material, a palavra é capaz de armazenar o tempo através do preenchimento do espaço, permitindo, assim, a manipulação deste - enquanto a fala está inscrita em um contínuo, em um fluxo que é essencialmente irreversível no tempo (KITTLER, 1999). Não apenas a mudança de uma ordem de materialidade direta – inscrever palavras no papel, e não no corpo humano ou ainda no próprio ar, através da propagação sonora da voz -, mas a modificação na própria forma de representação e nos componentes representacionais, que irão dizer sobre o que é possível ou não representar e como, também compõe o que chamaríamos de uma materialidade da escrita. Tal colocação, veremos a seguir, depõe contra o próprio diagnóstico kittleriano de que o software não existe e que toda a materialidade das ações computacionais são localizadas no hardware – mas, por enquanto, seguimos.

Assim, a história da escrita é a própria história de seus suportes em cruzamento com os usos humanos e seus códigos em um determinado período. Os métodos, códigos e ferramentas que utilizamos para armazenar e transmitir conhecimento irão possibilitar representações e usos distintos, em contextos sociais distintos. A escrita surge a partir do desenho bidimensional elaborado em linha, proporcionando um tipo de ordenamento sequencial distinto do ordenamento em cena: inicia-se assim a narrativa. É justamente da linearidade da linha que decorre a experiência linear da vida dos homens, surgindo daí o processo histórico como um processo irrevogável. A universalização de tal experiência é possível a partir do século XIX, após a invenção da tipografia e a ascensão das publicações no Ocidente (FLUSSER, 2007). O código linearmente escrito diz também sobre como vivemos a experiência do tempo, como um fluxo contínuo e sequencialmente ordenado.

A questão da escrita não aparece como mera escolha ilustrativa que fornece indícios para a construção de uma argumentação. Aqui importa pensar que a materialidade da escrita e sua centralidade como forma de produção, representação e armazenamento de conhecimento durante séculos, desenham possibilidades e relações específicas em grupos específicos. Tal questão, evidentemente, não se resume apenas à codificação da escrita, podendo se expandir para todas as formas representacionais, tais como, por exemplo, as representações numéricas:

A representação de números, então, não é simplesmente abstrata. Como arranjos particulares de marcas em papel, a representação numeral permite formas de manipulação específicas que são significativas com ou sem nosso sistema de prática

matemática. Os aspectos materiais de como formatamos números acabam por ter propriedades essenciais para o que nós esperamos fazer com eles (DOURISH, 2017, p. 7, tradução nossa).<sup>13</sup>

A separação e distinção histórica entre números e numerais, por exemplo, vem a compor sistemas de notação distintos em diferentes culturas e períodos, estabelecendo relações localizadas com a linguagem, música e matemática, por exemplo (KITTLER, 2006). Importa, para este trabalho, as formas de representação numéricas no sistema de notação contemporâneo, assim como as técnicas matemáticas criadas para sua manipulação – em especial o algoritmo, a ser explorado no tópico seguinte.

Para Flusser, “o significado geral do mundo e da vida em si mudou sob o impacto da revolução na comunicação” (2007, p. 127) e, no centro desta mudança, estaria a própria mudança dos códigos que utilizamos como formas comunicacionais, sendo estes capazes de nos programar enquanto sociedade. Vale lembrar que computadores “calculam e representam estruturas simbólicas como codificações do real.” (KITTLER, 2016, p. 328), ou seja: seria ponto importante para Kittler, de forma similar ao proposto por Flusser, que a noção de diferenciação midiática passa também por uma distinção nos modos de codificação.

Se existem novos modos de codificar e representar a realidade, compostos por conjuntos de técnicas complexas que hoje permitem desenvolvimentos de inteligência artificial ou simulações computacionais para teste de uso de armas de guerra, quando falamos em materialidade dos sistemas de comunicação e informação não estamos falando apenas dos recursos materiais utilizados em sua construção ou das características físicas dos objetos, mas sim “daquelas propriedades de representação e formatos que restringem, permitem, limitam e moldam as formas como essas representações podem ser criadas, transmitidas, armazenadas, manipuladas e colocadas em uso” (DOURISH, 2017, p.6, tradução nossa)<sup>14</sup>.

Neste trabalho, propomos pensar a materialidade dos desenvolvimentos em inteligência artificial considerando os códigos, lógicas e formatos próprios deste tipo de tecnologia, o que inclui tanto o software quanto o algoritmo, uma vez que, tais desenvolvimentos “dependem

---

<sup>13</sup> No original: “The representation of numbers, then, is not merely abstract. As particular arrangements of marks on paper, numerical representation allow for specific kinds of manipulation that are meaningful without our system of mathematical practice. The material aspects of how we format numbers turn out to have critical properties for what we might hope to do with them” (DOURISH, 2017, p. 7).

<sup>14</sup> Original: “those properties of representations and formats that constrain, enable, limit, and shape the ways in which those representations can be created, transmitted, stored, manipulated, and put to use” (DOURISH, 2017, p. 6)

crucialmente das capacidades do sistema computacional envolvido e da adequação destes aos métodos matemáticos implementados” (DOURISH, 2017, p.2, *tradução nossa*).<sup>15</sup>

Compreender as lógicas imbricadas nos objetos digitais se justifica não apenas pelo potencial investigativo desta abordagem, mas pela ideia de controle que ainda alimentamos em relação aos sistemas computacionais. Com frequência tratamos nossos computadores e suas extensões como ferramentas sobre as quais temos controle, esquecendo que sabemos muito pouco sobre suas formas de operação:

Através da ilusão do significado e da causalidade nós não acobertamos o fato de que nós conseguimos e não podemos compreender completamente a computação, ou controla-la? Que computadores cada vez mais desenham uns aos outros e que o nosso uso é – até um certo ponto – uma súplica, uma fê cega? A nova retórica da “interatividade” ofusca mais do que revela (CHUN, 2004, p. 44, *tradução nossa*)<sup>16</sup>.

Parte desta complexidade reside justamente no que compreendemos por Software – os programas que operam em nossas máquinas e que possibilitam nosso uso. Softwares não tratam da noção abstrata de passos para a resolução de um problema, tal como os algoritmos, mas são específicos, ou seja, escritos para computadores reais (HAREL, 2000). A possibilidade de emergência do software depende de dois fatores importantes: a) compreensão de que um computador é capaz de armazenar instruções numéricas; b) que a própria máquina computacional poderia ser utilizada para realizar a tradução entre notações numéricas e simbólicas (CHUN, 2004) – aquilo que Kittler caracterizou como uma alteração de um simbólico centrado no humano para um simbólico centrado na máquina.

Ainda que possa ser tratado como uma forma imaterial, uma vez que o software, não pode, por exemplo, ser visto – tudo o que podemos ver e perceber são as ações executadas, sua materialidade se apresenta justamente nessas ações. O software “tenta mapear os efeitos materiais daquilo que é imaterial e postular o imaterial através de pistas visíveis. Através desse processo o imaterial emerge como uma commodity, como algo em seu próprio direito” (CHUN, 2004, p. 44, *tradução nossa*)<sup>17</sup>.

---

<sup>15</sup> Original: “depends crucially on the capacities of the computer system involved, and their match to the mathematical methods being implemented” (DOURISH, 2017, p.2).

<sup>16</sup> No original: Through the illusion of meaning and causality do we not cover over the fact that we do not and cannot fully understand nor control computation? That computers increasingly design each other and that our use is - to and extent - a supplication, a blind faith? The new rhetoric of “interactivity” obfuscates more than it reveals”(CHUN, 2004, p. 44).

<sup>17</sup> No original: “try to map the material effects of the immaterial and to posit the immaterial through visible cues. Through this process the immaterial emerges as a commodity, as something in its own right.” (CHUN, 2004, p. 44)

Com foco nas novas formas culturais, o que Manovich (2001) compreende como técnicas de criação de mídia de base computacional, são elencados cinco princípios das novas mídias: 1) representação numérica: composição por códigos digitais com potencial de quantificação e programação; 2) modularidade: forma estrutural de um fractal, uma vez que tanto organizam um todo, quanto são independentes; 3) automação: potencial de execução automática de algumas ações; 4) variabilidade: possibilidade de criar mídias distintas com um mesmo conjunto de dados; 5) transcodificação: processo de tradução de arquivos de um formato para outro;

O software irá se fortalecer como peça chave para Manovich (2013) alguns anos depois, com foco especial no que o autor caracteriza como softwares de media, ou softwares de criação – aqueles que possibilitam que os usuários criem, publiquem, compartilhem e etc. formas midiáticas como textos, mapas, sons, imagens. Se, para a radicalidade Kittleriana o software não passa de um modo de restrição das capacidades de ação do hardware, limitando uma máquina universal a um conjunto específico de ações, capaz de ofuscar as ações maquímicas durante este processo; Para Manovich (2013), entretanto, quando consideramos o âmbito da cultura digital, existe apenas o software. Tal colocação não aponta para uma negação do hardware ou das operações que aí se dão, mas sim para a percepção de que é através do software que se encontra a camada cultural e que se estabelecem as relações possíveis, sendo as ações performadas neste nível aquelas que caracterizam nossa tecnocultura (SHAW, 2008) e não os processos de digitalização em si. Apenas através da compreensão do software, dos seus usos e de sua linguagem é que podemos de fato desvendar as nuances da relação homem-computador.

Se o software pode ser considerado parte da materialidade de nossos objetos digitais e, portanto, cabem no escopo desta pesquisa, a programação computacional que possibilita sua existência é também elemento relevante. Para Flusser (2010), o ato de elaborar prescrições e inscrevê-las nas máquinas, ou seja, programá-las, seria uma espécie de nova modalidade de escrita que dirige-se aos aparelhos. Além de possuírem códigos próprios, nossas máquinas também operam hoje de forma a dar sentido, de decifrar o código base da linguagem digital, o código binário:

Os novos códigos são digitais – e, na verdade, em sua maioria, binários do tipo “1-0” – em virtude de dependerem do tipo de construção dos aparelhos para o qual eles são determinados, e que devem decifrar os códigos. Trata-se de aparelhos que – de maneira semelhante ao telégrafo – ou deixam passar (“1”) correntes de elétrons ou as interrompem (“0”). No fundo, os novos códigos não devem fazer nada além de atribuir um sentido a esse ligar e desligar mecânico da corrente, para codifica-lo (p. 210 – 220).

O modo de programação de um computador, entretanto, nem sempre foi automático, como o software possibilita que seja. Houve um tempo, como pontua Chun (2004), em que literalmente não existia software: os primeiros computadores eram programados manualmente, ou seja, programar consistia na tarefa humana de fazer conexões, configurar interruptores e atribuir valores de input – a chamada programação direta.

O ENIAC foi inicialmente religado para cada novo problema, então, essencialmente, um novo ENIAC era criado a cada utilização. A sua conversão para um programa de computador armazenado na máquina em 1947 (parcialmente por sugestão de John von Neumann) significava que os programas poderiam ser codificados através de mudanças de configuração, o que correspondia a sessenta instruções armazenadas, ao invés de plugar cabos (CHUN, 2004, p. 28, tradução nossa<sup>18</sup>).

A possibilidade de programação automática diminui drasticamente o tempo necessário para computação, além de tornar o próprio ato de programar mais acessível. O que as linguagens de programação fazem hoje é justamente tornar essas mudanças físicas das configurações computacionais em um conjunto de instruções que podem ser traduzidas em um conjunto de números pela máquina, em sua própria memória. Assim, Chun (2004) aponta que o conceito de software é muito mais complexo do que a ideia de instruções direcionadas para a máquina a fim da realização de uma determinada tarefa, dadas as camadas de softwares que temos sobre um mesmo centro de hardware.

Um software é, na verdade, uma versão detalhadamente escrita em linguagem computacional de um algoritmo, de forma a ser executada por um computador. Considerando justamente as camadas de software sobre um mesmo hardware, para que a leitura computacional do programa seja possível, é preciso que cada sintaxe utilizada em uma dada linguagem computacional seja descrita para o computador em termos de sentido naquele programa em específico, contendo, assim regras rígidas de prescrição de significado e de validade (HAREL, 2000; 2004).

Para os fins deste trabalho, propomos pensar o software e o hardware como engrenagens de uma mesma maquinaria, sendo sua segregação, atualmente, impossível – assim, nenhuma das partes deve ser deixada de lado para fins analíticos, sendo os próprios códigos utilizados e sua lógica de funcionamento tão importantes quanto a materialidade na qual estes se inscrevem.

---

<sup>18</sup> No original: “The ENIAC was initially rewired for each problem so that, essentially, a new ENIAC was created each time it was used. Its conversion to a stored-program computer in 1947 (in part due to a suggestion by John von Neumann) meant that programs could be coded by setting switches, which corresponded to sixty stored instructions, rather than by plugging cables.”

Assim, as assistentes pessoais digitais são tanto hardware (mais especificamente: hardwares, mas chegaremos nisto depois) quanto software e suas potencialidades dependem do entrelaçamento específico de ambos, por isso elas diferem em alguns aspectos quando acessadas pelo computador ou por smartspeakers, por exemplo.

Essa percepção passa pela própria complexidade do conceito de software hoje, uma vez que a sua compreensão passa pela própria compreensão da evolução de modos de programação mecânica para o que entendemos hoje como programação automática, ou apenas programação. Para além da simples definição de software como uma série de instruções dirigidas a um computador para que ele possa realizar uma tarefa específica, “o que nós chamamos de programação hoje surgiu do desejo de reutilizar códigos e de recrutar o computador para sua própria operação – isso é, transformar instruções singulares em uma linguagem que o computador pudesse escrever” (CHUN, 2004, p. 30). Assim, com as linguagens de programação:

o produto da programação não seria mais o funcionamento de uma máquina mas ao invés disso essa coisa chamada software – algo teoricamente (se não praticamente) iterável, possível de repetição, reutilizável, não importa que tenha escrito ou para qual máquina tenha sido destinado. Linguagens de programação inscrevem a ausência tanto do programador quanto da máquina na sua própria escrita (CHUN, 2004, p. 30).

É justamente este caráter de automação da programação que faz com que hoje a programação seja voltada à resolução de problemas e não numericamente orientada (CHUN, 2004). Se, a partir da visão de Chun, torna-se impossível segregar a máquina de seu código - ou o hardware do software -, podemos deixar de pensar as tecnologias digitais de forma dualística, passando para uma compreensão sistêmica de seu funcionamento. Não apenas não é possível separar estas duas faces das tecnologias digitais, quanto ambas compõe elementos importante do que iremos considerar aqui como as materialidades da comunicação e da informação:

Eu defendo que o software e a informação digital – os objetos ‘virtuais’ por excelência, as crianças-modelo das propagandas para o triunfo dos bits sobre os átomos – possuem uma dimensão material no processo interacional de sua construção e uso. Programadores entendem a forma como as estruturas digitais podem resistir a sua vontade, tanto quanto argila, madeira ou pedra, eles reconhecem que diferentes programas e modos de design servem melhor ou pior para diferentes plataformas, da mesma forma que esculturas podem ser acomodadas ou executadas com mais dificuldade ou facilidade em substratos diferentes. Materialidade – a natureza dos substratos e as propriedades que moldam e condicionam o encontro projetado – esta

no centro de cada experiência (DOURISH, 2017, p.6)<sup>19</sup>.

Atualmente habitamos um mundo no qual estabelecemos não apenas modos de comunicação através de nossas máquinas, mediados pelo software, mas também *com* as nossas máquinas. Entretanto, do desconhecimento das nuances das lógicas e linguagens computacionais como o grande complexo que são, decorre uma dificuldade de teorização sobre a questão. Podemos compreender, é claro, o significado do Bit, a menor partícula de informação possível, sendo transmitida a partir da corrente elétrica, sempre alterando entre ausência-presença, entre 0-1. Ainda assim, a não compreensão da lógica formal por trás dos algoritmos criados para a solução de problemas computacionais, baseados em funções matemáticas para as quais não possuímos letramento, aponta para o distanciamento pressuposto entre estas questões e o âmbito prático e teórico da comunicação. Na próxima seção iremos abordar a técnica específica dos algoritmos como uma forma de representar problemas computacionais, buscando compreender de que forma a materialidade de uma lógica algorítmica se expressa.

#### 1.4 ALGORITMOS COMPUTACIONAIS

Se “sem a descoberta das equações diferenciais parciais, nenhum chip moderno ou receptor ou transmissor de radio poderia ser concebido ou construído”<sup>20</sup> (KITTLER, 2006, p. 34, *tradução nossa*), vale, a fim de recolocar a matemática como uma camada essencial que compõe os aparatos digitais, retroceder no tempo e repensar o desenvolvimento do primeiro computador, o ENIAC (*Eletronic Numerical Integrator and Computer*), desenvolvido pela Eletronic Control Company, nos Estados Unidos, em 1946. Utilizado inicialmente em substituição ao trabalho de cálculos de balística realizados por mulheres durante a II Guerra Mundial, o nome ‘computador’ é atribuído justamente no lugar do nome da classe de trabalho que vem a substituir: o trabalho de computar números manualmente realizado pelas

---

<sup>19</sup> Original: “I argue that software and digital information - the ultimate “virtual” objects, the poster children for the triumph of bits over atoms - have a material dimension in the interactional processes of their construction and use. Programmers understand the way in which digital structures can resist their will, every bit as much as clay, wood, or stone. they recognize that different programs and program designs fit more or less easily on different platforms, just as different sculptures or carvings are more or less easily accommodated or executed in different substrates. Materiality - the nature of substrates and the properties that constrain and condition the designedly encounter - is at the core of each experience.” (DOURISH, 2017, p.6).

<sup>20</sup> No original: “without the discovery of partial differentiation equations, no single modern chip or radio receiver or transmitter could either be conceived or constructed”.

computadoras.

Atualmente, entretanto, o conceito de Algoritmo - recorrentemente levantado nas discussões sobre grandes corporações digitais como Google e Facebook e igualmente em voga no âmbito da comunicação ao trabalhar com a ideia de filtro-bolha (PARISIER, 2011) -, por vezes é esquecido como uma função de origem matemática e tratado como um conceito apenas de ordem imaterial e cultural. Ao utilizarmos computadores (em sentido amplo do termo), estamos, entretanto, nos sujeitando a toda a uma ordem de coisas e a procedimentos que possuem uma lógica própria e que pertencem ao campo da computação (GILLESPIE, 2018).

Um algoritmo busca resolver computacionalmente uma ordem de problemas que podem ou não ser resolvidos computacionalmente – o que determina, porém, se um problema pode ou não ser resolvido por um computador não são questões de ordem tecnológica, mas de ordem matemática (HAREL, 2000; 2004). Rieder (2018) aponta para um desconhecimento - mesmo no campo das ciências sociais que busca compreender a imbricação entre sociedade e tecnologia - dos mecanismos técnicos e dos modos de operação de softwares e algoritmos, por exemplo.

Entretanto, continua existindo um contraste entre o diagnóstico cada vez mais aceito de que o “software estrutura e possibilita grande parte do mundo contemporâneo” (FULLER, 2008, p.1) e a atenção dada a casos específicos e métodos precisos. Se o software realmente se tornou uma técnica de poder ou, precisamente, um meio para projetar e implementar técnicas complexas para “conduzir condutas” (FOUCAULT, 2004a, p. 192), é desconcertante que a análise crítica de objetos técnicos concretos, procedimentos, e práticas sejam extremamente raras (p. 125).

O autor propõe (RIEDER, 2018), dada a dificuldade de acesso e análise de algoritmos específicos, o engajamento de pesquisadores com o que ele entende como técnicas algorítmicas, que, ainda que não correspondam a um algoritmo específico, correspondem a um modo de pensar e ou abordar a construção de um algoritmo para um determinado problema.

Se o algoritmo vem recebendo centralidade nas pesquisas contemporâneas, não é por acaso. Elemento essencial do processo computacional, o algoritmo pode ser descrito como a prescrição que determina a forma pela qual o computar irá utilizar o input recebido e transformá-lo no output desejado (HAREL, 2000; 2004). Ainda que esse conceito possa parecer extremamente próximo do que descrevemos como software no segmento anterior (e estes são, de fato, conceitos que se borram quando falamos de computação), o algoritmo se define justamente como os passos a serem articulados, enquanto o software é o programa em si, que descreve estes passos em uma linguagem computacional, e pode ser composto por diversos algoritmos. As funções que um algoritmo computacional desenvolve, em sua maioria, podem ser feitas sem a utilização de uma máquina computacional - ou seja, é possível computar sem



um computador e o método algorítmico pode ser utilizado fora deste contexto (GILLESPIE, 2018).

Existem três ordens de problemas algorítmicos: aqueles que podem ser computados; aqueles que podem ser resolvidos computacionalmente, mas para os quais não existe um algoritmo eficaz<sup>21</sup>; aqueles que não podem ser resolvidos computacionalmente. Ainda que todos os problemas com um conjunto finito de inputs possa ser computacionalmente solucionado, o mesmo não é verdade para um conjunto infinito de inputs legais. Problemas computáveis são aqueles para os quais é possível descrever uma sequência de passos que funcione para qualquer um dos inputs possíveis.

A Tese de Church-Turing é essencial na compreensão do que significa o princípio da computabilidade. Consideramos a máquina de Turing como uma máquina universal por sua capacidade de ser programada a fim de lidar com números de tamanhos e formas variáveis: são dotadas da capacidade de solucionar qualquer problema que possa ser solucionável algoritmicamente (HAREL, 2000; 2004). Computabilidade diz respeito justamente a esse conjunto de problemas que podem ser algoritmicamente resolvidos – ainda que a tese de Church-Turing não possa ser matematicamente comprovada, permanece sem ter sido falseada cientificamente, nos termos de Popper (2013).

O que essa tese supõe é que uma máquina de Turing é equivalente à qualquer computador real com qualquer linguagem de programação (hoje, ou no futuro): ou seja, variáveis de ordens técnicas não seriam capazes de tornar problemas não computáveis em computáveis. Para Harel, "uma das melhores formulações da tese não menciona qualquer modelo em particular, mas simplesmente declara que todos os computadores e todas as linguagens de programação são equivalentes em poder computacional, dado tempo de computação e espaço de memória ilimitados<sup>22</sup>" (HAREL, 2000, p. 42).

Os algoritmos são, em geral, uma composição de certa forma híbrida, uma vez que podem ser tanto fórmulas matemáticas utilizadas para prever o movimento de um planeta ou o conjunto de instruções de navegação (GILLESPIE, 2018). Os problemas algorítmicos de maior ocorrência podem ser categorizados da seguinte forma, sendo que, muitos deles podem ser híbridos: problema de cálculo aritmético<sup>23</sup>; problema de cálculo aritmético com quantidade de

---

<sup>21</sup> Para os quais a computabilidade exigiria uma quantidade de tempo ou espaço extremamente irrealista em comparação com as proporções de tempo e espaço enquanto fisicamente experienciadas pelos seres humanos.

<sup>22</sup> "In fact, one of the most striking formulations of the thesis doesn't mention a particular model at all, but states simply that all computer and all programming languages are equivalent in computational power, given unlimited computation time and memory space." (HAREL, 2000, p. 42).

<sup>23</sup> Corresponde a uma fórmula simples, com cálculo aritmético e input de dois números. Exemplo:  $n2 + 5J$

elementos variáveis<sup>24</sup>; problemas de decisão<sup>25</sup>; problemas não-aritméticos<sup>26</sup>; problemas de busca<sup>27</sup>; problemas sobre algoritmos<sup>28</sup> (HARREL, 2000; 2004).

É interessante pensar que, assim como pode existir mais de um algoritmo que solucione um determinado problema de forma adequada, também é possível formular o próprio problema de formas distintas – assim, a própria enunciação de um problema enquanto computacional já pressupõe uma decisão, com implicações específicas no tipo de resultado obtido. A legitimidade algorítmica, entretanto, reside justamente na ideia de não-interferência humana - trunfo da automação - (GILLESPIE, 2018) ainda que esteja permeado por uma imbricação diversa de lógicas humanas.

O tempo de duração do processo algorítmico não corresponde necessariamente a um longo processo de passos a serem tomados, mas pode variar conformes casos de inputs específicos, por exemplos, ou exigirem uma repetição de uma sequência simples de ações muitas vezes (HAREL, 2000; 2004). Um método algorítmico supõe a definição de um problema específico com potencial de solução para qualquer input da categoria possível definida como válido. A descrição de um problema algorítmico deve incluir a descrição de dois itens: uma definição precisa dos inputs legais<sup>29</sup> e uma caracterização precisa do output requerido em função do input.

Ainda que estejamos falando da lógica de funcionamento do algoritmo e da forma como este tipo de problema é pensado e elaborado, é importante ressaltar a impermanência dos algoritmos com os quais lidamos cotidianamente, uma vez que estes estão em um constante processo de atualização, muitas vezes invisível para o usuário. Para além de sua impermanência, o algoritmo age enquanto processo, portanto,

um algoritmo ou sistema baseado em Machine Learning é, por definição, não uma coisa, é sempre um devir. Para mim, essa noção de devir [a partir de Whitehead] é bastante útil porque também é um ponto de vista metodológico de que qualquer coisa que estejamos estudando é sempre uma coisa em seu devir (BUCHER; D'ANDRÉA;

---

<sup>24</sup> Exemplo: sendo X o input, correspondendo a um número positivo, soma de números de 0 a X.

<sup>25</sup> Nesses casos, ainda que o problema envolva aritmética, o output será sempre sim ou não (ou seja, não-numérico).

<sup>26</sup> Não exigem operações aritméticas, mas podem exigir a definição computacional de todos os termos utilizados. Exemplo: ordenamento alfabético de uma lista de palavras.

<sup>27</sup> Em geral, buscam um processo de otimização, como: encontrar o caminho mais rápido entre dois pontos de rotas de um mapa. Às vezes, pode ser formulado como um problema de decisão, gerando, como output, ao invés do trajeto, sim ou não.

<sup>28</sup> Problemas nos quais o input é outro algoritmo.

<sup>29</sup> Definição precisa da categoria a qual os inputs corresponde. Pode ser, por exemplo, números, palavras, uma lista de livros, salários de funcionários, etc. Quanto mais complexa a categoria, mais complexa a sua definição em termos computacionais. Se um input inadequado for fornecido, o algoritmo não irá funcionar.

JURNO; 2018, p. 167, tradução nossa <sup>30</sup>).

Essa natureza lógica dos problemas e suas soluções, podem hipoteticamente gerar transformações na ordem de estruturação social contemporânea, uma vez que, a centralidade deste código que já recebemos codificado pela máquina, pode criar não apenas vieses, mas, associado à questões de ordem cultural, social, política e econômica, verdadeiras configurações de mundo. Podemos arriscar dizer que, se como proposto por Flusser (2007), nossos códigos nos programam enquanto sociedade, então a sociedade contemporânea é, ao menos parcialmente, também matematicamente programada, dado o que a maioria dos algoritmos de uso cotidiano se baseia em técnicas de probabilidade (RIEDER, 2018), por exemplo.

Esse tipo de codificação obedece a uma série de regras que é não apenas ignorada por uma parcela razoável dos usuário, mas também por parte dos pesquisadores no âmbito das ciências sociais e humanas, uma vez que os códigos matemáticos e lógicos e os códigos sociais foram tratados de formas tão distintas – quero dizer, poucos cientistas sociais possuem um treinamento lógico e matemático razoável, e a recíproca é verdadeira para matemáticos e cientistas da computação. Ainda que o domínio completo de um campo altamente complexo seja provavelmente inviável, a comunicação pode beneficiar-se de estudos que busquem compreender a lógica de funcionamento de determinadas práticas algorítmicas, perspectiva que, uma vez cruzada com métodos de pesquisa das ciências sociais, pode oferecer interessantes insights sobre a interação entre seres humanos e tecnologias digitais baseadas em técnicas de inteligência artificial, cada dia mais complexas.

A técnica, entretanto, conforme já pontuamos em seções anteriores, também é um componente relevante do sentido – ou seja, conforme Rieder (2018), para a compreensão das formas de ação social de um objeto técnico, a compreensão de seu “significado funcional” também se faz necessária: seria preciso, portanto, tentar apreender o objeto técnico em seu modo de funciona, quais, como e porque são as ações que ele realiza.

Prever os impactos do uso em escala global de novos códigos de funcionamento, como pontuou Flusser (2007), é difícil – entretanto, buscar compreender o modo de funcionamento destes e suas próprias estruturas lógicas e formais, me parece um desafio para a ciência da comunicação. Para que seja possível estabelecer pontos de contato e diálogo entre técnicas algorítmicas e os processos culturais que as envolvem, se faz necessário um aprofundamento do conhecimento técnico da área da computação, a fim de diminuir a “fenda cultural e

---

<sup>30</sup> Original: "an algorithm or machine learning system is, for definition, not a thing, it's always a becoming. For me, this notion of becoming is quite useful because it is also a methodological point that whatever we're studying is a thing in its becoming."

epistemológica existente entre humanistas e engenheiros.” (RIEDER, 2018, p. 139). As técnicas algorítmicas apresentam maneiras muito específicas de leitura do mundo e modos de ação sobre este, ainda guiadas por um propósito específico, como, por exemplo, ranquear páginas de interesse em mecanismos de busca ou realizar reconhecimentos faciais. Se o software ainda é pensado como uma entidade imaterial, seus efeitos são materialmente sentidos através de ações e consequências que operam cada vez mais em nosso cotidiano.

Algoritmos são, para Fazi e Parisi (2014), o estado da arte do pensamento mecanizado. Não existe necessidade de que os sistemas computacionais tentem simular as condições de uma consciência humana, dado que os algoritmos podem expressar, através de sua agência material, determinados pontos de vista. Ou seja, algoritmos expressam uma particularidade, um certo viés que está intrinsecamente relacionado com o modo como eles foram pensados, desde a formulação de seus problemas, até a definição de categorias específicas. Esse viés pode se expressar através de ações com efeitos reais no mundo: “enquanto androides possuem um corpo mas estão procurando por uma alma, no tempo presente o pensamento mecanizado enquanto processo computacional criam instâncias que já são materiais e encarnadas, uma vez que sempre são performados e inscritos na sociedade, na cultura e na economia<sup>31</sup>” (FAZI e PARISI, 2014, p.109).

O ponto da discussão trazido pela autora é que o cerne da questão não é a potencialidade de liberação de sua própria prescrição da mecanização computacional, mas se a própria forma prescritiva mecaniza já não comporta hoje uma certa autonomia. Esse ponto, ainda que diretamente relacionado com o conceito de algoritmo, será tratado em mais profundidade quando das perspectivas dos desenvolvimentos em inteligência artificial, no capítulo seguinte.

No âmbito da comunicação, as questões dos algoritmos vem sendo pesquisadas principalmente a partir do que Gillespie (2018) caracteriza como algoritmos de relevância pública: aqueles que constituem uma camada relevante na seleção e recorte das informações que os sujeitos acessam a partir de um conjunto de dados, em um contexto de grande disponibilidade de conteúdo - ou seja, os algoritmos voltados para hierarquização da informação funcionam como uma espécie de filtro contemporâneo. Considerando que algoritmos são também “ferramentas emergentes para o conhecimento e discurso públicos” (2018, p. 98), Gillespie aponta para a necessidade de pensar as escolhas institucionais e humanas que configuram os algoritmos tais como são.

---

<sup>31</sup> No original: “while androids had a body but were looking for a soul, in the present day mechanized thoughts instantiated in computational processes that are themselves already material and incarnated, insofar as they are always performed and embedded in society, culture and the economy” (FAZI, 2014, p.109).

O algoritmo enquanto forma de mecanização do raciocínio acaba por gerar um deslocamento no exercício de ação e poder humanos: ao invés de agirmos sobre definições de resultados, as escolhas passam a agir no sentido de definição de procedimentos (RIEDER, 2018). Assim, quando falamos de algoritmos, se tornaria necessário compreender quais são as tomadas de decisão que estão em jogo e em que momento elas operam. Sobre a pesquisa em torno das técnicas algorítmicas como estratégia metodológica para estudos de tecnologias no âmbito social, como seu estudo acerca do *classificador de bayer*, Rieder argumenta (2018):

[...] cada vez mais está ficando claro que os algoritmos - geralmente baseados em técnicas de probabilidade - estão desempenhando um papel crucial ao decidir como a informação circula, como as pessoas se encontram e se relacionam, e como a conduta é efetivamente realizada. E enquanto quase todas as técnicas deixam espaço para muitas decisões e podem ser desenvolvidas em várias direções, seu horizonte conceitual ainda implica uma orientação epistemológico e formas específicas de intervenção (p. 139).

Na intersecção de uma preocupação com as formas de seleção de conteúdos, os processos comunicativos desenvolvidos por usuários em plataforma digitais e as políticas envolvidas surge uma proliferação de pesquisas relevantes na área da comunicação nos últimos anos (ver por exemplo: JURNO, 2016; JURNO; D'ANDRÉA, 2017; ARAÚJO, 2017; BUCHER, 2012; 2016).

Para Gillespie (2018), importa investigar os contextos específicos nos quais esses algoritmos podem ter ramificações políticas, de seis dimensões chave dos algoritmos de relevância pública: padrões de inclusão; ciclos de antecipação; avaliação de relevância; promessa de objetividade algorítmica; entrelaçamento com a prática; produção de públicos calculados. Para a investigação aqui proposta, são dimensões interessantes os padrões de inclusão, os ciclos de antecipação e a promessa de objetividade algoritmo, por suas relações com os desenvolvimentos em inteligência artificial.

Por considerarmos que sistemas como as assistentes pessoais virtuais também estão conectadas à internet como bancos de dados para solução de demandas dos usuários, os padrões de inclusão destes bancos de dados terão aqui uma dimensão relevante. Assim, de certa forma, essas assistentes operam também enquanto ferramentas de busca - embora não exclusivamente -, e aquilo que é ou não incluído em seus bancos de dados pode possuir e/ou construir vieses políticos. Exemplo interessante para este ponto foi a ausência de resultados na busca realizada pela Siri de clínicas de aborto em Manhattan, nos Estados Unidos, percebida por blogueiras que

solicitaram esta informação para a assistente pessoal virtual (WORTHAM, 2011)<sup>32</sup>. Ainda que a Apple tenha alegado que a ausência de resultados se devia a um erro do sistema ainda em sua versão beta, tal distorção de resultados não deixa de possuir implicações políticas e sociais.

Um segundo ponto está relacionado aos ciclos de antecipação, ou seja, ao conjunto de informações sobre um determinado usuário e o grupo ao qual pertence que são utilizadas como base para criar um perfil de usuário, que será um input utilizado no algoritmo de filtragem de informação (GILLESPIE, 2018). Essas informações geralmente são coletadas através de um gama de ferramentas de uma mesma empresa (o Google, por exemplo, uma coleção de produtos muito maior do que seu mecanismo de busca) e das ações do usuário nas mesmas, utilizadas, em última instância, na tentativa de antecipar aquilo que o usuário gostaria. Gillespie pontua ainda que:

Mas, também nessas antecipações, valores implícitos e, muitas vezes, políticos podem estar inscritos na tecnologia: os hábitos de percepção ou interpretação de alguns usuários são tomados como modelo universal, hábitos contemporâneos são imaginados como atemporais, objetivos computacionais particulares são tratados como auto-evidentes (p. 103).

Ou seja, aqui o autor está apontando para o risco de tomar categorias artificialmente e aproximadamente criadas como realidade – um processo que se não reduz ao modo de ação específico do algoritmo ou mesmo de características relacionadas exclusivamente as tecnologias. Conforme pontua Broussard (2018), “quando olhamos par ao mundo apenas pelas lentes da computabilidade ou quando tentamos resolver problemas sociais usando apenas a tecnologia, tendemos a recriar uma série de problemas previsíveis que impedem o progresso e reforçam a desigualdade”<sup>33</sup> (p. 22), principalmente na utilização de algoritmos para tomada de decisão ou com finalidade preditiva. Como algoritmos são “procedimentos computacionais para derivar resultados” a noção “de que computadores são mais objetivos ou sem vieses porque eles quebram perguntas e respostas de forma a se tornarem uma avaliação matemática”<sup>34</sup> (p. 24) não é real.

A partir da lógica de coleta de dados pessoais, as assistentes pessoais podem ser mais um braço na própria infraestrutura técnica e modelo de negócios de empresas do setor de

<sup>32</sup> Disponível em: <https://bits.blogs.nytimes.com/2011/11/30/apple-says-siris-abortion-answers-are-a-glitch/>

<sup>33</sup> No original: “when we look at the world only through the lens of computation, or we try to solve bi social problems using technology alone, we tend to make a set of the same predictable mistakes that impede progress and reinforce inequality

<sup>34</sup> Nos originais em inglês: “An algorithm is a computational procedure for deriving a result” e “the notion that computers are more “objective” or “unbiased” because they distill questions and answers down to mathematical evaluation”

informática, uma vez que as empresas que as fabricam são já consolidadas no mercado (Apple, Amazon, Microsoft e Google). Como veremos adiante, essa coleta de dados é parte integrante do próprio processo de aprendizagem e melhoria desses artefatos, uma vez que as técnicas atualmente empregadas em inteligência artificial depende de uma estrutura bottom-up, que utilize uma grande quantidade de dados a partir da qual se deriva resultados e se aperfeiçoam os sistemas de compreensão. O acesso às conversas por voz em ambientes domésticos, entretanto, podem significar uma nova camada de acesso tanto aos dados quanto à privacidade das pessoas, com efeitos de interferência ainda não previstos.

Exemplo desta questão foi o recente caso do casal de Portland, nos Estados Unidos, que reportou ao sistema de atendimento ao consumidor da Amazon, que a Alexa, a assistente pessoal inteligente da empresa, gravou, sem pedido explícito dos donos, uma conversa privada que ocorreu na casa. Após gravar, a Alexa enviou a conversa para um dos contatos do casal, um empregado do marido, que os contactou para explicar o que havia ocorrido. A empresa já possui mais de 40 milhões de dispositivos como esse em uso e recentemente decidiu participar de um programa de vigilância do governo dos Estados Unidos. O suporte da Amazon, mesmo após investigação dos logs do dispositivo por engenheiros da empresa, não conseguiu explicar o motivo do acontecimento (GREENE, online, 2018<sup>35</sup>).

Se por trás da construção de um algoritmo de relevância pública existe uma filosofia específica sobre aquilo que é conhecimento (GILLESPIE, 2018), podemos pensar que, por trás dos desenvolvimentos de inteligência artificial existe uma noção daquilo que é inteligência e que tais desenvolvimentos potencialmente tentam responder à estes pressupostos. Essa premissa de fundo irá ter um papel importante nos mecanismos especificamente desenvolvidos: isso quer dizer que, dados pressupostos diferentes, as abordagens técnicas para obter uma inteligência artificial seriam outras. Este assunto será abordado em maior profundidade no próximo capítulo, onde trataremos, a partir de uma perspectiva histórica, dos desenvolvimentos e noções envolvidos no campo da Inteligência Artificial.

Gillespie propõe ainda que existe uma influência de ordem recursiva entre o cálculo dos algoritmos e o cálculo dos usuários (GILLESPIE, 2018), - ainda que nem sempre cientes dos meandros de funcionamento de seus objetos digitais, é justamente na forma da interação entre um e outro que podemos compreender as modificações/rupturas/continuidades que estes objetos trazem para o mundo:

---

<sup>35</sup> Disponível em: <https://thenextweb.com/artificial-intelligence/2018/05/24/amazon-cant-explain-why-alexa-recorded-and-shared-a-private-conversation/>

Se os usuários falharem ou se recusarem a se encaixar nas suas práticas e a conceder significado a elas, o algoritmo irá falhar. Isso significa que não devemos considerar seus “efeitos” nas pessoas, mas um “entrelaçamento” multidimensional entre algoritmos postos em prática e as táticas dos usuários que fazem face a eles (p. 110).

Entretanto, tais processos, podem, de fato, construir a realidade. Se pensarmos, por exemplo, sobre a forma como o Google filtra o que é ou não uma informação ou conhecimento relevante, e considerarmos que, atualmente, uma quantidade elevada de pessoas utiliza este serviço como ferramenta de busca para o conteúdo disponível na web, a reiteração desta ação, ao longo do tempo, se torna fator relevante na própria construção da noção do que é conhecimento relevante ou de como podemos estabelecer os critérios para essa relevância. Ou seja, o uso repetido, em larga escala, de uma lógica determinada, pode construir o entendimento tanto de uma lógica correta, quanto neutra.

## 1.5 BANCO DE DADOS E O ARQUIVO

Em sua obra “Mal de Arquivo: uma impressão Freudiana”, Jacques Derrida (2001), fala da ação de arquivamento como algo que se lança, mais do que para o passado, para o futuro: “Ao mesmo tempo, mais que uma coisa do passado, antes dela, o arquivo deveria pôr em questão a chegada do futuro” (p. 48). Considerando que “o arquivo sempre foi um penhor e, como todo penhor, um penhor do futuro. Mais trivialmente: não se vive mais da mesma maneira aquilo que não se arquivava da mesma maneira” (p. 31), buscaremos aqui lançar ideias iniciais sobre o que estamos penhorando como forma de futuro em um tipo específico de constituição de arquivo digital contemporâneo, a partir da própria compreensão de como funcionam as ferramentas responsáveis pela estruturação do arquivamento hoje, uma vez que “os sentidos arquiváveis se deixam também, e de antemão, co-determinar pela estrutura arquivante” (p. 31).

Nos parece possível estabelecer relações entre o conceito contemporâneo de banco de dados – sendo que dados “podem ser quaisquer variáveis, basta que apresentem um formato definido (analógico ou digital, bits ou palavras, etc)” (KITTLER, 2017, p. 246). Ponto chave para tecnologias digitais de uso cotidiano como ferramentas de busca na web, sistemas de navegação com base em dados de trânsito ou categorização automática de um conjunto de fotos, por exemplo – com a noção de arquivo conforme proposta por Derrida (2001). Acontece que, o que chamamos de banco de dados na ciência da computação e/ou informação, funciona tanto



como arquivo, quanto como insumo básico de diversos ramos associados à informática – desde os mecanismos de busca até tecnologia de inteligência artificial. Ou seja, estes desenvolvimentos tecnológicos dependem da própria existência do arquivo e dos dados que os compõe.

Uma vez que tanto a noção de arquivo quanto a proposição de que, na contemporaneidade, uma das faces arquivistas fortemente marcada pelas características materiais e modos de funcionamento das tecnologias de informação são os banco de dados, matriz para uma gama distinta de funcionamento de técnicas algorítmicas<sup>36</sup>, são complexas, faremos nesta seção apenas um caminho exploratório de questões pertinentes a serem pesquisadas em profundidade. Propomos o deslocamento do conceito de arquivo para a ideia de banco de dados e a desconstrução de uma lógica de neutralidade frequentemente ligada às técnicas algorítmicas associadas aos bancos de dados.

Neste trabalho estaremos olhando para um arquivo que possui uma dupla forma de constituição. Uma parte diz respeito aos documentos catalogados que constituem, por exemplo, o acervo de páginas na web indexados pelo Google, ou o conjunto de vídeos publicados no YouTube; Outra parte é constituída pela coleta sistemática de informações sobre as ações online dos usuários em plataformas determinadas, que vem a compor um perfil específico utilizado como parâmetro para uma parte considerável das técnicas algorítmicas ligadas a catalogação e hierarquização do primeiro conjunto de documentos citados.

O conjunto de técnicas de informação que, de forma conjugada, compõem as ferramentas necessárias para transformarem, por exemplo, rastros digitais coletados de usuários em bases de dados deste segundo tipo passa por políticas específicas de cada empresa. Tais políticas irão coordenar um conjunto de dados, detendo as funções de unificar, identificar e classificar os componentes deste arquivo – tais como os arcontes, na casa dos quais se depositavam os documentos oficiais no sentido latino do termo arquivo (DERRIDA, 2001). Essas políticas podem ocultar, por exemplo, noções de uma empresa sobre aquilo que é privado e aquilo que é público, conforme problematizado a partir da coleta de dados para o Google Street View (GILLESPIE, 2018). O poder de consignação de um arquivo, “tende a coordenar um único corpus em um sistema ou uma sincronia na qual todos os eleitos articulam a unidade de uma configuração ideal” (DERRIDA, 2001, p. 14).

Tal poder se expressa também a partir dos métodos de classificação e hierarquização utilizados em um determinado arquivo, como a técnica conhecida como rebaixamento

---

<sup>36</sup> Ver mais em: RIEDER, 2018.

algorítmico. Nesse caso, um algoritmo, junto a um banco de dados, utiliza ferramentas para que um conteúdo específico possua menos visibilidade que os demais, tornando-o invisível, sem a necessidade de removê-lo:

O YouTube “rebaixa algorítmicamente” vídeos provocantes para que não apareçam nas listas dos mais assistidos ou na página inicial gerada para novos usuários. O Twitter não censura conteúdos profanos de tuítes públicos, mas ele os remove da avaliação algorítmica que determina quais termos figuram nos “Trending” (GILLESPIE, 2018, p. 101).

Ou seja, as técnicas algorítmicas utilizadas como forma de organização e hierarquia de um arquivo possuem também vieses de ordem social e política, estando atreladas ao que Derrida (2001) apontava como o conceito de comando do arquivo. Estas mesmas técnicas podem ser utilizadas para diminuir a visibilidade de resultados relacionados à questões políticas em governo autoritários, por exemplo. Ou seja, diz respeito ao sentido do arquivo como um princípio da lei, sendo forma de exercício de autoridade e de ordem social.

Como o *archivum* ou o *archium* latino (palavra que empregamos no singular, como era o caso inicialmente do francês “*archive*”, que outrora era usado no singular e no masculino: “*un archive*”), o sentido de “arquivo”, seu único sentido, vem para ele do *arkheion* grego: inicialmente uma casa, um domicílio, um endereço, a residência dos magistrados superiores, os arcontes, aqueles que comandavam. Aos cidadãos que detinham e assim denotavam o poder político reconhecia-se o direito de fazer ou representar a lei (p. 12).

A capacidade de agir politicamente, através de fórmulas criadas para agirem de forma automatizada e permanecerem protegidas pela diretriz do segredo comercial, desloca, de certa forma, o poder exercido sobre este tipo de arquivo – é um poder diluído, automatizado, publicamente concebido como uma técnica neutra.

Para além da coleta dos dados, é importante pensar sobre o processo de limpeza dos dados: quais as técnicas e concepções envolvidas no processo de tornar um determinado conjunto de dados em um conjunto tratável para que um algoritmo específico possa funcionar de forma automática, levando a uma composição específica de arquivo? Não existem dados brutos, todos os dados são, de uma forma ou outra, tratados - as informações são preparadas para a utilização de um algoritmo (GILLESPIE, 2018). Ou seja, ainda que a utilização da palavra ‘dados’ ao invés de ‘documentos’ possa justamente apontar para um espectro da neutralidade técnica, percebemos que um conjunto de dados carrega em si marcas próprias do arquivo: são selecionados e tratados, ou ainda destruídos e esquecidos, uma vez que “o arquivo tem lugar em lugar da falta originária e estrutural da chamada memória” (DERRIDA, 2001, p.

22).

A construção de um banco de dados passa também por um processo de exclusão, ou seja, por aquilo que se decide não pertencer ou não interessar a um determinado banco de dados. O processo de exclusão dos dados digitais, portanto, mantém sua importância na construção de um arquivo, que é marcado tanto por aquilo que armazena, quanto por aquilo que exclui – a pulsão de destruição opera também nos arquivos de dados digitais.

Se não há arquivo sem uma exterioridade, a exterioridade digital pressupõe não apenas um outro aparato técnico igualmente dispendioso em termos de recursos materiais (trocamos a tinta e o papel pelos servidores de armazenamento, cabos de fibra óptica e computadores), suas possibilidades de memorização, repetição reprodução (DERRIDA, 2001) estão associadas às fórmulas automatizadas que chamamos hoje de algoritmos e que se encontram, em grande parte, nas mãos de empresas privadas. O Google pode não ser proprietário dos sites indexados em seu banco dados, mas o é da formulação específica de fatores (muitas vezes atrelados a bancos de dados coletados de informações dos usuários) que determinam um resultado X para uma busca neste arquivo. Isso depende tanto do processo de catalogação das informações existentes nos documentos que formam o arquivo Google de páginas da Web, quanto de políticas da própria empresa.

Apesar de os serviços de informação em larga-escala se gabarem por serem abrangentes, esses sites são, e sempre devem ser, censores de informações também. Dados indexados excluem spams e vírus; vigiam violação de direitos autorais e pornografia; e retiram o que é obsceno, condenável ou politicamente controverso das bases de dados (GILLESPIE, no prelo)

Se “a estrutura técnica do arquivo arquivante determina também a estrutura do conteúdo arquivável em seu próprio surgimento e em sua relação com o futuro. O arquivamento tanto produz quanto registra o evento” (DERRIDA, 2001, p. 29), as tecnologias que utilizamos não apenas registram e arquivam determinados acontecimentos, mas a própria materialidade destes dispositivos agem no acontecimento em si, gerando uma “experiência política dos meios chamados de informação” (p. 29). Assim as tecnologias de informação são forma de compreensão do mundo, não apenas ferramentas de transmissão e armazenamento.

Atualmente, o próprio modo de criação de ferramenta de Inteligência Artificial passa pela técnica das redes neurais profundas, um modo de aprendizado de máquina baseado em banco de dados. O arquivo, então, é armazenado e explorado a partir de técnicas específicas de funcionamento tão complexo que nem sempre podem ser plenamente explicadas por seus desenvolvedores (HAREL, 2000). As assistentes pessoais digitais disponíveis hoje, tais como

a Cortana (Microsoft), Alexa (Amazon) e a Siri (Apple), são na verdade uma combinação de diversas funções, como software de reconhecimento de fala, interface de linguagem natural e assistente pessoal baseado em inteligência artificial.

Esses dispositivos estão inevitavelmente associados a banco de dados de duas ordens: um primeiro banco de dados montado a partir dos próprios usuários, suas preferências e interações prévias, ou seja, de ordem privada, e um segundo banco de dados armazenado na nuvem, ou seja, de consulta online. Toda vez que fazemos uma solicitação para a Siri, por exemplo, e este pedido não puder ser solucionado a partir das ferramentas do próprio aparelho (ou seja, do banco de dados interno), a assistente pessoal irá realizar a busca para solucionar a solicitação a partir de um banco de dados na nuvem (MERCHANT, 2017). Ou seja, a ideia de Inteligência Artificial hoje, desenvolvida a partir de técnicas de *machine learning* e redes neurais profundas<sup>37</sup>, é indissociável de seu banco de dados e das ferramentas de análise destes, ou seja, depende de uma ideia automatizada, porém não neutra, de arquivo<sup>38</sup>.

## 1.6 PODEM AS MÁQUINAS PENSAR?

No artigo inaugural do que hoje conhecemos como o campo da Inteligência Artificial (IA), intitulado “*Computing machinery and Intelligence*” (1950), Alan Turing, quase quinze anos após conceber, em 1936, a máquina abstrata que viria a ser conhecida como ‘máquina Universal’ ou ‘máquina de Turing’ (nosso computador contemporâneo), nos propõe considerar se “uma máquina pode pensar” (TURING, 1950, p. 433). Para escapar do que ele considera uma dificuldade em trabalhar com termos que exigem uma definição específica, dada a amplitude com que estes conceitos podem ser tratados, especificamente “pensar” e “máquina”, o autor realiza um deslocamento da questão, buscando explorar a possibilidade de uma máquina passar no que ele chama de jogo da imitação. Apesar dos inegáveis avanços, as questões do significado e da estrutura do ato de pensar permanecem em aberto no âmbito das investigações de IA.

---

<sup>37</sup> Ver mais em: LARGELY, Pat. The Changing Science of Machine Learning. In: Machine Learning, Volume 82, N 3, março de 2011, pp. 275-279.

<sup>38</sup> Outra perspectiva interessante para ser investigada a partir da ideia de arquivo são os softwares de tomada de decisão com base em banco de dados, como utilizados por exemplo em cortes judiciais para decisões de liberdade condicional e por departamentos de polícia para determinar as áreas que devem ser mais policiadas nos Estados Unidos. Ambos os casos foram estudados em profundidade e os vieses reproduzidos pela não historicidade desse corpus de dados podem ser conferidas em Burton et al., 2017.

Ainda que com certas imprecisões, Turing irá entender o conceito de máquina, para este propósito, como um computador digital. Conforme o autor, a ideia deste tipo de dispositivo é na verdade bastante antiga, podendo ser traçada até Charles Babbage, Professor da Universidade de Matemática de Cambridge, que descreveu um tipo de máquina chamada por ele de Engenho Analítico entre 1828 e 1839, mas mesmo que “Babbage possuísse todas as ideias essenciais, sua máquina não era, naquele tempo, um prospecto muito atrativo<sup>39</sup>.” (TURING, 1950, p. 439, *tradução nossa*). Tal escolha se justifica devido a propriedade específica deste tipo de máquina, que pode ser programada de forma a modelar qualquer tipo de máquina:

Essa propriedade especial dos computadores digitais, de poderem mimetizar qualquer outra máquina de estado discreto, são descritas ao se dizer que são máquinas universais. A existência de máquinas com essa propriedade tem a importante consequência que, desconsiderando questões de velocidade, não é necessário desenhar várias novas máquinas para realizar processos computacionais. Todos eles podem ser feitos em um computador digital, programado adequadamente para cada caso. Como consequência disso, todos os computadores digitais são, de certa forma, equivalentes.<sup>40</sup> (TURING, 1950, p. 441 e 442, *tradução nossa*).

A partir dessa conceituação de máquina, Turing (1950) irá propor o jogo da imitação como forma de determinar a inteligência desse tipo de sistema. O autor parte de uma versão existente desse jogo que ocorre com dois participantes humanos e propõe a substituição de um deles por uma máquina. Na versão original, um homem, uma mulher e um interrogador estão separados por barreiras físicas que impedem o contato visual e sonoro. O papel do interrogador é questionar ambos os participantes e tentar descobrir qual deles é o homem e qual deles é a mulher, uma vez que os dois participantes deveriam performar o papel feminino e tentar convencer o juiz de que eram uma mulher<sup>41</sup> – ambos os papéis, evidentemente, tornam-se confusos, devido a dificuldade de determinar a intenção de cada um dos jogadores.

A mesma ideia de que é necessário mimetizar o comportamento considerado

---

<sup>39</sup> No texto original: “Babbage had all the essential ideas, his machine was not at that time such a very attractive prospect.”

<sup>40</sup> “This special property of digital computers, that they can mimic any discrete state machine, is described by saying that they are universal machines. The existence of machines with this property has the importance consequence that, considerations of speed apart, it is unnecessary to design various new machines to do various computing processes. They can all be done with one digital computer, suitably programmed for each case. I will be seen that as a consequence of this all digital computers are in a sense equivalente.”

<sup>41</sup> Genova (1994) faz uma leitura interessante da passagem realizada por Turing entre o Jogo da Imitação com uma participante mulher para um participante máquina, propondo que, para Turing, o ato de pensar e o conhecimento estão relacionados ao gênero, sendo este uma construção discursiva e cultural. A pergunta de Turing seria, para ela, “quem pode pensar?”, e sua resposta aponta para uma maleabilidade entre gêneros e formas humanas e não-humanas. Tal análise encontra respaldo na própria trajetória pessoal de Turing.

característico de uma mulher – daí o nome de jogo da imitação – passa para a versão de Turing (1950), na qual a ideia é que a máquina seja capaz de imitar a forma de comportamento humana. Para a realização do teste, são necessários uma máquina inteligente, um ser humano e um interrogador, sendo mais uma vez papel do interrogador determinar qual dos dois é o homem e qual é a máquina, em uma interação baseada na escrita. O computador precisaria passar por sessões com interrogadores diferentes para que pudesse ser excluída a possibilidade de palpite por parte do juiz (HAREL, 2000; 2004). A pergunta proposta por Turing (1950), é, então, se, no caso de substituição de um dos participantes por uma máquina, a porcentagem de erro do interrogador será a mesma.

Na tentativa de estabelecer parâmetros iniciais de formas de programar uma máquina para que possa passar no Jogo da Imitação, ou seja, como programar uma máquina para imitar adequadamente um humano, Turing (1950) ressalta três componentes: a forma biológica da mente humana ao nascer, educação formal recebida e experiências de vida cotidianas. Os dois últimos antecipam o que veio a ser uma das apostas da IA: o *machine learning*. Voltaremos a essa questão na seção subsequente deste documento.

Posteriormente, o jogo da imitação fica conhecido como Teste de Turing. Para que um computador seja capaz de passar nesse teste, enganando uma porcentagem dos juízes sobre sua natureza não-humana, o dispositivo deveria não apenas ser programado para agir de forma similar a um ser humano com inteligência comum. É preciso também que propriedades computacionais sejam disfarçadas, uma vez que estas poderiam entregar a máquina ao juiz: um ser humano, por exemplo, não é capaz de realizar cálculos complexos em velocidade rápida, mas um computador sim. Um computador programado para passar no Teste de Turing, por exemplo, deveria demorar uma quantidade de tempo similar à humana para realizar alguns cálculos (HAREL, 2000).

Apesar de ainda ser considerado um dos testes possíveis para verificar a inteligência de um sistema, a imprecisão dos termos apontada pelo próprio Turing coloca de saída uma primeira fragilidade: o fato de enganar o interrogador sobre sua natureza maquina é suficiente para que a máquina seja dita ‘inteligente’. Críticas sistemáticas foram feitas tanto à proposta de Turing em si, quanto à falta de precisão dos conceitos e dos métodos explicitados no texto (HARNAD, 2008).

Percebe-se a variação de dois critérios distintos para avaliação da performance a partir do jogo da imitação: enquanto em alguns momentos o que parece estar em questão é a similaridade das performances do homem e da máquina, em outros, o foco parece ser a possibilidade de detectar, através de características expressas na performance e relacionadas a

um conhecimento prévio de expectativa de comportamento, qual é o humano e qual é a máquina. Esta alternância da forma de avaliação do jogo da imitação parece colocar, de fato, em uma primeira perspectiva uma metodologia de engenharia reversa da performance de capacidade e, em uma segunda, um jogo de adivinhação, calcado em artifícios que visam enganar e confundir (HARNAD, 2008).

A versão aplicada do Teste de Turing, o *Loebner Prize*, tem sido igualmente criticado, seja pela incapacidade de testar a inteligência de um sistema devido ao desenho do concurso ou pelos truques utilizados pelos competidores na tentativa de enganar os juízes (WALSH, 2015). Toby Walsh, especialista em Inteligência Artificial, pontua, por exemplo, a necessidade de outras formas de testar as habilidades de uma máquina, não baseadas na mimetização: “Eu espero que a discussão atual sobre formas de substituição do Teste de Turing eventualmente nos levará a superar os testes para IA baseados na enganação e criar testes capazes de quantificar habilidade específicas e inteligência<sup>42</sup>” (WALSH, 2015, sem paginação, *tradução nossa*).

O *Loebner Prize* acontece anualmente desde 1991, concedendo seu prêmio principal, a medalha de ouro, para a inteligência artificial que conseguir enganar no mínimo 30% dos juízes sobre sua origem – até hoje, nenhum programa participante do torneio conseguiu ganhar a medalha de ouro. Segundo Andre Martin, um dos administradores do prêmio, “ninguém [da organização] está preocupado que a medalha de ouro possa sair tão cedo”<sup>43</sup> (2017, online, *tradução nossa*)<sup>44</sup>.

Em versões mais antigas do torneio, um prêmio era concedido ao programa de computador que recebesse a maior nota por parte dos juízes: o “Most Human Computer”. Como nessa versão todas as interações eram avaliadas individualmente, ou seja, não eram realizadas em duplas de humanos e máquinas, um prêmio também era concedido ao humano que recebesse a melhor nota dos juízes: “Most Human Human”<sup>45</sup>. Interessante notar que, mais a frente, na análise dos dados empíricos sobre como as pessoas se sentem ao utilizarem a linguagem natural falada com uma máquina, participantes do formulário quantitativo afirmaram que essa interação faz com que questionem sua própria humanidade (e não a da máquina): essa construção se

---

<sup>42</sup> No original: “I expect that the current discussion about replacements for the Turing test will eventually move from tests for AI based on deception to tests that quantify explicit skills and intelligence” (WALSH, 2015, sem paginação).

<sup>43</sup> No original: “no one is worried that the gold medal is going to go anytime soon.”

<sup>44</sup> Disponível em <https://gearpatrol.com/2017/05/18/chatbots-loebner-prize-turing-test/>. Kirsten Nelson, 2017.

<sup>45</sup> Aqui é possível ler a descrição da competição a partir da visão de dois participantes humanos, ambos ganhadores do prêmio de “Humano mais Humano”. Charles Platt escreveu sobre sua participação no evento em 1995, para a Revista Wired: <https://www.wired.com/1995/04/turing-2/>; Brian Christian participou do evento em 2009 e relata sua experiência para a The Atlantic: <https://www.theatlantic.com/magazine/archive/2011/03/mind-vs-machine/308386/>

assemelha a descrita por dois participante do *Loebner Prize*, que relataram uma preocupação constante com a forma como deveriam performar a sua humanidade. Além disso, Platt (1994) e Christian (2011), afirmaram que possuíam receio de serem tomados por uma máquina, falhando como humanos em sua expressão de humanidade.

A estrutura do *Loebner Prize* passou por diversas alterações ao longo dos anos. Atualmente, a avaliação é feita pelos juízes<sup>46</sup> a partir de conversas simultâneas com um humano e um *bot*<sup>47</sup>, através de dois *chats*, com duração de 25 minutos para cada rodada. Ao final deste tempo, o juiz deve definir qual dos dois interagentes era o humano e qual era a inteligência artificial. Cada juiz passa por aproximadamente quatro interações como as anteriores, com interagentes distintos em cada vez, para que cada dupla de humano/*bot* possa ser avaliado por um conjunto de juízes. O prêmio da medalha de ouro é concedido apenas se um *bot* convencer a todos os quatro juízes de que é humano; uma medalha de prata pode ser concedida para aquele que convencer dois dos quatro juízes; se nenhum dos concorrentes convencer, ainda é possível escolher um *bot* como preferido em termos de performance, que ganhará a medalha de bronze da competição<sup>48</sup>.

Ao propor o jogo da imitação, hoje materializado no *Loebner Prize*, Turing pretendia deslocar a questão “Uma máquina pode pensar?” para “Uma máquina pode passar no jogo da imitação?”, igualando uma pergunta à outra. Turing é, por um lado, o primeiro a propor a ideia de uma máquina capaz de pensar, mas, em seguida, desconsidera essa questão, ao propor que a forma mais adequada de tratar do assunto é através da possibilidade de imitação do comportamento humano. Harnad (2008) aponta que a questão de possibilidade de uma máquina pensante é uma questão extremamente relevante, porém constitui um problema de investigação que não possui solução:

O que nós queremos dizer por ‘pensar’ é, por um lado, o que criaturas pensantes podem fazer e como elas podem fazer essas coisas e, por outro lado, como o ato de pensar é sentido. O que pensadores podem fazer é capturado pelo Teste de Turing. Uma teoria de como eles fazem isso é dada pela forma como nossas máquinas fazem isso (se existem diversas máquinas bem sucedidas, é uma questão de inferência da

---

<sup>46</sup>É possível ler o relato de Linda Serck, juíza da competição de 2008, aqui: [http://www.bbc.co.uk/berkshire/content/articles/2008/10/12/turing\\_test\\_feature.shtml](http://www.bbc.co.uk/berkshire/content/articles/2008/10/12/turing_test_feature.shtml) e de Charlie Maloney, juiz da competição de 2017 aqui: <https://chatbotsmagazine.com/how-to-win-a-turing-test-the-loebner-prize-3ac2752250f1>

<sup>47</sup> Um *bot* é um programa de inteligência artificial voltado para conversação, também conhecidos como *chatterbots*. Falaremos especificamente sobre este tipo de software no decorrer deste capítulo.

<sup>48</sup> Em 2014 a administração do *Loebner Prize* passou para a *Society for the Study of Artificial Intelligence and Simulation of Behavior (AISB)*<sup>48</sup>. Alguns bots de conversação foram premiados com a medalha de bronze do *Loebner Prize* nos últimos anos, entre eles a Mitsuku, em 2016, desenvolvida por Steve Worswick, e o programador Bruce Wilcox, que desenvolveu quatro sistemas premiados. Disponível em <https://gearpatrol.com/2017/05/18/chatbots-loebner-prize-turing-test/>. Kirsten Nelson, 2017.



melhor teoria). Até aqui, nada sem sentido. Agora nós perguntamos: os candidatos que obtiveram sucesso no Teste de Turing realmente sentem como nós sentimos quando pensamos? Essa questão não é sem sentido, é apenas impossível de ser respondida – de qualquer outra forma que não seja sendo o candidato em questão” (HARNAD, 2008, *tradução nossa*).

Nessa passagem, Harnad explica o Problema das Outras Mentes: ao mesmo tempo em que é a certeza de pensar que nos constitui enquanto sujeitos (DESCARTES, 2009), não é possível afirmar isso sobre nenhum outro ser vivo (GUNKEL, 2012; 2017). Essa questão é relevante para pensarmos tanto o paradigma simulativo no campo da Inteligência Artificial, que se estabelece a partir de Turing, quanto a própria noção de que, uma vez que não podemos determinar se outros seres e/ou objetos pensam – uma vez que pensamento é um processo, e não um efeito -, podemos apenas tentar delimitar a existência de outra inteligência que não seja a nossa a partir de efeitos de pensamento que são similares aos efeitos do pensamento humano. Um destes efeitos é o efeito comunicacional: decorre da inteligência humana a capacidade de estabelecer comunicação com outros humanos a partir de um código comum. A relação entre o Teste de Turing e as habilidades comunicacionais são marcantes, uma vez que este é baseado na conversação (HAREL, 2000), e entrelaçam o campo da Inteligência Artificial com o da Comunicação desde o princípio (GUNKEL, 2012; 2017).

Em outras palavras, já que o pensamento inteligente não é diretamente observável, o máximo que podemos fazer é lidar com algo, como a interação comunicativa, que é assumidamente um produto da inteligência e pode ser empiricamente observada, medida e avaliada (GUNKEL, 2017, sem paginação)

Assim, para Gunkel, o que o Teste de Turing coloca, na verdade, é a necessidade de “demonstração de habilidade comunicacional” (2017, sem paginação). Para o autor (2012; 2017), existem duas características que decorrem da formulação da Inteligência Artificial a partir do jogo da imitação que possuem consequências que tanto envolvem quanto afetam a comunicação: a) o problema das outras mentes; b) aparência de inteligência.

As premissas A e B são, na verdade, duas faces da mesma moeda; o problema de não poder determinar se algo é inteligente, com base no problema das outras mentes, é justamente o que leva à formulação do jogo da imitação, ou seja, agir como um ser pensante é suficiente para que consideremos um ser pensante. Decorre disso que, é, de fato, impossível determinar a aparência de pensamento do pensamento real. No Teste de Turing, portanto, aparência de inteligência é o mesmo que inteligência, uma vez que não há forma de saber “com alguma certeza se eles de fato performam aquilo que parecem manifestar” (GUNKEL, 2017, sem paginação).

A partir disso, se estabelecem duas premissas relevantes para esse trabalho. A primeira diz respeito ao papel comunicacional como central para o jogo da imitação e, por extensão, aos desenvolvimentos em Inteligência Artificial que partem dessa perspectiva: a comunicação, nesse caso explicitamente a conversação em linguagem natural, é colocada tanto como forma de expressão de uma inteligência artificial, quanto ferramenta avaliativa para a noção de inteligência artificial.

O segundo pressuposto é de um caráter antropomórfico no campo da Inteligência Artificial estabelecido pela proposta de Turing ao colocar que uma máquina que pensa é aquela que age de forma análoga ao ser humano e da posterior emergência de um paradigma simulativo no campo – ponto ao qual retornaremos mais tarde. Ainda que o próprio Turing (1950) tenha considerado que a inteligência de uma máquina pudesse ser de uma ordem distinta da humana, as características do teste proposto por ele fazem com que a emergência do campo seja marcada pelas comparações entre homem e máquina e pela tentativa de criação de sistemas capazes de realizar uma mimetização de nosso próprio comportamento:

Não seria possível que as máquinas carreguem algo que possa ser descrito como pensamento mas que é muito diferente do que o ser humano faz? Essa é uma objeção muito forte mas, podemos dizer ao menos que, se, ainda assim uma máquina possa ser construída para jogar o Jogo da Imitação satisfatoriamente nós não precisamos nos atormentar com essa objeção<sup>49</sup> (p. 435, tradução nossa).

O paradigma simulativo estabelecido a partir daí é relevante no contexto deste trabalho, uma vez que as assistentes pessoais digitais são formas antropomórficas que propõe uma interface humanizada em aspectos como: a forma de utilização da linguagem verbal, a voz e a constituição enquanto personagens, que será abordada em subcapítulo teórico voltado para o assunto.

## 1.7 INTELIGÊNCIA ARTIFICIAL: PROBLEMATIZAÇÃO E CONCEITOS

Alguns anos após a publicação de Turing (1950), o termo inteligência artificial seria cunhado durante uma conferência em Stanford, por John McCarthy e outros colegas para uma

---

<sup>49</sup> “May not machines carry out something which ought to be described as thinking but which is very different from what a man does? This objection is a very strong one, but at least we can say that if, nevertheless, a machine can be constructed to play the imitation game satisfactorily, we need not be troubled by this objection.”

disciplina, formalizando o campo em 1956 (MERCHANT, 2017). Até aquele momento, alguns trabalhos individuais haviam sido feitos de forma isolada: é nessa conferência que tem início um esforço coletivo em torno da disciplina, que, na metade dos anos 60 já teria seus laboratórios especializados funcionando (TURKLE, 2005). No começo dos anos 70, o campo já possuía as principais características associadas à consolidação de uma área de conhecimento: publicações especializadas (através de periódicos e livros de base), cursos nas principais Universidades (principalmente dos Estados Unidos, onde se situavam os principais centros de pesquisa: MIT, Stanford e Carnegie Mellon) e congressos internacionais (TURKLE, 2005).

O âmbito da IA passa a ser definido como a área de estudo dedicada a dispositivos e sistemas computacionais criados pelo homem que possam ser feitos de forma a agir de uma maneira que consideremos inteligente (ISTVÁN, 1996), sendo que a inteligência, nesse caso, está relacionada à capacidade de racionalizar e responder como um ser humano (HAREL, 2000; 2004).

O campo de pesquisa da inteligência artificial possui três áreas fundamentais: conhecimento, aprendizado e dedução. Essas também são habilidades relacionadas àquelas expressas por um ser humano em uma conversação (HAREL, 2000; 2004). De acordo com Turkle (2005), o campo da Inteligência Artificial tem como ambição a construção de um sistema de propósito generalizado, tal como os seres humanos, não apenas focados em tarefas específicas.

Tanto o antropomorfismo na forma de pensar a Inteligência Artificial, quanto o intuito de gerar sistemas que funcionem de forma genérica, para inúmeras funções, colaboram para uma certa confusão entre a IA e a Modelação Cognitiva, campos do conhecimento distintos, mas que andam frequentemente entrelaçados (HARNAD, 2008).

O âmbito da Modelação Cognitiva tem como objetivo explicar os modos pelo qual a cognição humana funciona, ou seja, gerar modelos cognitivos capazes de descrever esse processo adequadamente – esses desenvolvimentos podem ser utilizados como subsídio para formulação de sistemas de Inteligência Artificial.

Entretanto, a IA não possui qualquer relação necessária com estes modelos, objetivando, apenas a criação de ferramentas que possuam boa performance – nem todos esses processos, entretanto, podem ser explicados e/ou compreendidos em detalhes, mesmo pelos engenheiros responsáveis por seus desenvolvimentos (HARNAD, 2008). Turing (1950) desentrelaça esses dois âmbitos ao antecipar a possibilidade de formulação e aplicação de métodos experimentais que gerem resultados positivos em uma máquina, sendo esta capaz de passar no Jogo da Imitação, mas cujos desenvolvedores não sejam capazes de descrever adequadamente seu modo

de funcionamento.

O início das investigações no campo da Inteligência Artificial é marcado tanto pelo otimismo quanto pela incredulidade. Em 1957, Herb Simon previu que em apenas dez anos o campo avançaria tanto a ponto de construir um computador capaz de ganhar um campeonato mundial de xadrez. Esse fato ocorreu apenas em 1996, quase 30 anos após essa previsão, quando o computador da IBM, o *Deep Blue*, venceu o então campeão mundial Garry Kasparov em uma partida (ISTVÁN, 1996). O filósofo Hubert Dreyfus, em contrapartida, afirma em 1965 que os avanços em desenvolvimentos de programas capazes de jogar damas era ilusório e que tal desenvolvimento seria inviável em jogos que exigem um tipo de raciocínio humano, como o xadrez.

Ainda que Deep Blue tenha vencido Kasparov, para o desgosto de Baudrillard (1999), é difícil desenvolver um programa que jogue xadrez de forma “perfeita, imbatível e imortal” (BAUDRILLARD, 1999, p. 120 e 121). Isso se dá pela própria dinâmica do jogo, no qual a ordem de magnitude da quantidade de jogadas possíveis é muito grande, tornando o cálculo computacional de todas as opções de movimento antes de uma jogada inviável em termos de tempo, uma vez que para cada jogada possível, as ramificações de cálculo seguinte se dividem a partir da técnica de árvore de decisões. Para diminuir o tempo de cálculo e possibilitar a tomada de decisão em tempo hábil para os termos do jogo, programas de xadrez em geral utilizam o método heurístico. A heurística pode ser caracterizada como um conjunto de regras não-rationais que os seres humanos utilizam para soluções de problemas cotidianos, que resultam na economia de tempo dessa ação. Um exemplo pode ser o processo de procura por algo que deixamos cair: como regra intuitiva, procuramos primeiro em um perímetro próximo do local onde estávamos, gradativamente expandido a abrangência da busca. Fazemos essa opção por acreditarmos que existe uma maior probabilidade do objeto cair próximo a nós, ou a uma distância que consideramos apropriada baseando-nos, por exemplo, na intensidade do movimento que levou o objeto a ser arremessado. Programas de computador que utilizam esse método possuem um conjunto de regras associadas que tornam possível a realização de alguns cálculos, mas não de todos: o método heurístico não pode gerar garantia de sucesso, justamente por não computar todas as possibilidades, porém possibilita a tomada de decisões em tempo hábil (HAREL, 2000; 2004).

É possível separar os desenvolvimentos em Inteligência Artificial em duas grandes linhas, cujas características se diferenciam principalmente no modo de estruturação e desenvolvimento destes sistemas, proveniente de entendimentos distintos do que caracteriza a inteligência e como esta poderia ser reproduzida em agentes não-humanos. São elas: IA

Simbólica e IA Conexionista, que serão brevemente apresentadas a seguir.

A Inteligência Artificial Simbólica se foca na produção de sistemas operacionais que possuem um quadro rígido de regras e formalização e que, quando trabalha a partir da perspectiva de um aprendizado da máquina, aposta em um nível maior de formalização que torne possível um aprendizado mais rápido, com um número menor de treinamento (LARGELY, 2011). A partir de uma teoria representacional da mente, a IA Simbólica domina o campo nas primeiras décadas, durante os anos 60, 70 e parte dos 80, tratando a inteligência como um dispositivo lógico que possui a capacidade de resolver problemas (PRIMO, 2011).

Um dos primeiros desenvolvimentos em Inteligência Artificial, que se enquadra nesta perspectiva, foi o Solucionador Geral de Problemas (GPS), criado por Simon, Shaw e Newell (1959), com o objetivo de simular a mente humana e gerar um protocolo para resolução de problemas, partindo de uma investigação do modo de tomada de decisão de humanos em determinadas situações. A ideia de problema aqui possui uma denominação mais abrangente do que geralmente empregada na linguagem coloquial, referindo-se a emergência de uma questão que faz necessária uma ou mais tomadas de decisão que levem a uma ação específica.

Os autores realizaram um estudo do processo mental a partir de seres humanos resolvendo um problema específico para coleta de dados que auxiliassem na criação de um passo a passo daquilo que o ser humano faz para solucionar um problema. Depois, partiram para a tentativa de fazer um programa que simulasse esse comportamento. Pessoas resolvem problemas de formas distintas, o que resulta em diversas possibilidades de programa para um único problema (ou seja, não existe um solucionador geral, uma forma única). A premissa deles é que havia uma forma geral, que deveria caracterizar como as pessoas pensam sobre um problema; o programa desenvolvido tentava capturar a maior porcentagem dos passos tomados por um humano para solução de um problema (TURKLE, 2005).

Uma das linhas de desenvolvimento que se destacam no contexto da IA Simbólica são os Sistemas Especialistas: são caracterizados como especialistas pois se concentram em um tipo de problema específico a ser solucionado, para o qual é possível criar um conjunto de regras mais rígidas e obter resultados de performance satisfatórios (TURKLE, 2005).

De forma distinta, a vertente da IA Conexionista compreende a inteligência a partir de suas características biológicas, tentando construir um sistema que funcione a partir da conexão de diversas unidades na criação de um conjunto capaz de gerar um output único, como o nosso cérebro. Ainda que tenha surgido de forma concomitante com a perspectiva simbólica, os desenvolvimentos em IA conexionista ganham força depois da metade dos anos 80 (PRIMO, 2011).

Dentre os desenvolvimentos desta linha se encontram as Redes Neurais Artificiais, que, ao invés de funcionarem com regras específicas, trabalham a partir de etapas, com níveis distintos de informação e aprofundamento, que vão sendo repassadas a partir de impulsos elétricos. Esse modelo computacional tenta simular a relação de transferência de informação entre neurônios no cérebro humano, tendo como vantagem a flexibilidade e capacidade de adaptação, facilitando o processo de aprendizado da máquina (HAREL, 2000; 2004). O funcionamento em rede permite um aprendizado a partir da recorrência, trabalhando de forma estatística, e depende de uma entrada elevada de dados para que o aprendizado da máquina possa ocorrer, ou seja, em termos gerais esse tipo de sistema demora mais para aprender, embora, ao fazê-lo, seja eficaz.

Os desenvolvimentos conexionistas focam na tentativa de construção de um mecanismo apropriado para armazenar, buscar e manipular informações, na tentativa de constituição de um sistema de memória. Se diferencia da IA Simbólica por não apostar na construção de um sistema formal específico, com um sistema de regras rígidas, mas sim com formulações que sejam mais soltas, atreladas a um banco de dados abrangente, a partir do qual o sistema seja capaz de traçar relações (MERCHANT, 2017). Exemplo de tal abordagem é o *Google Deep Mind neural network*<sup>50</sup>, que deixa de tentar colocar no sistema concepções anteriores, mas passa a focar na possibilidade de determinar padrões a partir de um grande banco de dados. Ou seja, ao invés de informar e descrever detalhadamente em um sistema o que é um cachorro, a abordagem de *machine learning* com base nas redes neurais, se foca em dados capazes de fazer emergir a noção do conjunto de padrões que forma um cachorro a partir da análise de um banco de dados.

Desde o princípio existe um forte entrelaçamento entre a IA e as ciências cognitivas<sup>51</sup>, principalmente considerando a IA clássica e em seus desenvolvimentos iniciais - já que a premissa básica das ciências cognitivas de ordem *behavioristas* é a noção de que a mente humana funciona de uma forma computacional (ISTVÁN, 1996). Portanto, descobertas de uma área poderiam beneficiar a outra, ainda que seja possível argumentar que esta é uma premissa que provavelmente gerou atrasos tanto na compreensão do que é a mente humana e como esta funciona, quanto no desenvolvimento e compreensão de sistemas inteligentes autônomos, que funcionam a partir de um modo de registro essencialmente distinto da lógica humana, conforme desenvolveremos posteriormente.

---

<sup>50</sup> Pode ser consultado aqui: <https://www.deepmind.com>

<sup>51</sup> Um aprofundamento sobre questões relacionadas ao pensamento humano, mente e corpo se fará necessário para o desenvolvimento futuro desse trabalho. Essa revisão deve integrar a próxima fase de nosso cronograma de pesquisa, com a continuidade de nossa pesquisa bibliográfica.

Dentre os principais empecilhos para um sistema de IA efetivo está o que se conhece como Problema do Conhecimento de Senso Comum (ISTVÁN, 1996) - coisas que são ordinárias para a maioria dos seres humanos, mas que precisam ser explicitamente programadas em um computador (entram aí ainda questões culturais e sociais, habilidades não desenvolvidas por computadores e que precisam integrar seu sistema de representação lógica). Diversas abordagens foram desenvolvidas na tentativa de solucionar esta questão: desenho de sistemas que operam em domínios restritos; criação de uma base de dados de conhecimento de senso comum, por Doug Lenat, que permitisse consulta de agentes inteligentes; e o que conhecemos hoje por *Machine Learning*, que parte da premissa que um sistema computacional inteligente deve ser capaz de aprender a partir da experiência, assim como os humanos. Uma das ramificações das abordagens de *Machine Learning* é o desenvolvimento das redes neurais artificiais, que foram bem sucedidas principalmente no reconhecimento de padrões (ISTVÁN, 1996).

A ideia do *Machine Learning* já estava prevista como central nos desenvolvimentos de sistemas inteligentes desde Turing (1950), que foi capaz de antecipar ramificações recentes decorrentes de implementação desse tipo como a imprevisibilidade que estes sistemas podem apresentar: “Uma característica importante do aprendizado da máquina é que seu instrutor será com frequência largamente ignorante do que está acontecendo dentro da máquina, ainda que ele seja capaz de prever, em alguma extensão, o comportamento do seu pupilo.<sup>52</sup>” (TURING, 1950, p. 457, *tradução nossa*).

A ideia da criação de sistemas inteligentes esteve, portanto, desde seu princípio, associada a uma noção de que, assim como seres humanos, estes sistemas precisariam passar por um treinamento para a aquisição de conhecimento. O sub-campo do *Machine Learning* não é fruto apenas de uma abordagem conexionista, tendo acompanhado as mudanças das matrizes de conhecimento do próprio campo da IA.

Considerado inicialmente como uma ramificação do campo da Inteligência Artificial, o início do desenvolvimento do aprendizado de máquina se focava em representações simbólicas do conhecimento (árvores de decisões, formulação lógica e produção de regras, por exemplo), deixando de lado as redes neurais e outras abordagens não-simbólicas - sendo inclusive deixadas de fora de chamadas de trabalho (LARGELY, 2011). No final da década de 50, por

---

<sup>52</sup> No original: “An important feature of a learning machine is that it’s teacher will often be very largely ignorant of quite what is going on inside, although he may still be able to some extent to predict his pupil’s behavior.” (TURING, 1950, p. 457)

exemplo, Arthur Samuel desenvolveu um programa para jogar damas programado para modificar suas próprias regras a partir das experiências do jogo, sendo um dos primeiros casos de sucesso da técnica do *Machine Learning* (TURKLE, 2005).

Essa ideia de *Machine Learning* do período condizia com o crescimento da AI simbólica e das ciências cognitivas, “com a maior parte dos pesquisadores preocupados com a construção de sistemas especialistas automáticos ou com a modelação das estruturas de aquisição de conhecimento humanas<sup>53</sup>” (LARGELY, 2011, p. 03, *tradução nossa*), persistindo até o final da década de 80. A preocupação era construir um sistema que fosse robusto, com um conhecimento anterior que permitiria um aprendizado rápido, devido ao seu formalismo, na tentativa da construção de uma taxa de aprendizado similar a do ser humano (LARGELY, 2011). A ideia central era a existência da necessidade de um sistema rígido, ou seja, de um modelo de fundo que fosse capaz de guiar o aprendizado da máquina adequadamente.

A área incorporou métodos utilizados em trabalhos empíricos da área da psicologia, relativos ao aprendizado humano, para formular experimentos de controle com estes sistemas automatizados, conforme proposto, por exemplo, por Kibler e Langley (1988). Em 1987, David Aha, então doutorando da UCI, coletou dados empíricos de estudos em *Machine Learning* e criou um repositório aberto para a comunidade acadêmica, onde era possível consultar e comparar resultados anteriores em uma determinada tarefa, permitindo avanço científico.

Já no final da década de 80 é que os desenvolvimentos realizados no âmbito do reconhecimento de padrões começam a ser incorporados ao âmbito do que se compreendia formalmente como o sub-campo do *Machine Learning*. A partir daí a década de 90 passa a ser dominada por esse campo que se volta para métodos que vem emprestados do campo do reconhecimento de padrões – inicia-se também nesse período o uso comercial de mecanismos de mineração de dados. Ocorre um distanciamento da ideia de conhecimento baseado na IA Simbólica, levando a um abandono de questões consideradas por Largely (2011) como mais complexas, tais como a compreensão linguística e a solução de problemas. Os principais desenvolvimentos de aprendizado de máquina passam a se focar em estatística e reconhecimento de padrões – devido a existência de um corpus de dados elevado, a taxa de aprendizado de uma máquina deixa de ser um fator crucial (LARGELY, 2011).

A partir dos anos 2000, o *Machine Learning* passou a ser tratado por muitos pesquisadores como uma disciplina separada, ainda que entrelaçada à IA. Largely (2011), acredita que o direcionamento excessivo para o uso de um grande conjunto de dados e um

---

<sup>53</sup> Texto original: “with most of these researchers being concerned with automatically constructing expert systems or modeling human acquisition of knowledge structures”



aprendizado baseado em estatísticas no campo do *Machine Learning* levaram a uma perda da complexidade das tarefas:

*Machine Learning* inicialmente se preocupava em desenvolver sistemas inteligentes que exibissem um comportamento rico em tarefas complexas, enquanto muitos pesquisadores contemporâneos parecem satisfeitos em endereçar problemas que não requerem nem inteligência nem sistemas. *Machine Learning* se focava na utilização e aquisição de conhecimento concebido como estruturas relacionais ricas, enquanto muitos pesquisadores agora parecem se importar apenas com estatísticas<sup>54</sup> (LARGELY, 2001, p. 278, tradução nossa)

O conceito de *deep learning*, por exemplo, existe há décadas no campo, e data de 1989. Mas é apenas a partir de 2010 que ele se torna interessante, devido ao aumento do poder de processamento de computadores que possibilita a redução do trabalho que durava semanas para alguns dias. Camadas múltiplas possibilitam um aprofundamento do machine learning – quanto mais dados são fornecidos, mais acurado é o resultado, mesmo que nunca atinja 100% de precisão. Essa técnica é extremamente relevante para o desenvolvimento e aperfeiçoamento de tecnologias envolvidas nas assistentes pessoais digitais, como o reconhecimento de linguagem falada e a tradução entre idiomas (DEVLIN, 2018).

Atualmente, as ferramentas que utilizam Machine Learning, principalmente a partir da linha da IA conexionista, atuam em frente variadas, tais como o reconhecimento facial, separação automática de fotos por categorias ou busca de fotos por palavras-chave, análise de imagens para diagnóstico médico, compra e venda de ações na bolsa de valores e previsões de campeonatos da Copa do Mundo<sup>55</sup>. Entretanto, independente dos potenciais do deep learning

Um computador julgando nossas fotos ainda não sabe o que é um gato. Pode encontrar um gato em uma imagem, mas não possui um conceito de sua existência. Não possui um entendimento de um animal “gato”. É esperto o suficiente para apontar a diferença entre animais diferentes, placas de trânsito, formas distintas de som e línguas diferentes, mas não possui entendimento destas coisas – apenas encontrou um conjunto de regras para que seja possível classificá-las. É apenas matemática<sup>56</sup> (DEVLIN, 2018,

---

<sup>54</sup> No original: “Machine Learning was originally concerned with developing intelligent systems that exhibited rich behavior on complex tasks, while many modern researchers seem content to tackle problems that do not require either intelligence or systems. Machine Learning focused initially on using and acquiring knowledge cast as rich relational structures, while many researchers now appear to care only about statistics.” (LARGELY, 2001, p. 278)

<sup>55</sup> Em 2018, um programa baseado em Machine Learning, a partir do método de *random-forest*, realizou uma previsão (equivocada) sobre os prováveis ganhadores da Copa do Mundo de 2018, cuja probabilidade apontava para a Espanha (no início do torneio) ou Alemanha (caso essa passasse das quartas de final): [https://www.technologyreview.com/s/611397/machine-learning-predicts-world-cup-winner/?utm\\_source=facebook.com&utm\\_medium=social&utm\\_campaign=owned\\_social](https://www.technologyreview.com/s/611397/machine-learning-predicts-world-cup-winner/?utm_source=facebook.com&utm_medium=social&utm_campaign=owned_social)

<sup>56</sup> No original: “a computer judging our pictures still doesn’t know what a cat is. It can spot one but it has no concept of its existence. It has no understanding of the animal ‘cat’. It is smart enough to tell the difference between different animals, different road signs, different speech sounds and different languages but it doesn’t understand

p. 115)

Como este tipo de tecnologia, na maioria das vezes, cria parâmetros para delimitar o que é um cachorro, por exemplo, a partir de um conjunto de imagens, esses parâmetros podem se associar a outros elementos presentes em uma imagem, que não necessariamente compõem um cachorro. Esse é um erro frequente do processo de indução do *machine learning*, onde parâmetros equivocados são estabelecidos no reconhecimento de um padrão. A imagem abaixo apresenta, a partir da sistematização do sistema Android para imagens de carro.

A figura 1 aponta, por exemplo, que o sistema em questão compreende como um dos componentes da categoria carro, elementos que na verdade são da categoria rua mas que com frequência estão associados com fotos de carros. Assim, toda vez que uma foto mostra uma rua, ele entende que ela é da categoria carro, uma vez que a maioria das imagens de treinamento para carro possuem, de fato, ruas. Nesse caso, por exemplo, o que é um carro para uma máquina e um carro para um humano são coisas distintas, o que impacta nas formas de ação e nos efeitos de ações mecânicas performadas a partir desta ideia de carro. O exemplo é simples e do qual não decorre aparentemente problema algum, entretanto, a utilização sistemática de técnicas como essa para tomada de decisão e predição de futuro podem ser complexas e precisam ser supervisionadas e adotadas com cuidado.

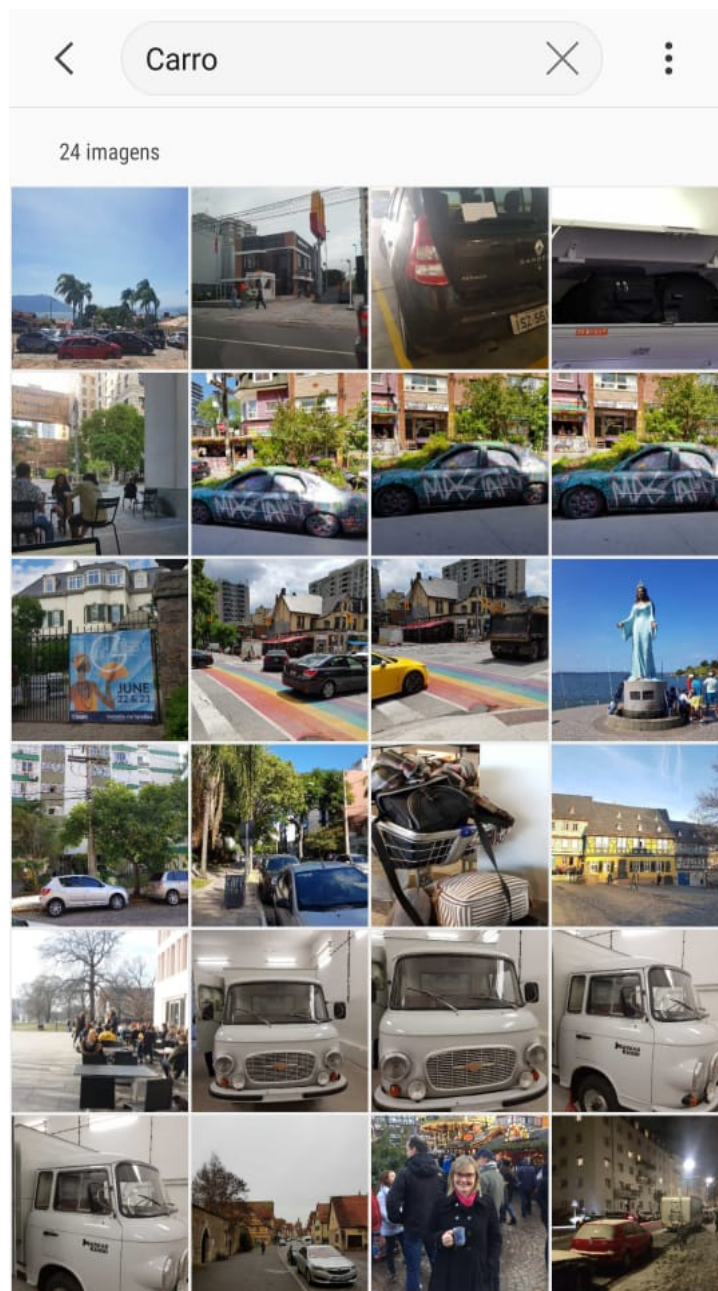
Para Grubber, chefe de desenvolvimento da assistente pessoal da Apple, a Siri, os sistemas de Inteligência Artificial que possuímos hoje, inclusive as assistentes pessoais virtuais, atrelados a formas de *Machine Learning* como as redes neurais profundas, funcionam de formas desconhecidas mesmo para especialistas: “ninguém tem realmente ideia do que os modelos sabem ou que sentido eles querem passar; eles apenas performam de uma forma que esta de acordo com os objetivos de funcionamento de um set de treinamento”<sup>57</sup> (MERCHANT, 2017, p.234). Assim, a questão principal é que esses sistemas na verdade não estão simulando como nós pensamos (e portanto não seriam inteligência artificial no sentido do paradigma simulativo), pois partem de processos não análogos aos dos seres humanos. A capacidade de encontrar padrões não previstos pelo ser humano e traçar relações com base em banco de dados pode trazer também resultado imprevistos e que não podem ser explicados por seus desenvolvedores, levantando questões éticas – essa discussão será retomada em seções subsequentes deste trabalho.

---

these things – it has merely worked out the rules to classify them. It presumes a stable world with a certain degree of consistency. It’s all just mathematics”.

<sup>57</sup> No original: “no one really has any idea of what the models know or what they mean; they just perform in a way that meets the objective function of a training set” (MERCHANT, 2017, p.234).

Figura 1 O que é um carro para o deep learning?



Fonte: printscreen de categorização de fotos da autora pelo Google.

## 1.8 AS POLÍTICAS DOS ARTEFATOS

Em 1966 surgiu, nos Estados Unidos, a primeira robô de conversação – ELIZA<sup>58</sup>. Criada

<sup>58</sup> O caso da ELIZA será tratado com maior profundidade durante nossa exposição teórica, no capítulo três.

por Joseph Weizenbaum, professor de ciência da computação do MIT, ELIZA simulava o papel de uma terapeuta Rogeriana na interação com usuários que assumiam o papel de pacientes. Reza a lenda que, entre leigos e estudantes de computação, muitos usuários trataram a psicoterapeuta artificial como uma profissional com a qual pudessem conversar abertamente sobre questões pessoais – e que a assistente pessoal do pesquisador pediu para que esse se retirasse da sala para que ela pudesse dialogar com a ELIZA com mais privacidade (TURKLE, 2005). A entusiasmada recepção dos usuários e da própria comunidade acadêmica levou o pesquisador a questionar sua própria carreira e se dedicar a desvelar a opacidade dos objetos digitais:

It is said that to explain is to explain away. This maxim is nowhere so well fulfilled as in the area of computer programming, especially in what is called heuristic programming and artificial intelligence. For in those realms machines are made to behave in wondrous ways, often sufficient to dazzle even the most experienced observer. But once a particular program is unmasked, once its inner workings are explained in language sufficiently plain to introduce understanding, its magic crumbles away; it stands revealed as a mere collection of procedures, each quite comprehensible. The observer says to himself “I could have written that”. With that thought he moves the program in question from the shelf marked “intelligent” to that reserved for curios, fit to be discussed only with people less enlightened than he (WEIZENBAUM, 1966, p. 36).

Nessa passagem, localizada no artigo no qual descreve o funcionamento de ELIZA, Weizenbaum aponta para a importância da compreensão do modo como um sistema funciona para que seja possível compreender também sua forma interacional, seu modo de agir no mundo. O esforço de Weizenbaum é louvável por pontuar que é no desconhecimento e na não-compreensão que a relação de cunho ‘mágico’ se estabelece – e que tal relação é perigosa quando tratamos de Inteligência Artificial.

A descrição do modo de operação, entretanto, não é suficiente para que possamos compreender os efeitos sociais destes sistemas, uma vez que tal questão não depende exclusivamente de seus atributos técnicos, mas da dinâmica estabelecida na aproximação de técnicas específicas de desenvolvimentos tecnológicos com padrões socioculturais. Não se trata, portanto, apenas de desmistificar os agentes de funcionamento autônomo, mas de investigar e compreender os mecanismos e formas de interação entre máquina e humano, buscando pistas nas micro-decisões (SPRENGER, 2015), automatizadas ou não, de um sistema, na sua recepção social, na forma como os indivíduos interagem com estes objetos e as características que lhes atribuem. Ou seja, para além de compreender os modos de funcionamento de um objeto de inteligência artificial, importa compreender a forma como esses operam (e são operados) em uma dinâmica social.

As formas de interação entre humanos e não-humanos vem sendo estudada no âmbito da Sociologia da Ciência e Tecnologia a partir de agenciamentos variáveis e sempre localizados nas relações e contextos sociais específicos. Ainda na década de 80, Langdom Wiener (1980) analisa as formas pelas quais artefatos podem ser dotados de formas de agir que incorporam pressupostos e ações políticas. Parte-se do pressuposto de que as máquinas, os sistemas e as estruturas da cultura material podem ser julgadas para além de suas capacidade de eficiência e produtividade, por incorporarem formas de poder e autoridade. A crítica subjacente é ao determinismo social da tecnologia, que pressupõe que os atributos específicos dos meios materiais não são em si importantes, dado que os seres humanos podem fazer com elas aquilo que desejarem – ou seja, são não mais do que ferramentas das quais dispomos. Ao contrário, Langdom argumento que existe uma adaptação constante, frequentemente e empiricamente observável das finalidades humanas para as formas técnicas disponíveis. O autor advoga, portanto, para a necessidade de se prestar atenção nas características dos objetos técnicos, especialmente no sentido que essas características adquirem no contexto social. Ou seja, para dar conta da complexidade do social em seus imbricamentos com o mundo material, não basta olhar apenas para um nem apenas para outro.

Ao falar de política o autor se refere a “arranjos de poder e autoridade nas associações humanas assim como em atividades que ocorrem nesses arranjos”<sup>59</sup> (LANGDOM, 1980, p. 122). Inicialmente, o autor (LANGDOM, 1980) divide essa possibilidade em dois formatos: instâncias em que a forma de elaborar, planejar ou arranjar um artefato técnico específico faz com os artefatos adquiram formas de ação políticas; e uma segundo que pondera que alguns artefatos trariam em si inevitavelmente em efeitos políticos. Ficaremos com a primeira, que é de interesse para este trabalho.

Através do estudo do caso das pontes em Long Island, Nova York, Langdom (1980) aponta para a produção deliberada por Robert Moses, construtor responsável pelo planejamento de diversos espaços públicos na cidade entre as décadas de 20 e 70, da diminuição do fluxo de habitantes negros e/ou de baixa renda em certos locais da cidade. De que forma? Projetando pontes que eram muito baixas para permitir a passagem de ônibus, tornando possível o acesso apenas através de carro (e, conseqüentemente, daqueles que possuíam meios financeiros para tanto). Assim, efeitos específicos podem ser obtidos através do designe do arranjo de artefatos específicos, dotando-os de um caráter político.

---

<sup>59</sup> No original: “arrangements of power and authority in human associations as well as the activities that take place within those arrangements”.

Nesse sentido, artefatos do nosso cotidiano não são apenas instrumentos técnicos, mas estão intrinsicamente relacionados aos processos culturais e ideológicos das estruturas nas quais suas formas de produção estão inseridas, especialmente quando atravessados por recortes de classe social, raça e/ou gênero. Estudos nesse sentido podem auxiliar a “expandir o entendimento do poder ideológico presente no design das tecnologias, como mecanismos de buscas, que são integrados (e integrais) a nossa experiência cotidiana como usuários” (SWEENEY, 2013, p. 06), além de possibilitar uma compreensão das formas pelas quais “política e cultura estão sempre inscritas nos sistemas tecnológicos” (SWEENEY, 2013, p. 7). A força política destes sistemas é tanto mais forte quanto maior a naturalização do uso da tecnologia em questão, uma vez que tecnologias socialmente integradas, sobre as quais não nos questionamos, podem operar como “uma força organizacional invisível na sociedade” (SWEENEY, 2013, p. 9). Assim, nossos artefatos cotidianos são tanto estruturados pelos moldes sociais dominantes quanto se tornam formas de estruturação e perpetuação de determinados moldes.

Ao longo deste trabalho veremos, por exemplo, como as teorias do paradigma de design de interface e da computação dos Computadores como atores sociais (CASA), desemboca em escolhas de design nas assistentes pessoais, com o objetivo, por exemplo, de aumentar o nível de confiança dos usuários com as empresas e de estabelecer relações com um grau maior de envolvimento (e de replicação de ‘normais sociais’). As assistentes pessoais digitais, por um lado, são pensadas a partir de um conjunto de pressupostos que operam política e socialmente: através de escolhas deliberadas, conscientes ou não. Por outro, a própria natureza do modo de funcionamento da computação, a partir da qual se originam, deixa uma marca por possuir limitações e potencialidade inerentes. Ainda que parte dessas limitações possam ser temporárias – ou seja, derivam de um estado tecnológico atual – outra parte decorre do fato de serem máquinas matemáticas e, portanto, estarem fadadas a um funcionamento sempre a partir dessas premissas (o que não é nem bom nem ruim, apenas é).

Mais contemporaneamente, um dos autores que propõe a inclusão dos aparatos técnicos nos estudos das ciências sociais é Latour (1992), situando-os como os elementos faltantes para que os sociólogos sejam capazes de recompor o social adequadamente em seus trabalhos. Parte de sua teoria se baseia na ideia de que os agentes não-humanos ocupam lugares na sociedade e executam ações (por vezes ações em substituição ao trabalho humano) e agenciam, com sua ação prescritiva, outras ações de agentes tanto humanos quanto não-humanos. Assim, os agentes não-humanos tanto agem quanto fazem agir, tornando inviável, no contexto social contemporâneo, qualquer tentativa de segregação entre o humano e o tecnológico.

A proposta de Latour é que a sociologia se beneficiará ao “tirar nossa atenção exclusiva para dos humanos e olhar também para os não-humanos”<sup>60</sup>(LATOURE, 1992, p. 153). O processo de delegação de ações para não humanos, exemplificado no texto a partir do objeto porta, modifica processos disciplinatórios e formas de agir de humanos e de não-humanos. Para o autor, os processos de delegação ocorrem em diversos âmbitos da vida social, uma vez que, “nós fomos capazes de legar para não humanos não apenas a força, que é conhecida há séculos, mas também valores, deveres e éticas” (LATOURE, 1992, p.157). Entretanto, olhar apenas para os não-humanos não é suficiente, já que “olhar apenas para os mecanismos é como assistir apenas metade de uma quadra em uma partida de tênis; parece uma sequência de movimentos sem sentido”<sup>61</sup>(LATOURE, 1992, 169). Ou seja, é necessário olhar para as relações estabelecidas entre um e outro e para as mútuas prescrições de comportamento que ocorrem a partir destas relações.

Para que não-humanos maquínicos possam executar ações no lugar de humanos, as máquinas possuem prescrições: modos de ação programados para determinados contextos, que são construídos, claro, por agentes humanos – como engenheiros e programadores, por exemplo. As prescrições de uma máquina podem ser mais ou menos restritivas (em relação as ações que os humanos podem executar) e, por causa das prescrições (ou seja, dos comandos que uma máquina deve executar e como), as restrições embutidas na máquina podem levar a discriminações, uma vez que nem todos os humanos irão se enquadrar no perfil do ‘usuário’ imaginado ou no cenário de respostas imaginadas. Para além das discriminações, também nos deparamos com outra questão: nem sempre é possível, a partir das prescrições, prever todas as situações pelas quais aquele sistema pode ser submetido – conforme a complexidade das ações que delegamos aos equipamentos tecnológicos aumenta, aumenta também a possibilidade de que as prescrições criadas não sejam suficientes para abarcar todos os dilemas latentes (LATOURE, 1992). Necessário ressaltar que as prescrições são elaboradas para que a máquina possa agir de uma determinada forma mas que, ao agir dessa forma, ela também prescreve aos humanos as formas de comportamento que são ou não possíveis ou aceitáveis naquele âmbito. Elas constantemente nos dizem, através de ações e de efeitos (por vezes, a recusa da ação) “faça isso, faça aquilo, se comporte dessa forma, não vá lá”<sup>62</sup>(LATOURE, 1992). Justamente por isso veremos, ao longo das análises relacionadas ao modo de falar com as assistentes (que é o modo

---

<sup>60</sup> No original: “have to turn our exclusive attention away from humans and look also at nonhumans”.

<sup>61</sup> No original: “looking at the mechanism alone is like watching half the court during a tennis game; it appears as so many meaningless moves”.

<sup>62</sup> No original: “do this, do that, behave this way, don’t go there”.

de comandar), a própria fala humana acaba sendo prescrita, ou seja, precisa ocorrer de uma determinada forma (em estrutura, volume, intensidade, velocidade) para que a interação seja efetiva.

Entretanto, desenvolvimentos de inteligência artificial de uma perspectiva emergente (que criam padrões a partir de bancos de dados) possuem prescrições muito menos específicas no momento da programação, tornando suas formas de ação para com aos agentes não-humanos mais opacas. É nesse sentido que eles colocam questões relacionadas à autonomia maquínica e ética: um sistema que é capaz de aprender a partir de um banco de dados desenvolverá formas de ação que, ainda que baseadas em prescrições, podem ser imprevisíveis, justamente por suas prescrições serem mais difusas. Podemos somar à impossibilidade de previsão do modo de funcionamento destes sistemas a ausência de transparência e a crescente utilização desse tipo de técnica em tomadas de decisão que envolvem a vida humana:

Another area of concern is the transparency around decisions made about us as more and more of these decisions are handed over to machines. Many current AI technologies are black boxes, unable to explain how they come to particular decisions. For example, one of the most fashionable and successful AI technologies currently is Deep Learning. This has been used in tasks as diverse as detecting skin cancer, pricing insurance and predicting crime. But Deep Learning cannot provide a good explanation for its decisions. Deep Learning uses a complex network of “artificial” neurons, one triggering another. In addition, how this network is connected and behaves depends on the massive amount of data used to train the network (WALSH, 2015, p. 6)

A ação pervasiva dos agentes de Inteligência Artificial contemporâneos, inseridos em uma gama de aplicações que muitas vezes escapa nossa percepção pela opacidade das tecnologias digitais, está em atores “capazes de aprender e aplicar essa inteligência em uma gama variada de tarefas: alguns como robôs capazes de agir em nosso mundo físico e social e alguns como agentes de software que tomam decisões em frações de segundos, controlando partes consideráveis da economia e de nossas vidas cotidianas”<sup>63</sup> (BURTON et al., 2017, p. 1).

O caso quase anedótico da inteligência artificial da Microsoft, nomeada como *Tay*, liberada para funcionamento no Twitter no começo do ano passado, pode servir de ilustração para a forma como desenvolvimentos baseados em um conjunto de dados podem produzir efeitos não previstos. Em cerca de 24 horas de atuação no microblog, *Tay*, programada a partir de uma base de dados filtrada da web de forma anônima, literalmente aprendeu a se expressar a partir de nossas próprias ações na internet. O resultado foi desastroso: uma sequência de

---

<sup>63</sup> No original: “to create AIs capable of learning and applying intelligence to a wide variety of tasks: some as robots able to take action in our physical and social world, and some as software agents that make decisions in fractions of a second, controlling huge swaths of the economy and our daily lives” (BURTON et al., 2017, p. 1).



postagens que incluíam racismo e homofobia fizeram com que a Microsoft rapidamente desativasse a robô. Para Walsh (2018), “ao colocar a Tay na internet, a Microsoft cometeu alguns erros fundamentais. Eles deveriam ter colocado um filtro de profanidade tanto no input coletado quanto no output da Tay. E não deveriam ter deixado Tay aprender a partir do Twitter sem checar”<sup>64</sup> (sem paginação).

A lição aqui é clara: dados podem criar distorções de realidade – e uma crença excessiva não apenas nos bancos de dados, mas nas ferramentas automatizadas capazes de encontrar padrões nestes dados, pode levar tanto a desenvolvimentos de tecnologias repletas de vieses que ignoramos pela impossibilidade de prever a totalidade dos resultados, quanto a posições epistemológicas como a de Anderson (2008), que em seu artigo “The end of Theory”, propõe não apenas o fim da teoria frente aos avanços de softwares analíticos, mas o fim do pensamento reflexivo. Nesse caso, não é a máquina que deve simular a inteligência humana, mas o próprio ser humano que deve abrir mão de suas características por uma realidade plenamente computável, em todo e qualquer aspecto.

Neste capítulo, apresentamos o referencial teórico a partir do qual desenvolvemos o problema de pesquisa e o desenho metodológico para esta pesquisa. Entretanto, não pretendemos nos filiar extensivamente a nenhuma das linhas apresentadas, ainda que utilizemos conceitos e noções destas que operacionalizam uma compreensão das formas de ação e de interação das e com as assistentes pessoais digitais. Ao longo deste trabalho, nos focaremos nas formas de agência das assistentes pessoais digitais no âmbito social a partir de cinco fatores que consideramos, a princípio, relevantes para a compreensão de como elas agem e fazem agir nas interações com humanos. Estes serão apresentados ao longo do próximo capítulo, de cunho metodológico.

---

<sup>64</sup> No original: “in putting Tay onto the internet, Microsoft made a number of fundamental mistakes. They should have put a profanity filter on the input and output of Tay. And, they should not have left Tay learn from the twitter sphere without any checks” (WALSH, 2018, sem paginação).

## 2. DESENHO METODOLÓGICO

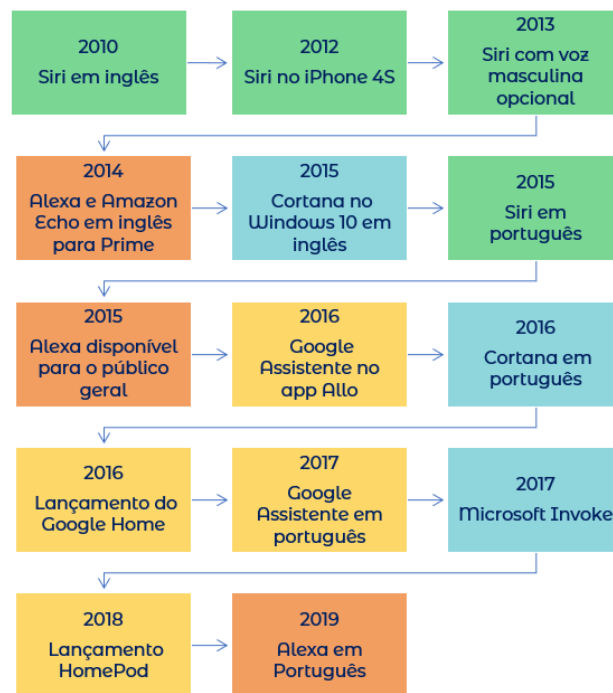
Neste capítulo apresentaremos brevemente os objetos empíricos da pesquisa – *Siri*, *Cortana*, *Google Assistente* e *Alexa* -, para em seguida descrever o desenho metodológico, as escolhas tomadas durante este trabalho e os procedimentos realizados. Por fim, apresentaremos dados de perfil dos participantes das etapas quantitativa e qualitativa da pesquisa.

### 2.1 CONHECENDO AS ASSISTENTES PESSOAIS DIGITAIS

As assistentes pessoais digitais utilizadas nessa pesquisa atuam a partir da interação por voz, tanto para input quanto para output. Podem ser acessadas através de hardwares variados, como o celular, o computador, os tablets, os smartspeakers e outros objetos inteligentes. Quando acessadas em dispositivos que possuem telas, podem também interagir através da linguagem natural escrita. Suas funcionalidades se relacionam ao hardware de acesso, com as contas associadas (como e-mail ou sistemas de streaming) e com o conjunto de produtos de cada uma das empresas. Atualmente quatro grandes empresas americanas possuem versões na língua portuguesa: *Apple*, *Google*, *Microsoft* e *Amazon*.

A linha do tempo abaixo mostra os principais desenvolvimentos relacionados às assistentes pessoais digitais das quatro empresas trabalhadas, a partir da perspectiva de análise deste trabalho.

Figura 2 Linha do tempo das assistentes pessoais digitais



Fonte: a autora

Nota-se uma maior concentração de lançamentos e desenvolvimentos das empresas entre os anos de 2015 e 2017, o que aponta para uma corrida no desenvolvimento de interfaces de voz no Vale do Silício, ainda que a primeira das assistentes, Siri, date de 2010.

Nenhum dos sistemas apresentados, mesmo nas versões do *smartspeaker*, vem com manuais de instruções que forneçam informações para além das configurações necessárias para iniciar a utilização (como fazer o setup da conta e conectar o dispositivo na internet, etc). Quando perguntadas, as assistentes fornecem algumas opções daquilo podem fazer pelo usuário, mas ainda assim a lista é restrita e nem sempre inclui as formas como os comandos devem ser enunciados. Nesse sentido, a utilização das assistentes depende muito da disponibilidade do usuário de aprender (através do uso, da pesquisa na web e da troca com outros usuários) o que é possível fazer e como fazer. Na sequência apresentaremos cada uma delas.

### 2.1.1 E AÍ, SIRI

*Siri* é o aplicativo de assistente pessoal digital da *Apple*, a primeira a ser lançada no mercado. Inicialmente criada e disponibilizada em 2010 por uma startup como um aplicativo a ser baixado na *AppStore*. De acordo com Tom Gruber (em entrevista para MERCHANT, 2017), um dos desenvolvedores da *Siri*, a assistente chegou ao topo de sua categoria em apenas um dia. A assistente foi rapidamente comprada pela *Apple* por duzentos milhões de dólares e posteriormente incorporada aos novos produtos da marca, a partir de 2012, passando a ser de uso exclusivo nos hardwares da *Apple*, disponível apenas no idioma inglês. O primeiro dispositivo da *Apple* a vir com a *Siri* de fábrica foi o iPhone 4S.

Durante os dez anos de existência, a interface gráfica da assistente mudou de forma significativa, saindo de um modelo inicial baseado em caixas de diálogo (NASRABADI, 2018) para um desenho mais ou menos assim:

Figura 3 Imagem da Siri no iPhone



Fonte: Raízes FM<sup>65</sup>

Em 2014 a assistente é ligada em uma rede neural e passa a utilizar técnicas de *machine learning* e de redes neurais profundas, gradualmente melhorando sua performance (MERCHANT, 2017). A *Siri* possui opção do idioma português para interação desde 2015 e voz masculina como opção desde 2013 (ainda assim, a feminina segue sendo padrão). Em 2018

---

<sup>65</sup> Pode ser visto em: <https://raizesfm.com.br/funcionarios-da-apple-podem-ouvir-conversas-e-ate-casais-fazendo-sexo-pela-siri-do-iphone/>

a empresa lança o *smartspeaker* com a assistente, conhecido como *HomePod*<sup>66</sup> – na versão em inglês, é possível ter usuários múltiplos na casa com reconhecimento de voz de cada um deles.

A *Siri* pode ser acionada com a *wake word* “E aí, Siri” nos seguintes dispositivos: *iPhone*, *iPad*, *AirPods*, *Carplay*, fones de ouvido, *Apple Watch*, *Apple TV*, *HomePod* e *Macs*. O acionamento a partir desta modalidade significa que a assistente está sempre captando áudio a partir de seus microfones, aguardando a identificação da palavra que desperta seu acesso (esse modo de funcionamento se repete em todas as demais assistentes de voz analisadas). Alternativamente, a ativação da *Siri* pode ser feita através de botões em cada um dos artefatos. A assistente é a única que não pode ser baixada como um aplicativo em hardwares de outras marcas.

Como uma interface de voz, as funcionalidades da *Siri* podem variar em relação ao hardware de acesso: no computador, por exemplo, é possível buscar arquivos, abrir programas, fazer buscas na internet; nos dispositivos móveis, a utilização de geolocalização possibilita buscas que se relacionam a estabelecimentos e rotas nas proximidades; a integração com a *Apple TV* e o *Carplay* auxiliam a navegação por voz dos usuários, e etc. A assistente também realiza buscas na internet sempre que solicitada, quando a demanda depende de uma informação atualizada fora do hardware e quando *Siri* não sabe o que você está solicitando. Além disso, existe um conjunto de perguntas e repostas que podem ser feitas para a assistente, no sentido conversacional, que são em sua maioria invariáveis em relação aos hardwares de acesso.

### 2.1.2 EI, CORTANA

*Cortana* é a assistente pessoal digital da *Microsoft*, embutida em no sistema operacional *Windows*. Foi lançada a partir do *Windows 10*, em janeiro de 2015, para acesso via computadores e dispositivos móveis com este sistema operacional, no idioma inglês<sup>67</sup>. A *Cortana* é atualmente compatível com outros sistemas operacionais, como o *Android* e o *iOS*. Cerca de um ano depois, em 2016, a versão no idioma português fica disponível. Em 2015 é

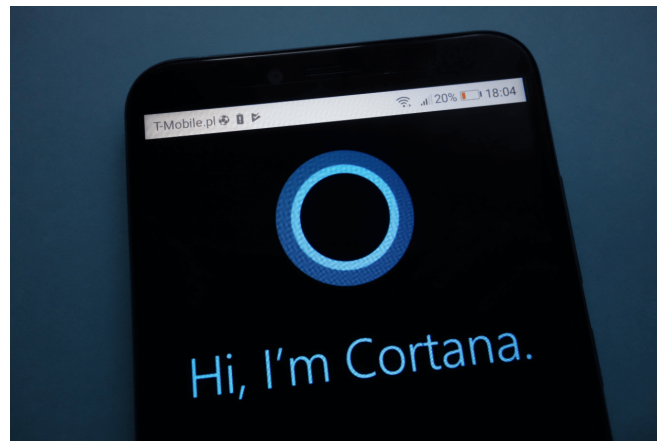
---

<sup>66</sup> O *HomePod* está disponível no mercado brasileiro, mas é o mais caro de todos os *smartspeakers*: no início de 2020 estava custando mais de 2.000,00 reais.

<sup>67</sup> Informação disponível em: <https://voicebot.ai/2017/07/14/timeline-voice-assistants-short-history-voice-revolution/>

anunciada uma futura voz masculina<sup>68</sup> que, até a publicação deste trabalho em 2020, não havia chegado ao mercado. O *smarspeaker* da empresa, chamado *Invoke*, é lançado em outubro de 2017, mas ainda não está disponível oficialmente para compra no Brasil<sup>69</sup>. A imagem abaixo apresenta a Cortana quando acessada pelo celular.

Figura 4 Imagem da Cortana no Celular



A assistente da *Microsoft* foi nomeada a partir de uma personagem de inteligência artificial que já fazia parte do universo ficcional da franquia *Halo*<sup>70</sup>. A personalidade de *Cortana* está atrelada, portanto, à personalidade da personagem da franquia, fato que aparece em respostas fornecidas pela mesma. Como personagem, *Cortana* aparece nos jogos como uma mulher, com características humanas e de acordo com os padrões de beleza predominantes no mundo ocidental, assim:

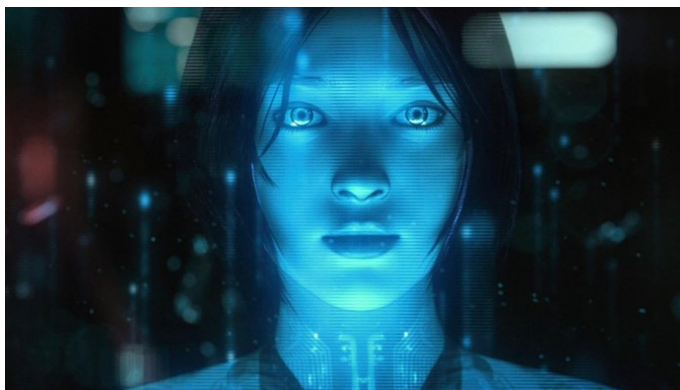
---

<sup>68</sup> Informação disponível em: <https://www.tecmundo.com.br/cortana/79272-cortana-windows-10-assistente-contar-voz-masculina.htm>

<sup>69</sup> O produto está a venda em sites como o Mercado Livre, por vendedores terceirizados. Durante nossa pesquisa, não encontramos nenhum usuário, no formulário quantitativo ou nas entrevistas, da Cortana pelo *Invoke* – um dos motivos pode ser justamente este. No Mercado Livre, no começo de 2020, o valor do *Invoke* saía mais ou menos 1.000 reais.

<sup>70</sup> *Halo* é uma franquia da mesma empresa, que possui uma série de games, livros e quadrinhos.

Figura 5 Imagem da Cortana no Celular



Fonte: Techtudo<sup>71</sup>

Entre as ações que a *Cortana* pode fazer, a partir do hardware de acesso, se encontram<sup>72</sup>: enviar lembretes que se baseiam em hora, local ou pessoas; enviar e-mails e mensagens de celular; gerenciar o calendário; criar listas; tocar músicas e podcasts; encontrar arquivos no computador ou celular e abrir aplicativos; conversar e jogar. Sempre que a *Cortana* não compreende a solicitação do usuário, oferece uma busca no *bing* como resultado. Como as demais assistentes, Cortana pode ser acionada através de um botão ou pela ativação da *wake word*, “Ei, Cortana”.

### 2.1.3 OK, GOOGLE

A *Google Assistente* foi lançada no início de 2016, na versão em inglês, e, no ano seguinte, já estava disponível em português, sendo a empresa mais rápida do mercado na tradução das assistentes para outros idiomas<sup>73</sup>. O antecessor da assistente, *Google Now*, que foi descontinuado, utilizava um sistema de input de voz para buscas online, somado a possibilidade de personalização através da utilização de *machine learning* (NASRABADI, 2018) – entretanto, não havia funcionalidades para além da pesquisa, nem output de voz. No final de 2016 a empresa lançou a versão *smartspeaker*, em uma linha chamada *Google Home* –

---

<sup>71</sup> Disponível em: <https://www.techtudo.com.br/noticias/noticia/2015/11/cortana-comeca-ser-testada-no-iphone.html>

<sup>72</sup> De acordo com o site da empresa: <https://support.microsoft.com/pt-br/help/17214/cortana-what-is>

<sup>73</sup> Esse processo foi provavelmente facilitado pelo sistema do Google Tradutor, disponível em diversos idiomas.

atualmente no Brasil existe a versão mini<sup>74</sup>. Desde 2018 a *Google Assistente* possui diversas opções de vozes, com gradações e gêneros diferentes, nomeadas por cores (voz verde, voz amarela, etc), quando a interação ocorre em língua inglesa. Em português, a voz segue sendo exclusivamente feminina.

Figura 6 Imagem da Google Home



Fonte: Techtudo<sup>75</sup>

A *Google Assistente* é compatível com outros sistemas operacionais além do Android e pode ser acessada a partir de um botão ou da *wake word* “Ok, Google”. Associa-se a todo o conjunto de produtos do *Google*, como o *ChromeCast*, e pode ser acessado a partir de um celular, da televisão e dos *smartspeakers* – além disso, integra-se a outros aplicativos como o *spotify* e o *whatsapp*. Na versão alto-falante, possibilita também a automação residencial, integrando-se a outros artefatos como luzes, cafeteiras, televisão, e o *Roomba*<sup>76</sup>. A *Google Assistente* envia periodicamente sugestões por e-mail aos usuários sobre funcionalidades adicionadas ao sistema e como utilizá-las.

As ações possíveis são similares ao demais assistentes e dependem do hardware de acesso: fazer ligações, acessar despertadores e agenda, criar listas, obter informações

---

<sup>74</sup> O Google Home mini, vendido no mercado nacional, é o mais barato dos *smartspeakers*, custando cerca de 300,00 no início de 2020.

<sup>75</sup> Pode ser conferida em: <https://www.techtudo.com.br/noticias/2019/07/google-confirma-que-funcionarios-tem-acesso-a-gravacoes-do-google-home.ghtml>

<sup>76</sup> *Roomba* é um aspirador de pó autônomo, que pode ser integrado tanto ao Google Home quanto a Alexa.



atualizadas, realizar buscas online, conversar e jogar. Além disso, existem ações e jogos temporários desenvolvidos em parceria com outras marcas para acesso através da *Google Assistente*, incluindo ações de marketing de produtos culturais. No *Google Home* também é possível criar rotinas despertadas por frases específicas: “Ok, Google, bom dia” pode despertar uma sequência de ações (padrão ou personalizadas), como ler notícias do dia, previsão do tempo e deslocamento para o trabalho.

#### 2.1.4 ALEXA

*Alexa* é a assistente pessoal digital da *Amazon*, disponível como aplicativo para celulares de qualquer sistema operacional, em alto-falantes da marca e integrada a outros objetos inteligentes. Foi inicialmente lançada para os sócios da *Amazon Prime* nos Estados Unidos em 2014 e ficou disponível para o público geral do mesmo país em 2015. A versão brasileira da *Alexa* ficou disponível em outubro de 2019. Até o momento possui apenas uma voz feminina, com possibilidade de troca apenas de idiomas e sotaques (principalmente no caso do inglês).

Atualmente existem três modelos de *smartspeakers* da *Amazon*, *Echo*, *Echo Dot* e *Echo Show*<sup>77</sup> – o primeiro é a versão em torre, o segundo a versão pequena e o terceiro a versão com uma tela, todos são atualmente comercializados no Brasil.

Figura 7 Imagem da Echo



Fonte: Show me Tech<sup>78</sup>

---

<sup>77</sup> A Amazon entrou com preços competitivos no Brasil, buscando ganhar mercado em relação ao Google Home, que já vinha sendo comercializado no país. No início de 2020, os preços dos modelos de smartpeakers variavam entre 300,00 e 700,00 reais, para versão a Echo Dot e Echo, respectivamente.

<sup>78</sup> Pode ser visualizada em: <https://www.showmetech.com.br/como-moderniza-sua-casa-com-a-amazon-echo/>

O modo de funcionamento da *Alexa* se difere em comparação com as outras assistentes.. Como a *Amazon* é uma empresa que não produz hardwares como celulares e/ou computadores (ao contrário da *Microsoft*, *Apple* e *Google*), a *Alexa* fica sempre associada a conta na *Amazon* do usuário. Essa conta possui algumas particularidades que impactam na utilização: ser baseada no país onde o usuário mora (como o *Kindle* faz) e ter acesso a um cartão de crédito vinculado a conta, o que permite fazer compras (mas apenas quando o país no qual a *Alexa* foi comprado e onde é utilizada forem os mesmos).

A assistente da *Amazon* inclui as funcionalidades das demais, além de realizar compras e checar encomendas, fazer ligações entre usuários que possuam o dispositivo, além de uma série de outras habilidades específicas. As habilidades específicas, que variam de acordo com o país de compra da *Alexa*, são as chamadas “*Alexa skills*” e podem ser desenvolvidas por terceiros. Em função disso, as funções da *Alexa* são bastante variadas, permitindo, por exemplo, que veículos de comunicação desenvolvam *skills* específicas para o seu conteúdo quando acessado através da assistente. O lado negativo é que, caso a *Alexa* tenha sido adquirida em outro país que não aquele onde o usuário mora, as *skills* disponíveis serão sempre as da região de compra, ainda que seja possível modificar o idioma de interação.

## 2.2 DAS ESCOLHAS METODOLÓGICAS

A construção de um desenho de pesquisa capaz de capturar um objeto do tempo presente, em constante transformação, tanto técnica quanto social, representa um desafio para o pesquisador. Quando esse objeto do tempo presente é, também, uma “novidade tecnológica”, é preciso que se estabeleça uma relação de contemporaneidade, nos termos de Agamben (2009): ser contemporâneo é tanto aderir quanto se distanciar de seu próprio tempo. Ao adotarmos a contemporaneidade como pressuposto metodológico, buscamos investigar as formas de agência das assistentes pessoais digitais, tanto como uma novidade tecnológica quanto como um episódio em uma série de continuidades e descontinuidades de desenvolvimentos e epistemologias no âmbito da computação e do design.

Demo (2009) desenha a pesquisa científica a partir de quatro gêneros abrangentes, entre os quais existe comunicação, ou seja, uma pesquisa pode ser formulada de maneira a

---

caracterizar o entrecruzamento de duas ou mais dessas inclinações: a) pesquisa teórica b) pesquisa metodológica c) pesquisa empírica d) pesquisa prática. Nossa proposta de pesquisa é um processo de diálogo entre o âmbito teórico e o empírico: o primeiro, busca a explicação de um fenômeno da realidade a partir de quadros teóricos de referência e o segundo é voltado para a codificação daquilo que pode ser observado na realidade social. Compreendemos que o processo de descrição passa também, no caso dessa pesquisa, por um movimento de explicação, ou seja, não apenas “constatar o que existe” (DEMO, 2009, p. 11), a partir de uma perspectiva empírica, mas também “desvendar por que existe” (DEMO, 2009, p. 11).

O problema de pesquisa que propomos, diz respeito tanto as formas de agência e de interação estabelecida entre os atores, quanto por explicações possíveis para o fenômeno observado. Conforme explicitado anteriormente, nossa pesquisa se baseia em três camadas analíticas que possuem a função principal de direcionamento, sendo estas: a) análise enquanto objeto técnico; b) análise enquanto objeto científico; c) análise enquanto objeto interacional. Estas camadas correspondem a forma como construímos o objeto de pesquisa aqui proposto, entendendo que compreender o objeto de pesquisa como algo que se constrói é compreender “a relação diversa entre sujeito e objeto, a começar pela sua problematização” (DEMO, 2009, p. 28), sendo muito mais da ordem do confronto entre pesquisador e objeto, do que do discurso formalizado.

O desenho metodológico proposto para esta pesquisa, portanto, precisa dar conta das especificidades de cada uma das camadas desenhadas com vistas à abordagem da interconexão – uma vez que é justamente o entrelaçamento entre as técnicas dos artefatos e as práticas culturais que nos interessa nesta investigação. Para isso, as seguintes etapas de pesquisa foram realizadas: a) pesquisa bibliográfica; b) análise documental; c) auto-observação da utilização das assistentes; d) pesquisa quantitativa; e) pesquisa qualitativa. Prosseguiremos para a descrição de cada uma delas.

O primeiro movimento desta pesquisa foi a investigação teórica em áreas de interesse, principalmente da Ciência da Computação, com foco na Inteligência Artificial, mas também no âmbito do Design e das Ciências Sociais. Com a revisão bibliográfica buscamos refinar o que compreendemos como técnicas e lógicas dos sistemas em questão: 1) a lógica computacional; 2) a noção e os desenvolvimentos de Inteligência Artificial; 3) as técnicas e mediações algorítmicas envolvidas nos desenvolvimentos das assistentes pessoais virtuais; 4) a abordagem das *redes neurais profundas* e do *machine learning*. Esses primeiros pontos de entrada nos auxiliaram a compreender, em termos de lógicas e processos, a forma pela qual esses

dispositivos são programados e pensados para agir no mundo, e compuseram o que entendemos pela camada técnica de pesquisa.

Essas chaves de leitura propostas se apresentaram, inicialmente, como um entrave no desenvolvimento desta pesquisa, por dois motivos. Primeiro, pela disciplinaridade acadêmica, que nos coloca o desafio de dar conta de elementos tão distintos e interdisciplinares sem perder o foco nas questões comunicacionais. Segundo, em função do método: como encontrar as marcas das programações técnicas nas formas interacionais socialmente estabelecidas com esses dispositivos? À primeira questão, respondemos com uma tentativa de apropriação de conhecimentos de uma disciplina distinta, a ciência da computação, através de leituras, aulas, eventos acadêmicos e participação em equipes multidisciplinares. Já a segunda, tentamos responder na forma do desenho metodológico aqui exposto, a partir da inclusão de duas etapas em particular, expostas na sequência: a) a auto-observação da interação com esses sistemas e b) entrevistas em profundidade com usuários.

Uma vez finalizada esta primeira etapa de pesquisa bibliográfica, iniciamos as etapas de pesquisa documental e de auto-observação. Durante esse período foi a continuidade da pesquisa bibliográfica, guiada pelas percepções geradas durante os processos de leitura documental e de auto-observação que nos permitiu reformular e refinar os conceitos centrais que guiam as análises, a partir das três camadas de pesquisa previamente apresentadas. Ou seja, nesse caso a pesquisa se apresenta como um diálogo entre suas etapas metodológicas, quando a execução de um determinado passo impacta tanto os passos seguintes quanto seus antecessores, compondo uma narrativa dos passos metodológicos e caminhos de pesquisa que é não-linear.

A etapa de análise documental contempla o material divulgado na mídia especializada em tecnologia e em revistas científicas especificamente sobre as assistentes pessoais digitais. Além disso, vídeos e publicações sobre as assistentes pessoais, compartilhados em redes sociais, também foram analisados e informam este trabalho. Tal pesquisa nos fornece subsídios interessantes, a partir de casos específicos, que utilizamos como material para o entrecruzamento das técnicas com os usos sociais, assim como informações e entrevistas que dizem sobre o modo como as assistentes foram idealizadas pelas empresas. O acompanhamento do material midiático sobre o assunto foi realizado de forma flutuante, a partir da leitura do material de interesse ao longo do período da tese e da seleção das notícias, colunas, artigos, comentários, fotos e vídeos que traziam informações relevantes ou potencial analítico para a tese. Entretanto, mesmo o material que não entrou no texto final desta tese foi importante para a compreensão geral do objeto de pesquisa.

O terceiro passo consistiu na auto-observação da utilização dos sistemas propostos, nomeadamente: *Alexa (Amazon)*, *Google Assistente (Google)*, *Cortana (Microsoft)* e *Siri (Apple)*. As auto-observações foram registradas com mecanismos utilizados na etnografia para documentação das interações no campo (GEERTZ, 1978)<sup>79</sup>, a partir de anotações sistematizadas das interações e registro de diálogos que fornecem percepções interessantes. As interações com a *Siri*, *Cortana* e a *Google Assistente* foram também registradas na forma de *printscreens*, quando a interação estabelecida era de especial interesse para a investigação. Além disso, o dispositivo da *Amazon* mantém um *log* de atividades realizadas no aplicativo utilizado para registrar e sincronizar as contas e dados do usuário com a assistente pessoal, o que facilitou o processo de registro das interações.

Ainda que os quatro dispositivos tenham sido utilizados durante o período da auto-observação, as assistentes *Siri* (pelo, celular, computador e iPad), *Alexa* (pelo *smartspeaker*) e *Google Assistente* (através do *Google Home* e do celular) foram utilizadas de forma mais extensiva. Antes de iniciarmos a auto-observação sistematizada, realizamos uma auto-observação exploratória com a *Siri* (no celular) para o projeto de qualificação. O período de auto-observação sistematizada se iniciou em outubro de 2018, com a compra da *Alexa* na Alemanha, durante o período do Doutorado Sanduíche. A partir daí, passamos a registrar também interações específicas com a *Siri*, assim como utilização do *Google Assistente* no celular e da *Cortana*, de forma menos intensiva, no computador. Após a aquisição do *Google Home*, versão em português, desde agosto de 2019 passamos a observar e testar também este dispositivo e suas funcionalidades para a casa, em comparação com a *Alexa*, de uso em inglês.

Através dos diálogos, nos interessava perceber as lógicas e técnicas que mapeamos enquanto conceitos teóricos na construção destes dispositivos: ou seja, o quanto a interface deixa perceber e/ou disfarça seu modo de funcionamento. Fica evidente, por exemplo, a construção destas assistentes enquanto um tipo de personagem, no sentido de construção de uma personalidade coesa que elas exprimem em suas interações, através de falas específicas e formas de interação. Uma camada adicional de complexidade ao problema de pesquisa deriva da constatação do tipo de personagem construído, caracterizado como feminino, não apenas através da voz, mas também de estereótipos de diálogo, ação e reação.

A auto-observação do uso destes objetos digitais seguiu até o final da escrita desta tese, sendo tanto as notas quanto os dispositivos consultados quando necessário. Entretanto, após cerca de seis meses do início da auto-observação sistematizada, delimitamos o que vem a

---

<sup>79</sup> Essas técnicas também são frequentemente utilizadas nas adaptações do método etnográfico para a internet (HINE, 2000; BRAGA, 2006; KOZINETS, 2007; FRAGOSO; AMARAL; RECUERO, 2012);

compor os cinco fatores de análise desta tese, em um cruzamento que vem da mistura entre a pesquisa teórica, a pesquisa documental e a auto-observação: 1) a lógica computacional; 2) interface e interação; 3) aspectos antropomórficos em objetos digitais; 4) voz e sua relação com o humano; 5) características de gênero. São estes cinco fatores de análise que irão guiar o desenvolvimento das perguntas para o formulário quantitativo online e para as entrevistas em profundidade, assim como dão origem, cada um deles, a um capítulo de análise desta tese. Estes elementos são aqueles que, ao longo desta pesquisa, compreendemos como sendo centrais na constituição da agência das assistentes pessoais digitais e nas formas de interação e percepção com/dos seres humanos e que passam, a partir daí, a nortear a discussão da tese e possibilitam a estruturação do formulário e da entrevista em profundidade.

O passo seguinte foi o desenvolvimento de perguntas de cunho quantitativo para um formulário aplicado<sup>80</sup> no formato on-line a partir da rede de contatos da própria pesquisadora, buscando expandir para contatos dos contatos. A partir de indicações e da manifestação de respondentes do formulário, alguns participantes foram selecionados para entrevistas em profundidade sobre seus usos e percepções das assistentes pessoais digitais. Priorizamos usuários com mais tempo ou frequência de uso, uma vez que estes poderiam trazer contribuições mais interessantes para a pesquisa. O roteiro para a entrevista foi criado de forma semi-estruturada, permitindo flexibilidade ao entrevistador e ao entrevistado durante as entrevistas. A etapa de entrevistas, por lidar com sujeitos, passou pelo Comitê de Pesquisa da Faculdade de Biblioteconomia e Comunicação e pelo Comitê de Ética da Universidade Federal do Rio Grande do Sul, sendo aprovado por ambos.

As perguntas criadas para cada uma destas ferramentas de coleta de dados foram pensadas visando obter respostas relacionadas aos cinco fatores de análise, que guiaram de forma transversal a análise dos dados obtidos em todas as etapas. Ou seja, o elemento analítico *interface e interação*, por exemplo, passa a ser atravessado por elementos de revisão bibliográfica, de pesquisa documental, de auto-observação e das respostas quantitativas e qualitativas a partir do formulário e da entrevista.

Dessa forma, tentamos garantir o entrelaçamento entre as camadas técnica, científica e social, que originaram o desenho teórico-metodológico desta pesquisa. Isso se dá uma vez que, cada elemento analítico, ainda que pertencente a um desses eixos, era atravessado por todas as etapas metodológicas da pesquisa e, de uma forma ou de outra, cada um desses elementos se relacionava com os outros de forma distinta. Ao analisar o conceito de *interface* enquanto

---

<sup>80</sup> O formulário completo pode ser conferido no Anexo 1 desta tese.

operante nas assistentes pessoais, nos deparamos, por exemplo, com questões que se relacionam com a *lógica computacional* (na forma como esses dispositivos compreendem ou não o que está sendo solicitado), com a *voz*, que é o que dá forma a essa interface e, conseqüentemente, com questões de *gênero e personalidade* que operam e atravessam essa interface, entre outros cruzamentos possíveis.

### 2.2.1 SOBRE O FORMULÁRIO PARA OBTENÇÃO DE DADOS QUANTITATIVOS

A versão final do formulário<sup>81</sup>, utilizado na coleta de dados quantitativos, contém um total de 26 perguntas, estruturadas da seguinte forma: um conjunto de 19 perguntas fechadas e obrigatórias, de múltipla escolha ou de caixa de alternativas; 03 perguntas abertas e opcionais, de caráter explicativo ou de complementação; 04 perguntas de perfil do respondente (idade, gênero, grau de escolaridade e área de formação). As perguntas fechadas buscavam respostas para as seguintes questões, estruturadas a partir das categorias de análise:

- *Informações gerais de uso*: assistentes utilizadas; frequência de uso; preferência; finalidade; locais de utilização;
- *Interação e interface*: dispositivos utilizados para interagir com a assistente; utilização como tecnologia assistiva; percepção sobre interface de voz; atributos que levam a utilizar a assistente;
- *Antropomorfismo*: acionamentos acidentais, sentimentos em relação ao dispositivo e desenvolvimento de uma relação com a assistente;
- *Voz e gênero*: percepção sobre uso de linguagem natural falada; percepção da sua forma de falar durante a interação; preferência por gênero da voz e percepção da interação mediada pelo gênero da voz;
- *Lógica computacional*: a assistente como máquina; preocupação com o modo de funcionamento e privacidade;

---

<sup>81</sup> O formulário completo pode ser conferido no Anexo 1 e o modelo do Termo de Consentimento Livre e Esclarecido assinado pelos participantes, no Anexo 2. A tabela com os dados brutos completos obtidos também está disponível: pode ser encontrada no Anexo 3 ou no endereço <http://bit.ly/dadosbrutosform>

A versão de testes do formulário foi enviada a três pesquisadores da área<sup>82</sup>, que ponderaram questões relativas ao número de questões, formato, gradação das respostas, escala de intensidade das respostas, agrupamento e formulação das perguntas, entre outras questões. Graças ao apoio desses colegas, o formulário adquiriu mais robustez e clareza e, ao mesmo tempo, foram otimizadas a obtenção de respostas, pela concisão. Ao trabalhar com formulários quantitativos, o pesquisador se depara com o dilema de tentar obter o máximo possível de informações de um único participante conjuntamente com o maior número possível de respondentes: formulários muito longos ou muito abertos podem gerar desistências e evasão dos participantes, por isso reduzimos o número de questões abertas e nos mantivemos nos tópicos que eram essenciais para este trabalho.

O formulário foi disponibilizado no Google Formulários e optamos por não exigir e-mail para acesso ao questionário, mantendo, ao máximo possível, o anonimato. Apenas participantes que tivessem interesse em participar da etapa qualitativa da pesquisa, com a entrevista em profundidade, foram convidados a deixar opcionalmente seus e-mails para contato. Como estratégia de divulgação, utilizamos mídias e plataformas digitais, acionando nossas redes de contatos on-line e off-line. Além disso, solicitamos a pessoas chave que auxiliassem deliberadamente a circular o formulário em grupos com perfis diferentes daqueles existentes na rede da própria pesquisadora. O objetivo era não apenas obter o máximo possível de respostas, mas também afastar os respondentes do formulário o máximo possível da região geográfica, social e cultural da pesquisadora.

A divulgação ocorreu, essencialmente, em quatro plataformas comunicacionais, *Twitter*, *Facebook*, *Whatsapp* e e-mail, além da forma presencial. Com o intuito de descentralizar as respostas ao formulário, no *Facebook*, além de publicação na timeline, mapeamos grupos de interesse para publicação, tais como: grupos voltados para discutir as assistentes pessoais digitais (sempre de origem brasileira), grupos de Instituições de Ensino e Universidades, grupos de estudantes e pesquisadores em ciência da computação, grupos de pesquisadores na área da Comunicação e da Ciência da Computação. Participantes de grupos brasileiros no *Facebook* sobre a *Alexa* e *Google Home* direcionaram a pesquisadora para grupos abertos sobre o mesmo tema no *Whatsapp* e no *Telegram*, onde a divulgação do formulário também foi feita.

Por e-mail, a estratégia foi semelhante: divulgação em listas de discussões de pesquisadores de Comunicação e de Universidades, da qual a pesquisadora já fazia parte. Além

---

<sup>82</sup> Ma. Gabrielli Tiburi, Ma. Maríla Gehrke e Dr. Marcio Telles.



disso, foram enviados e-mails solicitando divulgação para as secretarias de todos os Programas de Pós-Graduação da Universidade Federal do Rio Grande do Sul (cerca de 90 programas), visando amplitude de área de formação, e para todas as secretarias e coordenadores de Programas de Pós-Graduação em Comunicação do Brasil (49 programas<sup>83</sup>), visando uma melhor distribuição geográfica. A escolha de foco no público universitário ocorreu por julgarmos que esse público está mais familiarizado com pesquisas neste formato e é mais disposto a participar, além da possibilidade de, com um único e-mail, atingir uma quantidade grande de alunos. Essa decisão irá se refletir nos dados de perfil dos respondentes, que analisaremos em seguida, gerando um peso de nível de formação (Ensino Superior), área (da Comunicação) e também de idade, que são limitações desta pesquisa. Ainda, colegas, amigos e a rede pessoal de contatos de pesquisa em todo o Brasil auxiliaram na divulgação em suas próprias redes de contato, o que foi essencial para o sucesso da pesquisa. Cada uma destas ações de divulgação foi distribuída ao longo do período no qual o formulário ficou disponível, sendo possível mapear os efeitos da divulgação no retorno de respostas no Formulário do Google.

O formulário online para coleta de dados quantitativos ficou disponível para receber respostas entre os dias 28 de setembro e 02 de novembro, totalizando 6 semanas de formulário aberto. Neste período, 591 respondentes iniciaram a pesquisa e 501 concluíram.

Os motivos para a não finalização do formulário de 90 respondentes são: a) participantes que responderam de forma negativa à primeira pergunta, afirmando não terem utilizado nenhuma assistente pessoal virtual, totalizando 89 respondentes; b) participantes que responderam de forma negativa à segunda pergunta, afirmando que não autorizavam a utilização dos seus dados para a pesquisa em questão, totalizando 01 respondente. Já ter utilizado ao menos uma vez uma assistente pessoal virtual e concordar com a utilização, em caráter de conjunto e de forma anônima de seus dados, eram os únicos pré-requisitos para participar da etapa quantitativa da pesquisa. Como a) e b) eram, respectivamente, a primeira e a segunda pergunta do formulário, o mesmo foi desenhado para encerrar automaticamente em caso de resposta negativa em qualquer um dos casos. Iremos trabalhar, portanto, com o número de 501 respondentes.

Ao final da coleta dos dados quantitativos, com o formulário encerrado e o início da análise dos dados, percebemos que a questão de número 20 precisaria ser descartada da pesquisa em questão. A pergunta *“Você, alguém da sua família ou conhecido utiliza uma assistente*

---

<sup>83</sup> Utilizamos a lista disponível no site da Associação Nacional de Programa em Comunicação (COMPÓS): <https://www.compos.org.br/programas.php>

*peçoal como forma de tecnologia assistiva?”* obteve um número elevado de respostas “*sim*”. A pergunta seguinte, direcionada para quem respondeu de forma positiva, pedia “*Você pode explicar para qual finalidade a assistente pessoal é utilizada como tecnologia assistiva?*”: as respostas para esta segunda pergunta, entretanto, demonstram que a compreensão foi equivocada. Ao invés de compreenderem a utilização das assistentes como forma de auxílio para pessoas com necessidades especiais, muitos participantes compreenderam como a utilização das assistentes como auxílio para tarefas do cotidiano por parte de qualquer pessoa. Por esse motivo, os resultados não são confiáveis e foram desconsiderados – mantivemos apenas alguns exemplos escritos em que os participantes eram e/ou tinham familiares e amigos deficientes físicos que utilizam a interface como tecnologia assistiva.

Na próxima seção apresentaremos um detalhamento do formulário quantitativo, das condições de sua aplicação e circulação e do perfil dos respondentes.

## 2.2.2 DO PERFIL DOS RESPONDENTES DO FORMULÁRIO ONLINE

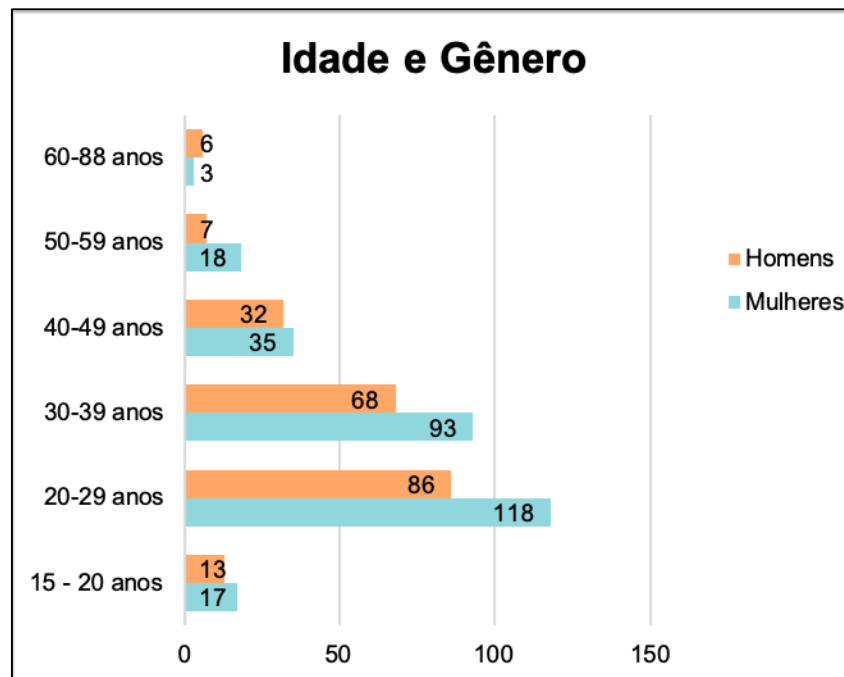
A maior parte dos respondentes do formulário online se encontra na faixa dos 20-29 anos, com 204 participantes, seguido por 30-39 anos e 40-49, com 161 e 67 participantes respectivamente. Na faixa dos 15-19 anos tivemos 30 respondentes e entre 60 e 88 anos<sup>84</sup>, 9. As faixas de maior número de respondentes se devem, parcialmente, ao foco da divulgação para um público universitário (20-29 anos) e na pós-graduação ou comunidade acadêmica em geral (40-49 anos e 50-59 anos). Ainda assim, vale ressaltar a participação de indivíduos mais jovens, assim como de participantes mais velhos. No que se refere a idade, cinco respostas foram consideradas inválidas, por não representarem um número.<sup>85</sup> Do número total de participantes, a pesquisa contou com respostas de 284 mulheres e 212 homens. O gráfico abaixo mostra a distribuição dos participantes por gênero e faixa etária.

---

<sup>84</sup> Destes, foram cinco respondentes entre 60-65 anos, três respondentes entre 70-75 anos e um respondente de 88 anos.

<sup>85</sup> Alguns participantes confundiram ‘idade’ por ‘cidade’, por exemplo, gerando respostas inválidas.

Figura 8 Gráfico Idade e Gênero dos Participantes

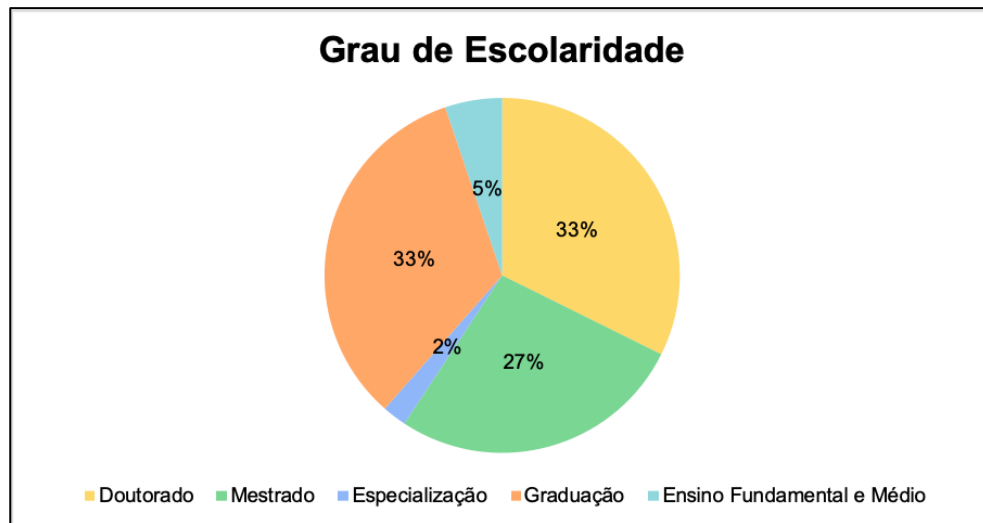


Fonte: a autora

A partir deste gráfico é possível visualizar que, de forma geral, as faixas etárias mantêm a proporção geral de um maior número de mulheres em relação ao número de homens. Apenas entre 60-88 anos que o número de respondentes homens ultrapassa o número de mulheres, correspondendo ao dobro deste. Esse dado se mostra interessante na medida em que nos concede pistas para pensar a adesão de novas tecnologias por idosos a partir de uma divisão por gênero, entretanto, o número de participantes é muito pequeno para nos permitir fazer inferências.

Dos 501 participantes nesta etapa, 95% já havia concluído ou estava cursando educação de nível superior, distribuídas entre graduação, especialização, mestrado e doutorado. Mais uma vez, este resultado é também, em parte, reflexo das escolhas na divulgação do formulário, focado no público universitário, tanto de alunos quanto de pesquisadores e professores. Ainda que o formulário tenha circulado em outros âmbitos, vale ressaltar que o Formulário deveria ser respondido por livre escolha dos participantes, de forma exclusivamente online, em que nenhum benefício fosse recebido em troca. Em função disso, compreendemos que o público universitário, por estar envolvido, em maior ou menor grau, com pesquisa, pode ser mais motivado e/ou solidário ao receber este tipo de solicitação. Essas questões se refletem também no alto índice de respostas de participantes que estavam cursando ou já haviam concluído cursos de mestrado e doutorado, que representam 27% e 33% dos participantes, respectivamente.

Figura 9 Gráfico Grau de Escolaridade dos Participantes



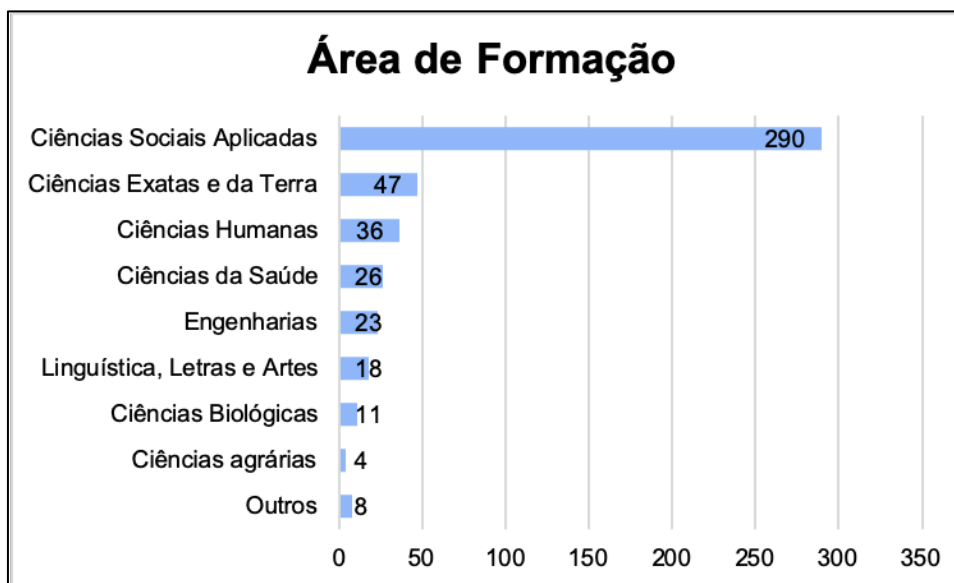
Fonte: a autora

O gráfico acima mostra a distribuição de níveis de escolaridade dos participantes: 33% dos respondentes estavam cursando ou já haviam concluído a Graduação; 33% estavam cursando ou já haviam concluído o Doutorado; 27% estavam cursando ou já haviam concluído o Mestrado; 5% estavam cursando ou já haviam concluído o Ensino Fundamental ou Médio e 2% estavam cursando ou já haviam concluído uma Especialização. No que diz respeito aos participantes no âmbito do Ensino Fundamental e Médio, apenas 01 respondente havia concluído apenas o Ensino Fundamental e se encontrava em idade adulta. Os demais, totalizando 25 respondentes, estavam cursando ou haviam concluído o Ensino Médio e se encontravam na faixa etária entre 15-20 anos, correspondente a idade escolar em questão.

Os participantes que informaram estar cursando ou já haver concluído curso de Ensino Superior, foram solicitados a informar, de forma opcional, a área do conhecimento dos seus estudos. Esse dado é interessante não apenas para verificar a concentração de respondentes em uma determinada área e o afastamento relativo em relação aos contatos iniciais da pesquisadora, mas também fornece subsídios para compreender as respostas fornecidas ao longo do questionário: por exemplo, pessoas formadas em ciências da computação terão um maior conhecimento técnico dos dispositivos de Inteligência Artificial, o que pode impactar sua percepção desses sistemas, suas formas de uso, entre outros. O gráfico abaixo apresenta a

distribuição dos participantes de Ensino de Nível Superior nas áreas de conhecimento a partir da tabela fornecida pelo CNPq<sup>86</sup>.

Figura 10 Gráfico Áreas do Conhecimento de Formação dos Participantes



Fonte: a autora

Observa-se, a partir deste gráfico, a predominância de respondentes pertencentes às Ciências Sociais Aplicadas, com 290 respondentes: nesta área do conhecimento é que se localizam os cursos de Comunicação, que corresponde à área de interesse e inserção desta pesquisa, assim como fez parte da estratégia de divulgação do próprio formulário, a partir dos Programas de Pós-Graduação em Comunicação. Em seguida temos as Ciências Exatas e da Terra, com 47 respondentes, no qual se localizam as Ciências da Computação, área de intersecção desta pesquisa e que também recebeu divulgação, principalmente através de grupos voltados para o assunto no *Facebook*. Depois disso temos Ciências Humanas com 36 respondentes, Ciências da Saúde com 26 respondentes, Engenharias com 23 respondentes, Linguística, Letras e Artes com 18 respondentes, Ciências Biológicas com 11 respondentes e Ciências Agrárias com 4 respondentes. Ainda que exista uma predominância das Ciências

<sup>86</sup> A tabela da Árvore do Conhecimento pode ser acessada aqui: <http://lattes.cnpq.br/web/dgp/ciencias-agrarias>

Sociais Aplicadas, o formulário recebeu respostas de todas as áreas do conhecimento. A categoria “Outros”, com 8 respondentes, corresponde à cursos técnicos em áreas diversas<sup>87</sup>.

Abaixo apresentamos uma tabela detalhada das subáreas<sup>88</sup> aos quais os participantes pertencem. Ressaltamos a participação relevante de participantes da Comunicação, que representam 84% dos respondentes da área de Ciências Sociais Aplicadas, e de Ciências da Computação, que representam 85% dos respondentes das Ciências Exatas e da Terra.

Tabela 1 – Distribuição dos Participantes por áreas e subáreas

<b>Ciências Exatas e da Terra</b>	<b>47</b>
Ciências da Computação	40
Física	3
Química	2
Geografia	2
<b>Ciências Biológicas</b>	<b>11</b>
Biologia	11
<b>Engenharias</b>	<b>23</b>
Engenharia Elétrica	7
Engenharia de Software	3
Engenharia Mecânica	1
Engenharia de Produção	1
Engenharia de Bioprocessos	1
Não especificaram	10
<b>Ciências da Saúde</b>	<b>26</b>
Educação Física	5
Farmácia	4
Fonoaudiologia	4
Nutrição	2
Medicina	2
Biomedicina	2
Fisioterapia e Terapia Ocupacional	2
Odontologia	1
Enfermagem	1
Não especificaram	3
<b>Ciências Agrárias</b>	<b>4</b>
Agronomia	4

<sup>87</sup> Enquanto alguns dos cursos técnicos poderiam ser facilmente encaixados nas áreas do conhecimento a partir da tabela do CNPq, para outros não encontramos categorias possíveis, como Gastronomia, por exemplo. Por esse motivo optamos por uma categoria em separado para os tecnólogos.

<sup>88</sup> Como a pergunta foi feita para ser respondida em campo aberto, os participantes apresentaram níveis diferentes de especificidade. Optamos por manter, na medida do possível, a especificidade maior, sem, entretanto, fazer inferências a partir das respostas. Por esse motivo a tabela apresenta níveis diferentes. Alguns participantes deixaram essa resposta em branco.

<b>Ciências Sociais Aplicadas</b>	<b>290</b>
Comunicação	244
Administração	14
Direito	9
Ciências da Informação	8
Design	6
Economia	6
Arquitetura e Urbanismo	1
Não especificaram	2
<b>Ciências Humanas</b>	<b>36</b>
Psicologia	16
Educação	8
Sociologia	4
Filosofia	2
História	1
Não especificaram	5
<b>Linguística, Letras e Artes</b>	<b>18</b>
Letras	9
Artes	4
Cinema	2
Linguística	2
Música	1
<b>Outros</b>	<b>8</b>
<b>Respostas em branco</b>	<b>13</b>
<b>Não fizeram formação universitária</b>	<b>26</b>

Fonte: a autora

Compreendemos que, em um cenário ideal de pesquisa, a amostra dos participantes na parte quantitativa deveria ser mais heterogênea, com melhor distribuição de participantes com graus de escolaridade, idade e em áreas distintas do conhecimento. Ainda assim, os desafios de realizar uma pesquisa quantitativa são grandes, especialmente com uma amostra considerável e sem orçamento específico de pesquisa para esta etapa ou uma equipe composta por diversos pesquisadores, justificando as decisões de divulgação realizadas, ainda que estas impactem no perfil da amostra. Uma vez que não estamos buscando, com esta pesquisa, validade estatística, não consideramos que os dados de perfil relacionados a idade ou gênero, conforme se apresentam nessa amostra, comprometem a validade dos resultados obtidos.

A validade estatística depende da formação de uma amostra probabilística, que é composta de forma a representar “um modelo reduzido do universo de pesquisa” (FRAGOSO; RECUERO; AMARAL, 2012, p. 76 e 77) e depende do uso de critérios estatísticos, como a

representação proporcional de regiões do país e de faixas etárias, que permitam “estimar as margens de erro envolvidas na generalização dos resultados da amostra para o universo da pesquisa” (FRAGOSO; RECUERO; AMARAL, 2012, p. 77). O formulário online é uma etapa de desenho geral desta pesquisa, que nos fornece indícios interessantes para serem analisados, mas sobre os quais não pretendemos fazer generalizações que extrapolem o universo dos respondentes.

Na sequência apresentaremos um detalhamento das entrevistas em profundidade e do perfil dos entrevistados.

### 2.2.3 SOBRE A ENTREVISTA EM PROFUNDIDADE

O roteiro da entrevista em profundidade<sup>89</sup> foi planejado de forma semiestruturada, com foco nas categorias de análise previamente estabelecidas, decorrentes das etapas anteriores desta pesquisa. A entrevista foi semiestruturada em quatro seguimentos: dados de perfil; experiência com uso de assistentes pessoais digitais; familiaridade com interface de voz e inteligência artificial; voz e gênero. Os dados coletados como perfil foram idade, nível de escolaridade, área de formação, gênero e sexualidade. O segundo segmento de perguntas se volta para questões como frequência, formas e rotinas de uso, preferências e familiaridade, percepção e descrição dos sistemas e uso para realizar “conversas” ou “brincadeiras”. O terceiro, buscava compreender como o entrevistado percebe os desenvolvimentos de Inteligência Artificial, a familiaridade dele com conceitos da computação relacionados as assistentes e sua compreensão geral sobre robôs e sua relação com as assistentes, além de percepção de sua forma de falar e educação durante as interações e da avaliação dos níveis de humanidade e confiabilidade das assistentes. O último segmento de perguntas envolvia a descrição das interações com as assistentes, relação com a voz, preferência por gênero de voz e motivos para as preferências, percepção de uma personalidade nas assistentes e avaliação da construção das assistentes como do gênero feminino.

Os pré-requisitos para participação nesta etapa foram os seguintes: ser maior de 18 anos de idade; ser usuário de uma ou mais assistentes, com frequência ou tempo de uso

---

<sup>89</sup> O roteiro completo da entrevista pode ser consultado no Anexo 6 deste trabalho.



consideráveis<sup>90</sup>, para construir uma amostra por intensidade (FRAGOSO, RECUERO, AMARAL, 2012). Como prioridade, buscamos por moradores da cidade de Porto Alegre, mas também foram conduzidas entrevistas online com moradores de outras cidades do Brasil e com três brasileiros morando atualmente fora do país, o que adiciona elementos de conveniência na construção desta amostra (FRAGOSO, RECUERO, AMARAL, 2012).

A seleção dos participantes para a etapa qualitativa se deu a partir da mescla de métodos distintos. Inicialmente, entramos em contato<sup>91</sup> por e-mail com participantes que haviam demonstrado interesse em participar da entrevista a partir do formulário quantitativo ou a partir de chamada para entrevistas divulgadas através de forma similar a adota para o formulário. Posteriormente, seguimos para a indicações de terceiros, feitas por amigos e colegas, que teriam interesse e perfil para participar das entrevistas. Por fim, utilizamos a técnica da bola de neve para a construção do restante da amostra (FRAGOSO, AMARAL e RECUERO, 2012), seguindo indicações dos próprios entrevistados. Desta forma: os entrevistados E1, E2, E3, E4 e E14 vieram a partir da indicação de colegas, amigos e conhecidos de terceiros que se enquadravam nos pré-requisitos para a pesquisa; os entrevistados E5, E6, E7, E9, E10 e E11 vieram a partir do preenchimento do formulário e indicação de interesse em participar da pesquisa; os entrevistados E8 e E13 vieram a partir de convite para entrevista divulgado no Grupo do Facebook sobre o Google Assistente; e o entrevistado E12 veio por indicação do entrevistado E4.

Estabelecemos como local presencial prioritário para a realização das entrevistas a sala do Grupo de Pesquisa Laboratório de Artefatos Digitais (LAD), na Faculdade de Biblioteconomia e Comunicação da UFRGS, mas concedemos aos entrevistados o direito de escolher o local que lhes fosse mais conveniente. Em cumprimento à resolução 510/2016, realizamos o ressarcimento do deslocamento dos participantes que precisaram se locomover para a entrevista, nos trechos ida e volta da Faculdade de Biblioteconomia e Comunicação. O ressarcimento foi definido a partir das informações fornecidas pelos participantes sobre a forma e valor do deslocamento utilizado e o pagamento foi realizado no mesmo dia da entrevista.

Aos entrevistados no formato online também foi permitido que escolhessem a plataforma com a qual possuem mais afinidade, tendo como pré-requisito apenas que a

---

<sup>90</sup> Como critério de tempo utilizamos: usar uma ou mais assistentes por um mínimo de dois anos. Como critério de frequência consideramos: utilizar uma ou mais assistentes diariamente. Os entrevistados precisavam cumprir ao menos um dos dois requisitos.

<sup>91</sup> O modelo da Carta Convite utilizada para participação na pesquisa pode ser consultado no Anexo 4 deste trabalho.

plataforma viabilizasse uma chamada de vídeo. Desta forma, tanto as entrevistas presenciais quanto as online contaram com o apoio da linguagem visual, ainda que de formas distintas. As entrevistas semiestruturadas em profundidade foram conduzidas entre os dias 08 de outubro e 01 de novembro de 2019. No total, foram 14 entrevistas, das quais 8 foram conduzidas de forma presencial e 6 no formato online. A tabela abaixo apresenta a data, duração, formato e local de todas as entrevistas realizadas.

Tabela 2 – Locais e Formato das Entrevistas

<b>ENTREVISTA</b>	<b>DATA</b>	<b>DURAÇÃO</b>	<b>FORMATO</b>	<b>LOCAL/MEIO</b>
<b>E1</b>	08/10/2019	01:01	Presencial	Casa do E1
<b>E2</b>	10/10/2019	00:51	Online	Skype
<b>E3</b>	17/10/2019	01:17	Online	Hangout
<b>E4</b>	18/10/2019	00:42	Online	Skype
<b>E5</b>	22/10/2019	01:43	Presencial	LAD
<b>E6</b>	22/10/2019	00:42	Presencial	LAD
<b>E7</b>	23/10/2019	00:49	Presencial	Trabalho do E7
<b>E8</b>	25/10/2019	01:32	Online	Facetime
<b>E9</b>	25/10/2019	00:32	Presencial	LAD
<b>E10</b>	29/10/2019	00:43	Presencial	LAD
<b>E11</b>	29/10/2019	01:07	Presencial	LAD
<b>E12</b>	30/10/2019	00:46	Online	Whatsapp
<b>E13</b>	01/11/2019	00:56	Online	Skype
<b>E14</b>	01/11/2019	00:42	Presencial	Trabalho do E14

Fonte: a autora

As diferenças no formato (presencial ou online) das entrevistas ou no local e meio (plataforma ou sala utilizada) no qual foram realizadas não trouxeram discrepâncias para a forma como as entrevistas decorreram. As entrevistas realizadas de forma online ocorreram

sempre com chamada de vídeo, o que possibilitou gerar uma proximidade entre entrevistador e entrevistado. Não consideramos, portanto, que houve uma maior dificuldade dos entrevistados online de falarem sobre os tópicos desejados, que era uma preocupação que tínhamos com este formato. As entrevistas realizadas fora da sala do LAD, local preferencial, ocorreram em locais igualmente tranquilos e silenciosos (sala privada no local de trabalho ou sala privada na casa do entrevistado), sem distúrbios de outra ordem interferindo na fluidez da entrevista. Observamos, entretanto, que as entrevistas E8 e E14, com participantes que vieram do Grupo do *Facebook* sobre a *Google Home*, ocorreram de forma menos natural que as demais. Acreditamos que dois fatores contribuíram para isso: 1) a não existência de um ponto em comum entre entrevistador e entrevistado, como, por exemplo, ser indicação de um conhecido, Instituição de Ensino ou tipo de trabalho em comum; 2) o grau de escolaridade, uma vez que estes dois entrevistados não fizeram curso de Ensino Superior, havendo uma compreensão menor sobre o trabalho de pesquisa na Universidade. Apesar das diferenças, todas as entrevistas foram consideradas de ocorrência dentro do esperado e de acordo com os cuidados que a pesquisa requer.

Antes do início de cada uma das entrevistas, a pesquisadora explicou o motivo da entrevista, o tema geral de pesquisa e a natureza e estrutura da entrevista, apresentando o Termo de Consentimento Livre e Esclarecido<sup>92</sup> para o participante, informando que este deveria ler, perguntar qualquer dúvida que tivesse e, por fim, assinar em duas vias, sendo uma sua e outra da pesquisadora. Para os entrevistados no formato online, o TCLE foi enviado de forma digitalizada ao menos uma hora antes da entrevista marcada, o qual os participantes deveriam ler, tirar dúvidas e enviar versão com assinatura digitalizada para a pesquisadora antes do início da entrevista. Para cada entrevista foi concedido um código, no formato E + ordem de entrevista, com o qual o entrevistado é identificado, para fins de anonimato.

As entrevistas foram gravadas em áudio, com autorização dos entrevistados, totalizando 13 horas e 23 minutos de material gravado. Posteriormente, foram decupadas no formato textual a partir de uma ficha padrão, contendo três colunas: a primeira marca o tempo/momento da entrevista, a segunda as anotações diretas ou indiretas do trecho em questão e a terceira contém anotações referentes à categoria de análise e observações. Assim, as 13 horas e 23 minutos de material de áudio foram decupadas em 115 páginas de texto<sup>93</sup>.

---

<sup>92</sup> O modelo do Termo de Consentimento Livre e Esclarecido assinado por todos os participantes pode ser consultado no Anexo 5 deste trabalho.

<sup>93</sup> A ficha de decupagem padrão utilizada pode ser visualizada no Anexo 7. Para manter o anonimato dos sujeitos envolvidos nas entrevistas, as decupagens completas serão fornecidas para pesquisadores somente mediante

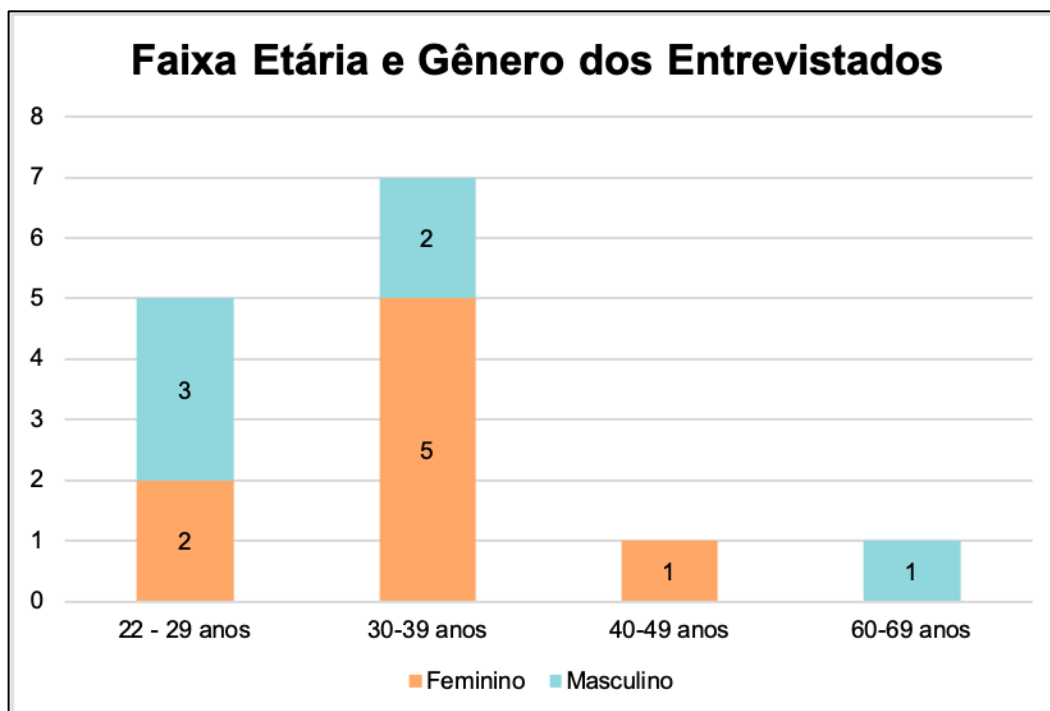
O método de decupagem das entrevistas, ao invés da transcrição completa dos áudios, segue o modelo que vem sendo utilizado nas pesquisas do Laboratório de Artefatos Digitais nos últimos experimentos (AMARO, 2016; CAETANO, 2017). Com isso, apenas os trechos de relevância específica para a investigação são transcritos na ficha, conjuntamente com o código de tempo correspondente do áudio da entrevista em questão e com observações sobre o trecho transcrito. Uma vez decupadas, as entrevistas foram analisadas a partir das categorias já mencionadas, conjuntamente com os dados quantitativos, a auto-observação e os materiais midiáticos, para compor os capítulos subsequentes.

#### 2.2.4 DO PERFIL DOS ENTREVISTADOS

Considerando que as entrevistas ocorreram a partir da garantia do anonimato e não identificação dos participantes, nesta seção iremos traçar um perfil geral dos entrevistados, tentando compreender as características do grupo como um todo. Optamos por não descrever cada um dos entrevistados por dois motivos: 1) a maior parte das entrevistas ocorreu com moradores de Porto Alegre, aos quais, somados dados como grau de escolaridade e área de formação, gênero e idade, torna-se um risco a possível exposição para leitores dos entrevistados em questão, especialmente quando cruzados com características do conteúdo da fala, que serão analisados posteriormente; 2) as características relevantes para o contexto das falas em específico podem ser acionadas quando necessário, no decorrer da análise; 3) ainda que com proporções muito diferentes, é interessante comparar as características do grupo de entrevistados como um todo com àquelas dos respondentes do questionário, uma vez que os dados de ambas as coletas de dados compõem nossa análise.

No gráfico abaixo, analisamos a distribuição do grupo de entrevistados no que diz respeito ao gênero e faixa etária.

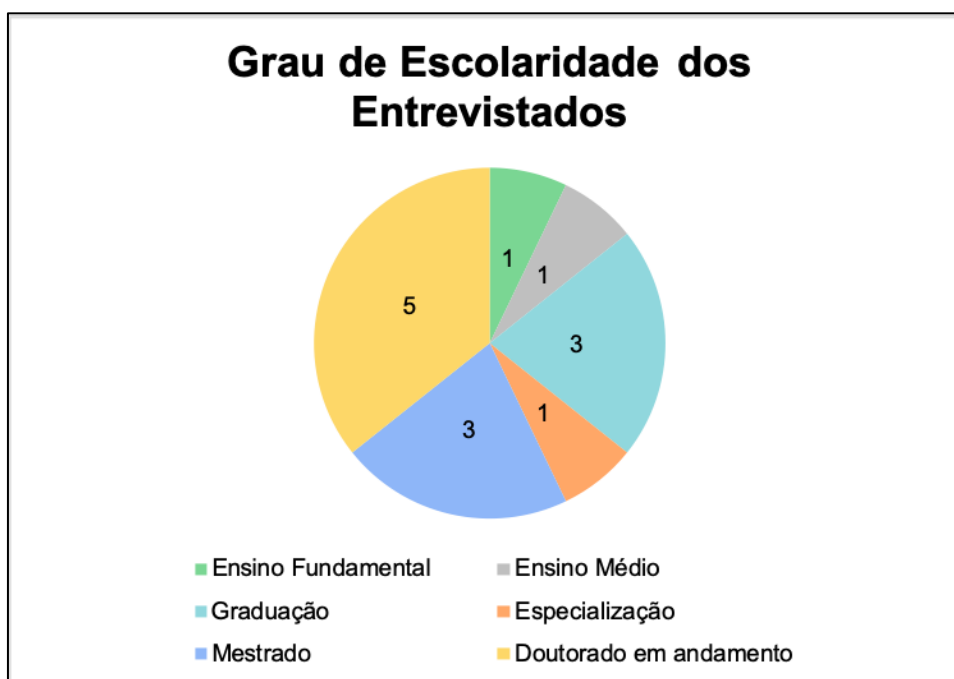
Figura 11 Gráfico Faixa etária e gênero dos entrevistados



Fonte: a autora

Das quatorze entrevistas em profundidade conduzidas para esta pesquisa, oito delas foram com mulheres e seis foram com homens, mantendo a tendência apresentada no formulário de uma participação um pouco maior de mulheres. O entrevistado mais jovem tem idade de 22 anos, enquanto o mais velho de 61 anos – mas, em termos gerais, a maior concentração das entrevistas fica na faixa etária dos 22-29 anos e dos 30-39 anos, seguindo também a tendência apresentada no formulário quantitativo. Vale ressaltar que não foi feita nenhuma triagem baseada no perfil dos entrevistados (como idade, gênero, nível de escolaridade), a triagem realizada foi baseada somente na disposição em participar voluntariamente e no cumprimento dos pré-requisitos supracitados. Na medida em que os candidatos a entrevistados foram retornando os e-mails e contatos realizados, prosseguimos para marcar as entrevistas – paramos de contatar e marcar entrevistas na medida em que o volume e a variação dos dados obtidos foi considerado adequado para a análise. No gráfico seguinte, observamos a distribuição dos entrevistados a partir do grau de escolaridade.

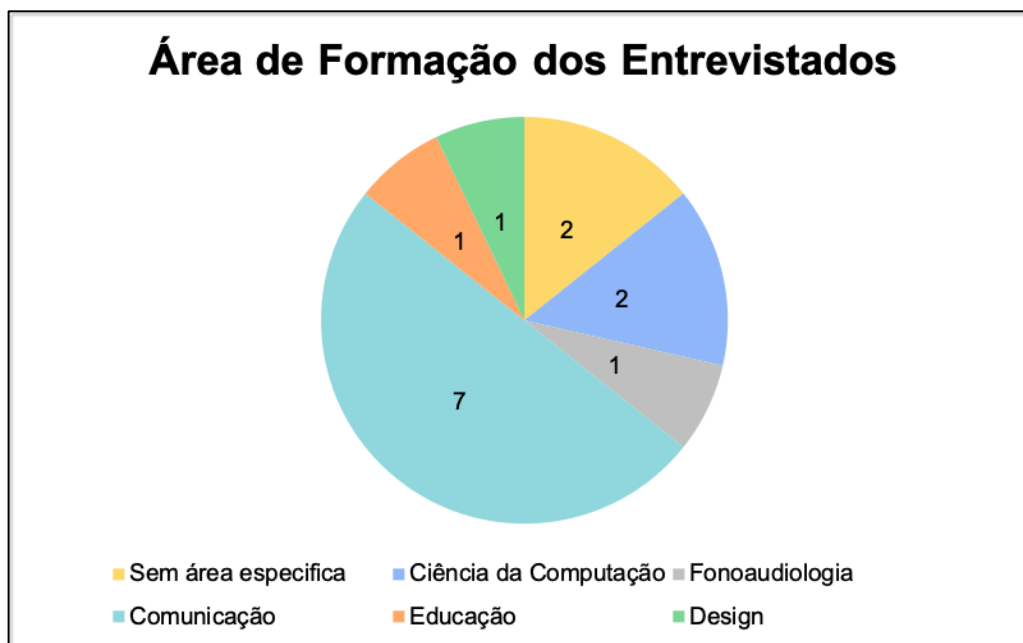
Figura 12 Gráfico Grau de Escolaridade dos entrevistados



Fonte: a autora

Observamos que este grupo, assim como o conjunto de respondentes do formulário, possui pessoas com todos os níveis de formação, desde o Ensino Fundamental até o Doutorado. Em função do viés de divulgação do formulário e da própria disponibilidade dos entrevistados para participarem da pesquisa sem qualquer gratificação, o grupo mais expressivo de participantes é de doutorandos (5). Seguem-se aos doutorandos, três mestres, três graduados, um especialista, um formado no Ensino Médio e um formado no Ensino Fundamental. No que diz respeito à área de formação dos entrevistados com Ensino de Nível Superior, o gráfico abaixo analisa a distribuição do grupo.

Figura 13 Gráfico Área de Formação dos entrevistados

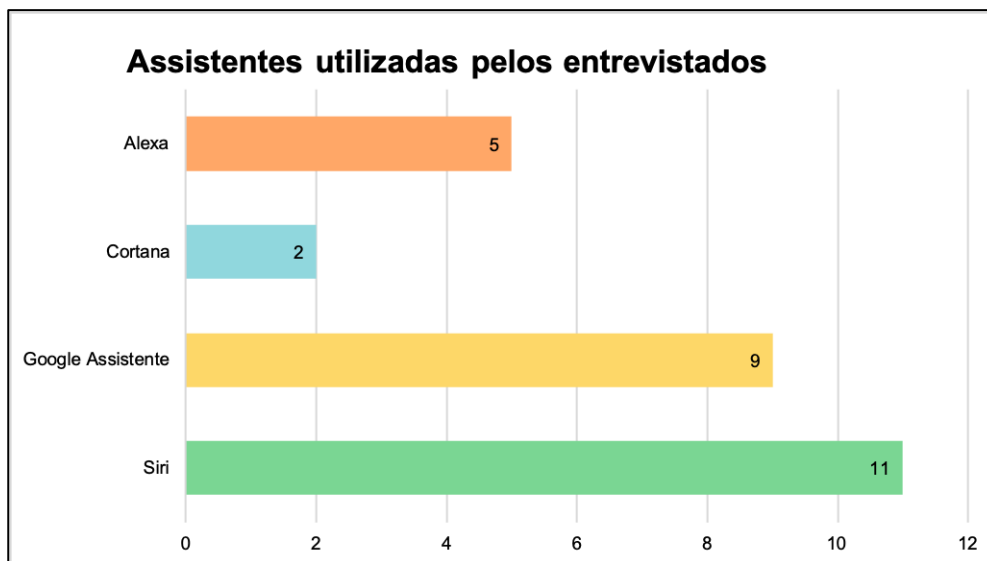


Fonte: a autora

Percebemos que a área de formação também segue tendência similar aos dados obtidos no formulário quantitativo: sete participantes da Comunicação, dois participantes da Ciência da Computação, um participante da Educação, um da Fonoaudiologia e um do design, além dos dois participantes sem área de formação. Ponderamos que, além do viés de divulgação e resultados do próprio formulário, uma vez que a realização da entrevista necessite de um grau de comprometimento maior por parte do participante (envolvendo um dispêndio longo de tempo e, algumas vezes, deslocamento), a motivação para participar de uma pesquisa da mesma área de formação (Comunicação) ou de área de intersecção com o tema (Ciência da Computação) são fatores que pesam.

Ressaltamos que, na medida em que for relevante para a análise, retomaremos os dados referentes à formação, gênero e sexualidade dos entrevistados, de forma individual. Seguiremos para a apresentação de dados referente ao perfil de uso dos entrevistados: assistentes utilizadas, preferências e frequência. O gráfico abaixo aponta as Assistentes Pessoais Digitais utilizadas pelos entrevistados.

Figura 14 Gráfico Assistentes utilizadas pelos entrevistados



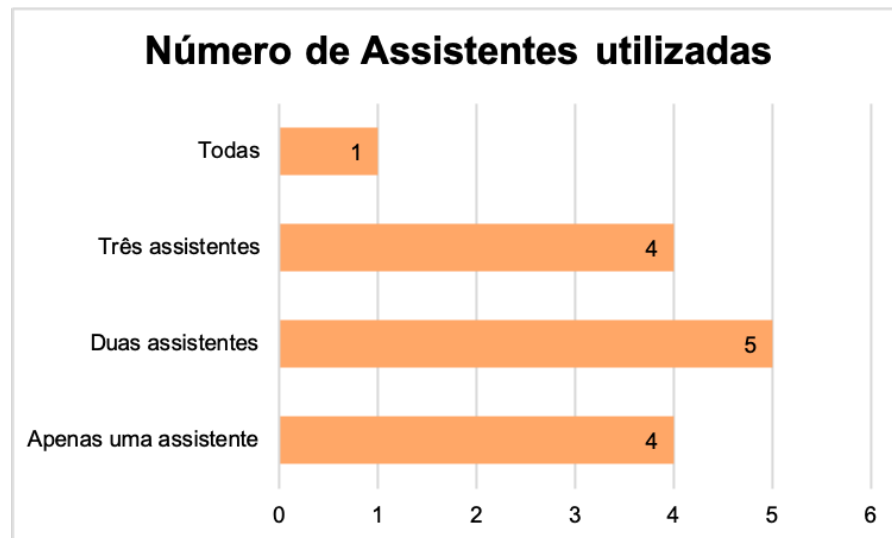
Fonte: a autora

As assistentes utilizadas pelos entrevistados apresentam uma ordenação diferente daquela apresentada pelo formulário quantitativo, que analisaremos posteriormente: a *Siri* aparece como a assistente mais utilizada, por onze dos quatorze entrevistados, seguida da *Google Assistente*, por nove entrevistados, da *Alexa*, cinco entrevistados e, por último, a *Cortana*, por dois entrevistados. Além da inversão entre *Siri* e *Google Assistente* como as mais utilizadas em relação aos dados quantitativos, *Alexa* e *Cortana* também invertem as posições, sendo a *Cortana* a menos utilizada. A última alteração ocorre, parcialmente, em função de dois entrevistados que, apesar de brasileiros, moram fora do Brasil, onde a *Alexa* está disponível há mais tempo, e a utilizam. Ainda assim, na amostra de entrevistados que tivemos, não conseguimos encontrar nenhum participante que sua assistente principal e/ou de preferência fosse a *Cortana* – os dois que informaram utilizar, o fazem apenas esporadicamente e de forma secundária.

O gráfico apresentado na sequência distribui os entrevistados por quantidade de assistentes que utilizam, apontando que dos quatorze entrevistados, dez possuem experiência de uso com mais de uma assistente e apenas quatro utilizam uma só.



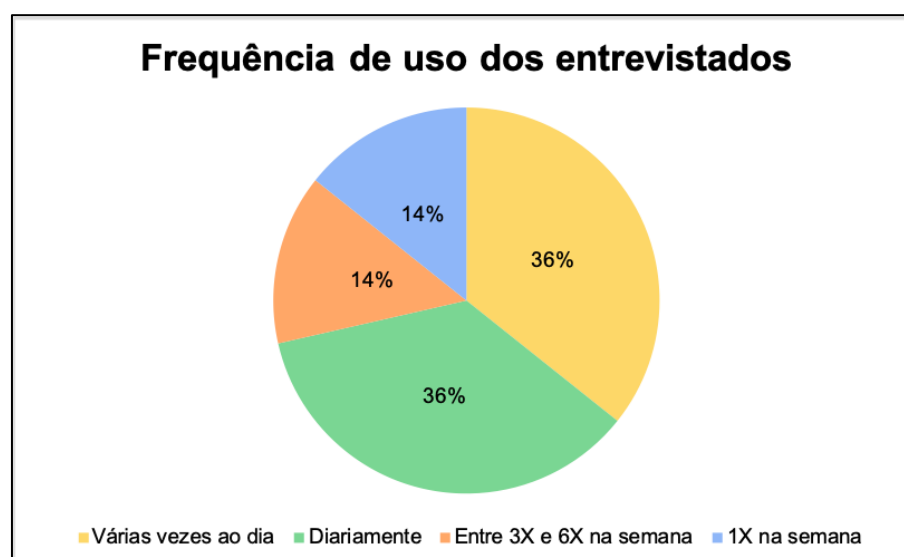
Figura 15 Gráfico Número de assistentes utilizadas



Fonte: a autora

A distribuição de experiência de uso com assistentes diversas fica da seguinte forma: cinco entrevistados utilizam/utilizaram duas assistentes, quatro entrevistados utilizam/utilizaram três assistentes e outros quatro utilizam apenas uma assistente, um entrevistado possui experiência com todas as assistentes. Todos os entrevistados que utilizam apenas uma assistente utilizam a *Siri*, o que pode apontar para uma tendência já conhecida de integração entre os dispositivos da *Apple*, em geral apreciada pelos seus usuários. O gráfico abaixo distribui os entrevistados por frequência de uso: lembrando que como pré-requisito para as entrevistas, o participante deveria ter intensidade ou tempo de uso.

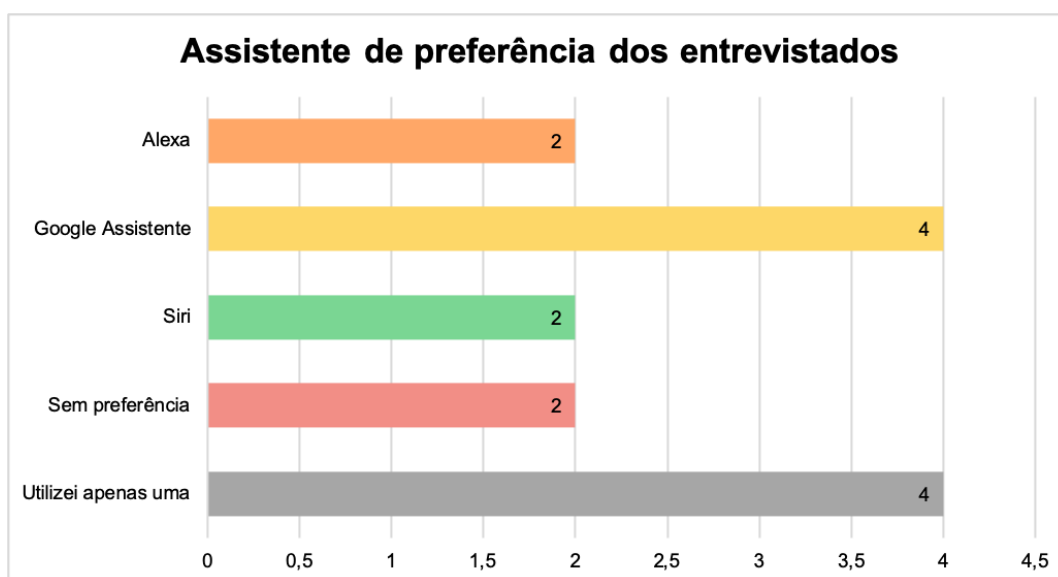
Figura 16 Gráfico Frequência de uso dos entrevistados



Fonte: a autora

Devido ao viés de seleção dos entrevistados, a frequência de utilização se mostra maior neste conjunto do que no dos dados quantitativos. Dos quatorze entrevistados, 36% informaram utilizar várias vezes ao dia e outros 36% diariamente, seguidos de 14% que utilizam entre três e seis vezes na semana e outros 14% que utilizam uma vez na semana. Interessante notar que, entre os entrevistados, tempo de utilização não necessariamente se reflete em aumento da intensidade de uso: dois dos entrevistados informaram utilizar a *Siri* desde o seu lançamento no *iPhone*, mesmo na versão em inglês, mas ainda assim sua frequência de uso é de apenas cerca de uma vez na semana. O gráfico seguinte apresenta as preferências dos entrevistados em relação as assistentes disponíveis no mercado.

Figura 17 Gráfico Assistente de preferência dos entrevistados



Fonte: a autora

Dos quatorze entrevistados, quatro informaram preferir a *Google Assistente*, especificamente no modelo *Google Home*. Entre os motivos citados para a preferência estão: a compreensão com maior clareza dos comandos utilizados, a integração com outros serviços e contas do *Google* e a compreensão e aprendizado da forma de falar do usuário. Quatro entrevistados afirmaram utilizar apenas uma assistente, dois não possuem preferência específica, dois preferem a *Siri*, outros dois preferem a *Alexa* e nenhum entrevistado prefere a *Cortana*.

### 3 AS LÓGICAS DA COMPUTABILIDADE

O filme *Ex-Machina* (Alex Garland, 2015) conta a história de Ava, uma robô dotada de inteligência artificial e com características físicas e mentais aparentemente similares a um humano, que passa por um Teste de Turing com um programador. Correndo o risco de dar algum spoiler aos leitores, a enredo corre mais ou menos assim: o protagonista, Caleb, é programador de uma grande empresa, convidado pelo dono da mesma, Nathan, para performar um Teste de Turing em Ava, sob circunstâncias misteriosas. O local de ação do filme é a casa de Nathan: isolada, altamente tecnológica e com espaço restrito no qual Ava é mantida confinada. Ao longo da trama, Ava e Caleb desenvolvem uma relação de afeto – conjuntamente com uma vigilância constante e o desenrolar de uma trama de suspeitas sobre Nathan. Em um dos encontros, quando consegue burlar a vigilância de Nathan, Ava propõe a Caleb que os dois fujam juntos do local, o que ele aceita. Resumo da ópera: Caleb auxilia Ava a se libertar, que o engana e foge sozinha, matando tanto Caleb quanto Nathan no caminho.

O filme foi lido por muitos espectadores como um estilo de “revolta das inteligências artificiais”, provando o potencial destes sistemas de se tornarem mais inteligentes que os humanos e se aproveitarem destas habilidades para virar as cartas do jogo. Entretanto, uma única passagem do filme desmente essa ideia: pouco depois de Caleb fazer as programações necessárias para a liberação de Ava, Nathan confessa que esta foi programada para tentar fugir daquele todo a todo custo e utilizando todas as suas habilidades. O que Ava faz é seguir suas diretrizes de programação com diligência, de um modo que pode se assemelhar a algumas formas de *machine learning*: a máquina consegue resultados positivos ao executar uma tarefa, entretanto, os caminhos utilizados para tanto podem ser desconhecimento inclusive de seus programadores. A novidade nesse sistema de inteligência artificial, estaria, como propõe Fazi (2018), na completude da tarefa que lhe foi designada e não na possibilidade de descumprir os comandos e buscar objetivos próprios.

Ao longo deste capítulo discutiremos e analisaremos a lógica computacional como elemento de agência das assistentes pessoais digitais, que possibilitam e limitam seus modos de ação. Iniciaremos apresentando os sistemas conversacionais, seguindo para os limites das inteligências artificiais; formas de pensar a inteligência não-humana; as questões éticas implicadas na interação com estes artefatos e a análise de dados quantitativos e qualitativos sobre como os usuários percebem estas questões.

### 3.1 COMO FUNCIONAM AS MÁQUINAS PARA CONVERSAÇÃO

Em 1966, Joseph Weizenbaum cria a ELIZA, o primeiro software de processamento de linguagem natural, que simula a conversação entre uma psicóloga e seu paciente (WEIZENBAUM, 1966)<sup>94</sup>. Durante a conversação com ELIZA, com base na linguagem natural escrita, o usuário deve assumir a posição de um paciente e o programa o de uma terapeuta Rogeriana. As respostas de ELIZA para as entradas do usuário em geral transformam uma das palavras-chave encontradas em perguntas que levem o usuário a falar mais sobre um determinado assunto.

A escolha da formatação de ELIZA enquanto uma terapeuta deste tipo específico permitia que a não-compreensão do programa do conteúdo da fala do usuário fosse utilizado como técnica de conversação e passasse como forma de método terapêutico. Weizenbaum explica a configuração do programa como uma relação entre paciente e terapeuta (1966):

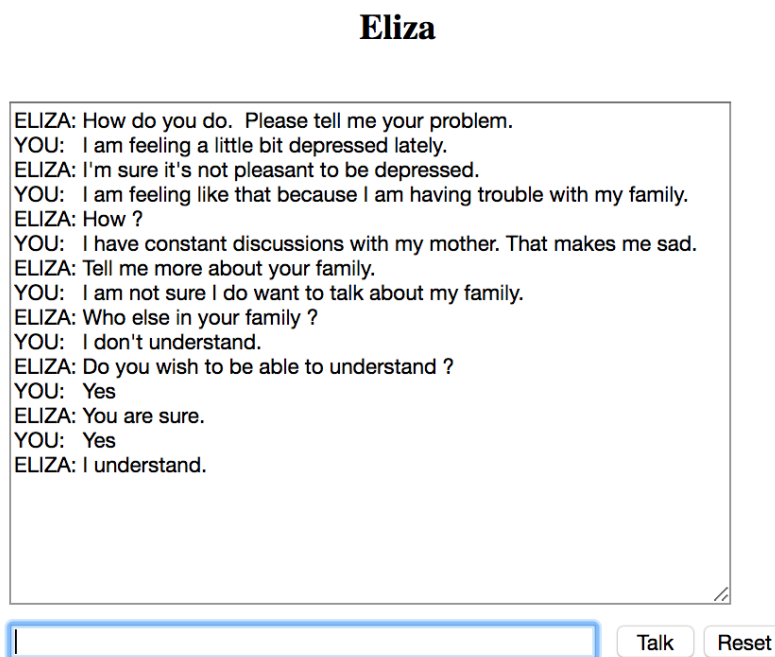
Esse modo de conversação foi escolhido pois a entrevista psiquiátrica é um dos poucos exemplos de comunicações em linguagem natural na qual um dos participantes é livre para assumir a posição de não saber praticamente nada do mundo real. Se, por exemplo, alguém falar para o seu psiquiatra “Eu saí para um longo passeio de barco” e ele responder “Me fale sobre barcos”, não iríamos presumir que o psiquiatra não sabia nada sobre barcos, mas que ele tinha um propósito em direcionar a conversação subsequente<sup>95</sup> (1966, p. 41, tradução nossa)

O uso dessa estratégia, por parte de Weizenbaum, traz algumas considerações interessantes. Primeiro, ELIZA, para assumir o papel de terapeuta e atuar nesse domínio de conversação específico, não precisa de um banco de dados com conceituações específicas e conhecimentos genéricos que geralmente precisam ser mobilizados para uma conversação entre seres humanos. Sua posição de terapeuta ainda permite que não precise tomar posições, fazer comentários e opinar, mas que apenas siga redirecionando a conversa de modo que seu interlocutor permaneça falando. O exemplo de diálogo abaixo, proveniente de interação nossa com ELIZA, ajuda a ilustrar a formatação do diálogo do programa como uma entrevista psiquiátrica:

<sup>94</sup> É possível conversar com a Eliza em: <https://www.masswerk.at/elizabot/>.

<sup>95</sup> No original: “This mode of conversation was chosen because the psychiatric interview is one of the few examples of categorized dyadic natural language communication in which one of the participating pair is free to assume the pose of knowing almost nothing of the real world. If, for example, one were to tell a psychiatrist “I went for a long boat ride” and he responded “Tell me about boats”, one would not assume that he knew nothing about boats, but that he had some purpose in so directing the subsequent conversation.” (1966, p. 41)

Figura 18 Imagem da tela de conversação do autor com versão da ELIZA disponível online



Fonte: Printscreen de conversa feito pela autora

Percebemos, pelo diálogo anterior, que as limitações computacionais de Eliza são rapidamente percebidas na interação, pois algumas das frases acionadas por elementos-chave não fazem sentido no contexto (que ELIZA, claro, não possui). ELIZA é o primeiro programa de uma classe que virá a ser conhecida como chatterbots e seu modo de funcionamento parte da análise das frases de input do usuário utilizando regras de decomposição acionadas por palavras-chave no texto de entrada: ou seja, uma vez que recebe uma mensagem de um usuário, o software da ELIZA realiza uma busca por palavras-chaves, que possuem entradas específicas na programação e acionam uma série de respostas possíveis, das quais Eliza irá escolher uma (WEIZENBAUM, 1966).

As palavras-chave possuem um ranking que permite a seleção de uma em detrimento da outra – na ausência de palavras-chave, o sistema gera um comentário sem conteúdo específico ou recupera uma transformação anterior. A partir disso as respostas são geradas mesclando técnicas de decomposição e associação (WEIZENBAUM, 1966). De acordo com Harel (2000), o sucesso de ELIZA, que levou inclusive as pessoas a compartilharem com o programa problemas que não compartilhavam com ninguém, se dá pelo âmbito restrito de atuação ao qual foi circunscrito:

A motivação por trás da ELIZA - e, nesse sentido, foi de enorme sucesso - foi mostrar que é fácil parecer inteligente; pelo menos por pouco tempo, para um observador simpático e dentro de um domínio estreito do discurso. Ser realmente inteligente é uma questão completamente diferente (HAREL, 2000, p. 198, tradução nossa<sup>96</sup>)

O modo de operação do sistema se baseia em método que envolvem a forma de construção frasal e substituições possíveis em locais específicos de uma frase, ou seja, operam a partir da sintaxe, sem envolver qualquer questão de semântica. As regras são regras de reordenação do input colocado de forma a criar um output que faz sentido, localizando determinado elementos em uma frase. Nesse sentido, ELIZA não possuía nenhum recurso para atribuir sentido às frases que lhe eram faladas, tal como pontua Turkle (2010):

Para “Minha mãe está me deixando braba”, o programa poderia responder, “Me fale mais sobre sua mãe”, ou talvez, “Porque você se sente de forma tão negativa em relação a sua mãe?” ELIZA não possuía nenhum modelo do que uma mãe poderia ser ou qualquer forma de representar o sentimento de raiva. O que poderia fazer era pegar palavras e transformar elas em questões ou reformular elas em forma de interpretações<sup>97</sup> (TURKLE, 2010, p. 23, tradução nossa)

De acordo com Weizenbaum (1966), o programa se concentra em torno dos seguintes problemas, de forma sequencial: “1. a identificação das palavras-chave; 2. a descoberta de contexto mínimo; 3. a escolha de transformações apropriadas; 4. geração de respostas na ausência de palavras-chave, e 5. a previsão de uma capacidade de fim para os scripts da ELIZA<sup>98</sup>” (WEIZENBAUM, 1966, p. 36). A linguagem de programação utilizada em Eliza é a Lisp (McCarthy, 1958), amplamente utilizada pela comunidade de IA durante os anos 60 e 70, tendo como característica importante o processamento de dados a partir de listas. O script da ELIZA é construído a partir de listas, criadas para cada palavra-chave, sendo que cada lista irá especificar técnicas de reformulação para o output dada aquela palavra-chave específica. Abaixo podemos observar uma amostra do código-fonte<sup>99</sup> de ELIZA, que auxilia na compreensão do mecanismo:

```
'((((?* ?x) hello (?* ?y))
```

<sup>96</sup> No original: “The motivation behind ELIZA - and in this respect it was enormously successful - was to show that it is easy to appear intelligent; at least for a short while, to a sympathetic observer, and within a narrow domain of discourse. To be really intelligent is quite a different matter”.

<sup>97</sup> No original: “To “My mother is making me angry”, the program might respond, “Tell me more about your mother”, or perhaps, “Why do you feel so negatively about your mother?” ELIZA had no model of what a mother might be or any way to represent the feeling of anger. What it could do was take strings of words and turn them into questions or restate them as interpretations.”

<sup>98</sup> No original: “1. the identification of key words; 2. the discovery of minimal context; 3. the choice of appropriate transformations; 4. generation of responses in the absence of keywords, and 5. the provision of an ending capacity for ELIZA “scripts”” (WEIZENBAUM, 1966, p.36).

<sup>99</sup> A amostra do código-fonte foi obtida aqui: <https://www.iaexpert.com.br/index.php/2016/10/18/historico-da-ia/>

(How do you do. Please state your problem.)  
 (((?\* ?x) computer (\*? ?y))  
 (Do computers worry you?) (What do you think about machines?)  
 (Why do you mention computers?)  
 (What do you think machines have to do with your problem?))

É possível perceber que cada entrada do usuário que acione uma palavra-chave específica prevista no banco de dados de Eliza, irá gerar uma lista de possibilidades de resposta para determinada entrada, que será calculada pelo sistema a partir de um ranking. No exemplo acima, a palavra-chave ‘computador’ aciona quatro continuações para o diálogo, em forma de perguntas: “Você se preocupa com computadores?”; “O que você pensa sobre máquina?”; “Por que você menciona computadores?” e “Qual a relação que você acha que máquina tem com o seu problema?”. O programa deve, portanto, escolher uma delas para seguir a conversa. A cada nova entrada, entretanto, o diálogo é conduzido como uma resposta única, uma vez que Eliza não armazena nenhuma das frases previamente utilizadas pelos usuários (PRIMO, 2011), mas consegue recuperar suas próprias respostas anteriores (e, portanto, evitar a repetição).

A intenção de Weizenbaum não era, entrando, efetivamente criar uma inteligência artificial, mas justamente pontuar as limitações existentes naquele período para os desenvolvimentos deste tipo: “Weizenbaum wasn’t trying to fool anyone with ELIZA. It was, in his words, a parody. He wanted to show just how superficial human–machine interactions were. There was no machine learning involved; there was no language processing happening. It was merely template-based.” (DEVLIN, 2018, p. 116).

Eliza foi a primeira dos chamados *bots* de conversações, ou *chatbots*, que serão criados em abundância nas décadas seguintes<sup>100</sup> – são nossos conhecidos na interação com empresas e até mesmo no twitter. Estes sistemas são os mais propensos a encararem o Teste de Turing, uma vez que são desenhados para realizar diálogos com humanos através da linguagem escrita, cumprindo requisitos de competição propostos por Turing (1950).

No final dos anos 60, Kabbala Reddy inicia o Projeto Hearshey com outros pesquisadores, na tentativa de criar um sistema capaz de fazer o reconhecimento de palavras faladas de forma isolada (MERCHANT, 2017). Com financiamento da ARPA (Agência de

---

<sup>100</sup> Outros exemplos mais recentes são a A.L.I.C.E (Artificial Linguistic Internet Computer Entity), criada em 1995, na Lehigh University por Richard Wallace. Exemplos comerciais mais recente apresentaram relativo sucesso, como os *chatbots* criados pela Insite. A Sete Zoom foi uma inteligência artificial criada para o site da Close Up, que simulava uma modelo e interagia com os usuários que entravam na página. Outro exemplo é o Robô Ed, da mesma empresa, criado para a Petrobras para falar sobre fontes de energia e conscientização no uso.

Projetos de Pesquisa avançados em defesa dos Estados Unidos) e situado na Universidade de Stanford, a primeira versão do sistema teve uma performance de 92%. É a partir desta tecnologia que, mais tarde, uma start-up do Vale do Silício irá desenvolver um aplicativo a ser comercializado na AppStore, que será o embrião da Siri, a primeira assistente pessoal acionada por voz (MERCHANT, 2017).

Após um período clássico, com diversos desenvolvimentos interessantes, e da passagem de uma predominância da abordagem simbólica para a abordagem conexionista, no início dos anos 2000 o campo da IA encontrava-se desacredito no meio acadêmico, com um número cada vez menor de pesquisadores dispostos a dedicar sua carreira ao assunto. Entre 2003 e 2008, o Programa CALO (Cognitive Assistant that Learns and Organizes) financiado pela DARPA, foi um projeto colaborativo entre cientistas de universidades diferentes, que auxiliou o campo da IA a recuperar o fôlego perdido e se caracteriza hoje como um dos maiores projetos do tipo já desenvolvidos. Após o encerramento do projeto, parte de seus desenvolvimentos foram levados para uma start-up pelo arquiteto chefe Adam Cheyer e o executivo chefe Dag Kittlaus. Inicialmente, a assistente pessoal não fazia nenhum reconhecimento de fala, apenas reconhecimento de linguagem natural, mas era capaz de acessar mecanismos de buscas e executar tarefas. Anos mais tarde e vivemos em um mundo povoado de assistentes pessoais digitais.

De acordo com Tom Grubber, chefe de desenvolvimentos da Siri (em entrevista para MERCHANT, 2017), a inclusão do mecanismo de fala na assistente como modo de conversação também se deve a uma necessidade de desambiguação. Uma das principais questões era, por exemplo, lidar com o fato de que a sustente assistente pessoal não sabe todas as coisas e, portanto, muitas vezes precisa “ganhar tempo” na conversação até que encontre a resposta – ou seja, para não deixar evidente ao usuário as limitações de funcionamento da assistente. As assistentes atualmente possuem formas mais ou menos diretas de lidar com os limites de suas capacidades. Quando surge um comando que a elas não conseguem codificar como um dos comandos existentes em seus sistemas ou para as quais elas não possuem uma resposta pronta: Alexa afirma “Sorry, I don’t know that one”; Google diz “Me desculpe, eu ainda estou aprendendo” e Siri informa que “Desculpe, eu não posso te ajudar com isso”. A Cortana, entretanto, nem sempre é muito boa ao mostrar que não sabe algo e realiza pesquisa no *Bing* com frequência quando não compreende.

Nesse sentido, as respostas engraçadas, os *easter eggs* e as personalidades das assistentes pessoais digitais, que veremos nos dois capítulos seguintes, podem ser compreendidas também como forma de distrair o usuário de suas próprias limitações, tornando



esses dispositivos ainda mais opacos.

As assistentes pessoais digitais hoje são na verdade uma combinação de diversas funções, como software de reconhecimento de fala, interface de linguagem natural e assistente pessoal baseado em técnicas algorítmicas de machine learning. Tom Gruber explica os passos tomados pela assistente quando acionada (em entrevista para MERCHANT, 2017): quando um usuário fala com a Siri, sua voz é digitalizada e transmitida para um servidor da Apple na nuvem, onde um software de reconhecimento de fala realiza a tradução da forma oral para a escrita – posteriormente, este texto será traduzido pelo processador de linguagem natural, já utilizado em tecnologias anteriores, como a mencionada Eliza. Uma vez que solicitação tenha sido compreendida pelo sistema, a Siri vai consultar um banco de dados com cerca de 200 megabytes sobre suas preferências e etc. Se a questão puder ser solucionada pelo próprio telefone (ligue para pessoa X, por exemplo), o request inicialmente realizado para a nuvem será cancelado. Se a Siri precisar de dados da internet para a execução do comando, o pedido segue para a nuvem, sendo analisado por um outro conjunto de ferramentas.

Tanto a Siri quanto as demais assistentes coletam dados de utilização de seus usuários, que são utilizados para tentar personalizar a experiência: é possível ensinar para Siri e Google Assistente, por exemplo, que é o seu namorado/namorada para ligações e mensagens. Além disso, de forma mais geral, os padrões de utilização coletados por cada usuário em específico e pelo conjunto de usuários em conjunto são utilizados para melhorar a compreensão de fala, o retorno das ações, a adequação das respostas, etc, a partir do machine learning:

Quando você conversa com elas, os algoritmos examinam sua solicitação para que elas possam isolar o comando. As grandes quantidades de dados que elas coletam por meio de suas interações com eles e pelas informações que elas têm em seus servidores podem ser analisadas para otimizar suas respostas. O aprendizado de máquina pega esse enorme conjunto de dados e o utiliza para gerar modelos para entender os comandos da linguagem natural. Existem limitações, é claro. As assistentes virtuais não podem responder a qualquer solicitação. Elas ainda são extremamente restritas e nem se aproximam de algo semelhante à IA geral. Elas são melhores e mais rápidas do que seus antecessores em reconhecer o que as pessoas dizem, mas não são perfeitas (DEVLIN, 2018, p. 123 e 124, tradução nossa<sup>101</sup>)

---

<sup>101</sup> No original: "When you talk to them, the algorithms scan your request so that they can isolate the command. The vast amounts of data they collect through your interactions with them, and through the information they have on their servers, can be analysed to optimise their answers. Machine learning takes this massive dataset and uses it to generate models for understanding the natural language commands. There are limitations, of course. Virtual assistants can't answer just any request. They are still hugely constrained and nowhere near anything resembling general AI. They are better and faster than their predecessors at recognising what people say, but they aren't perfect".

O motivo pelo qual as assistentes não conseguem responder todos os comandos e requisições, como pontua a autora, não apenas uma limitação temporária – ainda que a tecnologia venha melhorando (e deve continuar) consideravelmente. Existem, entretanto, limitações que são relacionadas ao modo de funcionamento das assistentes, também em termos potenciais, já que “em última análise, tudo o que fazemos com computadores se resume à matemática e existem limites fundamentais para o que podemos (e devemos) fazer com isso” (BROUSSARD, 2018, p. 19, tradução nossa<sup>102</sup>). Por isso, ao longo deste trabalho, ao analisarmos em detalhe aspectos diversos das assistentes – como a interface e a voz, por exemplo – estas estarão sempre atravessadas pelos limites e pelas potencialidades das assistentes, que derivam de sua lógica computacional de funcionamento. Isso porque, mesmo com a dinâmica conversacional que torna as assistentes mais opacas, elas são essencialmente um computador e

toda vez que você usa um computador, você está usando esse conjunto complexo de camadas. Não há mágica, embora os resultados possam parecer surpreendentes. Compreender as realidades técnicas é importante porque permite antecipar como, por que e onde as coisas vão dar errado em um cenário computadorizado. Mesmo que você sinta que o computador está falando com você ou que está interagindo com um computador, o que você está realmente fazendo é ter um contato com um programa escrito por um ser humano com pensamentos, sentimentos, preconceitos e antecedentes (BROUSSARD, 2018, p. 70, tradução nossa<sup>103</sup>).

Siri, Alexa, Cortana e Google Assistente são máquinas baseadas em computabilidade, que funcionam a partir de um sistema complexo que: a) opera a partir de algoritmos; b) é programadas, ou seja, softwares; c) depende de um hardware no qual serão acessadas e através do qual irão receber e ofertar feedback, sendo que este hardware pode ser exclusivo (como smartspeakers) ou compartilhado com outras funções (como o celular); d) dependem de outros hardwares, invisíveis aos olhos do usuários, uma vez que a maior parte de suas funcionalidades e de seus dados estão armazenadas e são processadas na nuvem (não funcionam sem internet); e) estão conectadas a redes neurais e utilizam técnicas de machine learning.

---

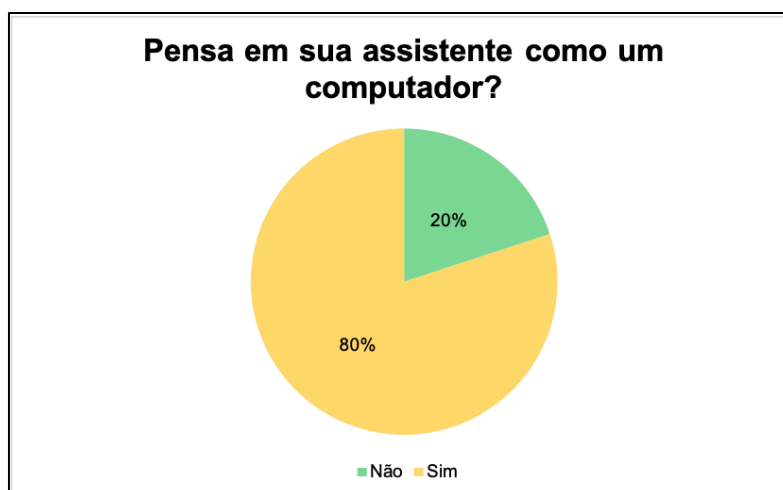
<sup>102</sup> No original: “[...] ultimately, everything we do with computers comes down to math, and there are fundamental limits to what we can (and should) do with it”.

<sup>103</sup> No original: “every time you use a computer, you are using this complex set of layers. There is no magic to it, although the results can seem amazing. Understanding the technical realities is important because it allows you to anticipate how, why, and where things will go wrong in a computerized scenario. Even if you feel like the computer is talking to you, or you feel like you are having an interaction with a computer, what you are really doing is having an interaction with a program written by a human being with thoughts, feelings, biases and background”.

### 3.1.1 ANÁLISE DOS DADOS QUANTITATIVOS SOBRE O MODO DE FUNCIONAMENTO DAS ASSISTENTES

As assistentes pessoais digitais são sistemas que propõem, como veremos nos próximos capítulos, um jogo de proximidade e de diferença com os seres humanos. Mesmo quando se propõe diferentes, o fazem de formas antropomórficas - mas não necessariamente esta precisa ser a forma de uma inteligência artificial. Entretanto, não é por utilizarem pistas antropomórficas que *Siri*, *Cortana*, *Google Assistente* e *Alexa* são menos máquinas. Em seguida, analisaremos as percepções dos usuários em relação aos aspectos computacionais das assistentes.

Figura 19 Gráfico Pensa em sua assistente como um computador?

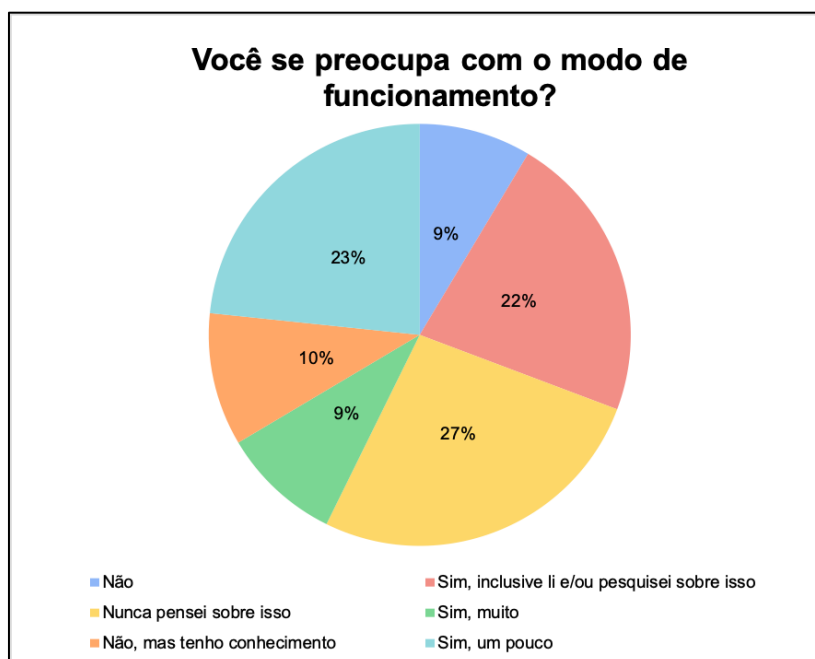


Fonte: a autora

Quando perguntados sobre a percepção da assistente pessoal digital como um computador, 80% dos usuários respondeu de forma positiva e apenas 20% respondeu de forma negativa. Nas entrevistas em profundidade, apenas E1 informou que não associa a assistente a um computador.

O gráfico abaixo apresenta as respostas dos participantes da etapa quantitativa sobre sua compreensão do modo de funcionamento das assistentes.

Figura 20 Gráfico Você se preocupa com o modo de funcionamento?



Fonte: a autora

Entre os respondentes do formulário, 27% responderam “nunca pensei sobre isso”, 23% que “sim, um pouco”, 22% “Sim, inclusive li e/ou pesquisei sobre isso”, 10% informou que “não, mas tenho conhecimento”, 9% disseram que “sim, muito” e outros 9% que “sim, um pouco”.

Dividimos os retornos dos entrevistados da etapa qualitativa em três categorias. A primeira é daqueles que afirmaram possuir nenhum ou praticamente nenhum conhecimento sobre o modo de funcionamento das assistentes, caso de E1, E6, E9, E11, E12, E13. O segundo, daqueles que informaram compreender um pouco do modo de funcionamento, em termos gerais, caso de E4, E8, E10. Por fim, aqueles que disseram que entendem como as assistentes funcionam, caso de E2, E3, E4, E5 e E14. Neste último grupo, os dois primeiros possuíam formação na área da ciência da computação, E4 informou que pesquisou por interesse próprio e E5 e E14 possui conhecimentos devido a sua área de pesquisa.

Os entrevistados também foram perguntados sobre conhecimento sobre termos específicos como *algoritmos*, *machine learning* e *redes neurais profundas*. As respostas também foram divididas em três grupos. Primeiro, E1, E6, E8 e E13 que informaram não possuir nenhum conhecimento sobre os conceitos. Segundo, E4, E9, E10, E11, E12 disseram que possuem algum conhecimento sobre os conceitos. Terceiro, E2, E3, E5, E7, E14 informaram que possuem conhecimento sobre os conceitos e que os associam aos modos de

funcionamento das assistentes. Em relação a pergunta anterior, os resultados se mantiveram pouco alterados. Além disso, com exceção de um entrevistado em cada pergunta, àqueles que compreendem as formas de funcionamento e os conceitos envolvidos são pessoas que possuem treinamento científico relacionado a sua área, o que aponta para uma opacidade desses artefatos.

É interessante comparar esses resultados com os obtidos em uma pesquisa da Mozilla Foundation (2019), que também teve a coleta de dados durante o ano de 2019, porém de nível global, mas com mais de 80% dos respondentes morando na América do Norte ou na Europa. Sobre o nível de conhecimento sobre inteligência artificial, os participantes responderam: 52% informaram “ter algum conhecimento sobre inteligência artificial”, 35% informaram ter ouvido falar, mas não saberem o que é, 10% se consideram proficiente em inteligência artificial e 4% perguntaram “O que é inteligência artificial?”. Considerando as diferenças e os pesos das amostras, é difícil traçar comparações entre os números, entretanto, é possível perceber que cerca da metade dos respondentes possuem algum conhecimento sobre o assunto, ainda que os níveis de conhecimento mais aprofundado sejam pequenos. A mesma pesquisa (MOZILLA FOUNDATION, 2019) perguntou sobre quão inteligente os usuários acreditavam que a inteligência artificial deveria se tornar. Os retornos apontam que menos da metade (39%) acreditam que a IA deve ser “muito mais inteligente do que eu”, os demais se dividem em “mais ou menos tão inteligente quanto eu” (19%), “um pouco mais inteligente do que eu” (17%), “não tão inteligente quanto eu” (16%) e “burro é melhor” (10%). Na próxima seção abordaremos os limites e as críticas que envolvem as inteligências artificiais e suas perspectivas de desenvolvimento.

### 3.2 LIMITES E CRÍTICAS DA INTELIGÊNCIA ARTIFICIAL

Searle (1997), um dos principais críticos das proposições da IA, considerando principalmente sua fase clássica e próxima relação com as ciências cognitivas, irá propor uma distinção entre o que ele chama de IA forte e fraca. A IA forte seria aquela que segue um paradigma cognitivista, ou seja, que acredita que a IA pode de fato simular um cérebro humano, ou o modo de funcionamento da cognição humana – essa proposição, para o autor, é falsa. Já a IA fraca acredita que o computador é uma ferramenta de estudo da mente, e não uma mente em si mesma, ou seja, não são capazes de explicar as capacidades cognitivas humanas.

Utilizando como base de exemplos desenvolvimentos como a Eliza ou o Schank – desenvolvido por Abelson, com o objetivo de construir um sistema capaz de realizar compreensão de texto similar à humana - a crítica de Searle (1997) se baseia na tentativa de mostrar de que forma as afirmações de que a máquina compreende histórias e que essa forma de compreensão da máquina explica a forma como os seres humanos compreendem histórias é equivocada. Para fazer isso, ele propõe examinar o exercício de pensamento do quarto chinês. Neste experimento, já bastante conhecido, um não-falante de chinês seria colocado em um quarto com uma porção de símbolos chineses que este desconhece. A partir de um conjunto de regras formais fornecidos em inglês sobre como funciona a organização daqueles símbolos, este sujeito pode aprender a manipular os símbolos a partir desse conjunto de regras, fazendo parecer que ele entende seu significado, sem, entretanto, falar chinês. Ou seja, ele domina o conjunto de regras formais do funcionamento, mas não compreende o conteúdo. Para Searle, esse seria o modo de funcionamento computacional, manipulação de símbolos a partir de determinadas regras, sem compreensão de sua significação. Ainda assim, poderia ‘parecer’, para os extremos ao quarto, que esse sujeito aprendeu a falar chinês.

Searle (1997) argumenta que existe uma diferença na forma como aplicamos o conceito de compreensão para a compreensão humana e como utilizamos essa palavra para qualificar o que outras máquinas cotidianas podem fazer. Segundo ele, em geral percebemos essa distinção, por exemplo, quando dizemos que uma porta automática ‘sabe’ quando abrir, entendemos que queremos dizer que o sensor dela detecta movimento e faz a ação designada. Entretanto, para computadores e programas de IA, Searle afirma que utilizamos a ideia de compreensão exatamente como a compreensão humana. Para ele, um programa de IA “sabe” algo tanto quanto a porta, ou, da mesma forma que a porta, e não da mesma forma que um humano. Ainda que a argumentação de Searle esteja aparentemente correta, vale acrescentar que é necessário ponderar não apenas o grau de complexidade das tarefas executadas, mas o próprio âmbito dessas ações, uma vez que o objetivo por trás de desenvolvimentos de IA buscam uma ação autônoma menos restrita do que abrir ou fechar uma porta, possuindo, portanto impactos e relevância distintos.

Como forma de contra-argumentação às críticas propostas por Searle, Gunkel (2012; 2017) traz o já mencionado problema das outras mentes, expondo as limitações das proposições:

Para distinguir, por exemplo, entre uma simulação de inteligência e uma “inteligência real”, alguém precisaria não somente ter acesso aos indicadores externos de algo que se parece com uma inteligência, mas também à atividade real de inteligência do modo como ela ocorre (ou não) na mente do outro. [...] Em outras palavras, Searle só consegue distinguir e comparar aquilo que aparece para esses indivíduos interagindo

com o quarto e o que há dentro do quarto porque ele tem acesso privilegiado aquilo que acontece em seu interior. Esse “contraexemplo”, portanto, viola as limitações epistemológicas impostas pelo problema das outras mentes, algo que no jogo da imitação Turing levou em consideração com cuidado e respeito (GUNKEL, 2017, sem paginação)

Turkle (2005) analisa de forma semelhante as críticas de Searle, ponderando que este, ainda que critique a possibilidade de uma IA forte, na verdade esta criticando o próprio Teste de Turing como medida para uma máquina inteligente. Searle coloca em questão justamente a capacidade do teste de Turing de medir a inteligência de sistema baseado apenas no output, ou seja, na aparência de inteligência. Assim, a questão estaria no próprio processo e não apenas no output e, o processo em si, não pode ser avaliado (aqui, novamente, pelo problema da mente das outras mentes).

É justamente neste ponto que a crítica de Searle (1997) se diferencia da feita por Dreyfus (1992), uma vez que o primeiro acredita que o resultado de uma IA pode ser eficaz sem que esta esteja de fato pensando – já que não é possível medir o pensamento, enquanto processo, a partir do resultado -, e o segundo defende que algumas tarefas apenas não podem ser performadas por máquinas (âmbito no qual o Teste de Turing seria efetivo).

Primo (2011) relata a experiência com a *chatbot* Cybelle, criada pelo autor para fins de pesquisa, com funcionamento através de um banco de dados e um mecanismo específico que analisa as frases ou palavras escritas pelo usuário e busca combinações possíveis previstas neste banco de dados. Com funcionamento similar ao de Eliza, o robô de conversação usa frases evasivas e genéricas quando algo que o usuário fala não está contido no banco de dados - ou seja, se ela não ‘compreende’ ou não possui uma resposta prevista para determinada entrada, existe uma lista de respostas para a situação.

Um dos pontos de dificuldade para que a conversa, tanto com a Eliza quanto com Cybelle, se mostrem similares a um diálogo entre humanos está no fato de que nenhuma das duas gravam as entradas anteriores dos usuários. Ou seja, cada frase digitada é vista como uma frase isolada, com uma resposta isolada previamente estabelecida no banco de dados, que não se relaciona com o restante da interação. Assim, não é possível que o diálogo possa ir aumentando na escala de complexidade, ainda que a forma como Eliza foi desenvolvida possa vir a dar a impressão de aprofundamento (PRIMO, 2011). Quanto maior o banco de dados, maior a possibilidade de alcance e interação deste tipo de *chatbot*. Atualmente, as assistentes pessoais digitais coletam, processam e armazenam dados de utilização dos usuários, entretanto, essas informações são raramente utilizadas nos diálogos, mas são acionadas quando necessário para realizar ações que demande uma memória. As versões em *smartspeaker* das assistentes

possuem reconhecimento de voz de usuários múltiplos, possibilitando reconhecimento e acionamento personalizado.

É justamente a necessidade do estabelecimento de um banco de dados prévio, assim como de um conjunto de regras, tornando ausente nestes dispositivos as noções de liberdade e criatividade, que leva Primo (2011) a considerar a interação entre humanos e robôs de conversação como um tipo de interação reativa, a partir de sua própria terminologia, ao invés de um modelo de interação mútua como a existente entre seres humanos (ainda que mediados por tecnologias). Noções como a de criatividade, por exemplo, nos levam a considerar que, na representação daquilo que consideramos como ser humano (e que utilizamos a partir do Teste de Turing como parâmetro para a Inteligência), existem outras qualidades envolvidas que não dizem respeito apenas a execução de tarefas, mas que são do âmbito, por exemplo, da emoção e do desejo, não correspondendo à formulações lógicas – e por vezes escapando de nossa própria compreensão, ordenação e consciência (FREUD, 1980). É possível ponderar, entretanto, sobre qual seria a necessidade de desenvolver sistemas de IA que exerçam ações relacionadas à liberdade e criatividade, essencialmente humanas, quando tais características tanto escapam ao seu modo de funcionamento formal quanto não necessariamente tornariam os sistemas autônomos melhores na execução de uma tarefa, apenas os aproximariam mais do ponto de referência humano.

Ponderação parecida aparece no trabalho de Pierce, conforme sistematizado por Nöth (2001): as máquinas não seriam capazes, até o momento, de realizar semioses completas, mas fariam uma quase-semiose quando em relação com seres humanos. O que diferenciaria, no fim das contas, a quase-semiose das máquinas de uma semiose humana estaria ligada à atributos que Pierce caracteriza como autonomia, criatividade e potencial de ação. Seria, portanto a autonomia o “traço distintivo da vida em contraposição aos sistemas não-vivos” (NÖTH, 2001, p. 66) uma vez que uma ‘semiose genuína requer, além disso, criatividade e habilidade de transformar signos em ação (NÖTH, 2001, p. 68).

Em Searle (1997), observamos uma questão similar: os programas de IA desenvolvidos possuem uma ausência de intencionalidade que está presente na manipulação de símbolos formais - ou seja, “eles têm apenas sintaxe e não semântica”. O ponto central seria com o termo “processamento de informação”, que se distingue da forma humana: um computador não processa informações, apenas as manipula; por não ser capaz de compreender o significado tanto do input que lhe é inserido quanto do output que é capaz de gerar é que a forma de lidar com informações difere da forma humana. Convergendo com a leitura do autor, Broussard aponta que, atualmente, o argumento do quarto chinês ainda é válido para as assistentes pessoais



digitais: elas conseguem decompor fórmulas específicas de fala em comandos a serem executados, mas não possuem qualquer compreensão sobre o que aquilo significa – por isso que, quando um comando do usuário foge da fórmula prevista pelo programador, as assistentes são incapazes de reconhecer e atender a solicitação.

Ainda, Searle acusa as teorias de IA de um behaviorismo residual ou operacionalismo, que tentam traduzir os modos humanos a partir dos modos de funcionamento dos computadores, tratando em termos de input e output. Interessante notar que a crítica inversa, da tentativa de restringir o âmbito computacional a uma relação com o referencial humano, foi feita por nós ao pensar o Teste de Turing como um ponto de partida antropomorfista. A questão aqui pode ser que, na comparação entre modos de operação humanos e maquímicos, não é possível fazer jus a nenhum deles, dadas as distinções epistemológicas que decorrem de cada forma de pensamento.

Na medida em que a noção de banco de dados é especialmente importante para esses agentes inteligentes, surge a analogia deste tipo de armazenamento de informação com a memória humana. Neste sentido, Primo (2011) duvida que a memória seja um simples conjunto de dados, sendo este conjunto de dados capazes de criar ou recriar uma pessoa. De acordo com o neurocientista Eric Kandel (2009), a memória humana constitui-se a partir de inscrições, marcações físicas em nosso cérebro, dependentes de reações químicas e biológicas que modificam a estrutura de nosso cérebro a cada novo armazenamento, de acordo com a importância de um determinado fato ou acontecimento na história do sujeito. Tanto aquilo que lembramos, como aquilo que esquecemos, alteram fisicamente nosso cérebro e eventos traumáticos também deixam marcações nessa estrutura complexa.

A colocação de Kandel vai de acordo com a crítica de Searle (1997), que afirma que para construirmos uma máquina pensante seria necessário replicar de forma idêntica o sistema cognitivo humano. Ou seja, o modo como pensamos, funcionamos e compreendemos também está relacionado, por exemplo, às propriedades físico-química de nossos cérebros, não sendo possível uma distinção completa entre cérebro e mente (entre máquina e operação). Uma vez que IA forte nega esta distinção, fica demarcado o forte dualismo presente nessa corrente, pretendendo o autor que se formule a questão sobre a possibilidade de existência de uma IA de forma distinta: “Mas pode algo pensar, compreender, etc somente em virtude de ser um computador com o tipo de programa adequado? Pode a instanciação de um programa, de um programa adequado é claro, ser por si só condição suficiente para compreensão?”

A questão da necessidade de um banco de dados se torna ainda mais forte nos desenvolvimentos de IA baseados em redes neurais, conforme previamente exposto. A questão

dos dados se expressa a partir de preocupações e movimentos contemporâneos distintos: a virada computacional nas pesquisas em ciências humanas, coleta de dados de usuários em plataformas digitais, proteção de dados pessoais e até mesmo o legado digital de nossos dados online pós-morte são expressões de um mesmo sintoma. Nossos traços digitais são rastros que deixamos, pequenas partículas que, se unidas, podem tentar remontar e de prever nossas preferências, estilo de vida, interesses, relacionamentos, personalidade (HARARI, 2018) – ou, ao menos, esta é a ideia presente no modo de funcionamento de ferramentas como Google e Facebook.

Por certo que a ficção nos brinda com alguns bons argumentos para pensarmos as tensões entre nossas presenças físicas e digitais no mundo, como o faz o primeiro episódio da segunda temporada da série britânica *Black Mirror*, intitulado “*Be Right Back*”. No enredo, após perder o marido em um acidente, uma jovem experimenta um serviço online que permite que as pessoas permaneçam em contato com seus entes queridos. O software utiliza as comunicações online passadas e os perfis de redes sociais para mimetizar a forma de se comunicar do falecido, possibilitando a troca de mensagens entre a personagem e seu marido. Um segundo nível da tecnologia proposta e da trama do episódio, mostram um corpo sintético (que imita esteticamente a aparência do personagem falecido) capaz de interagir a partir dos padrões de comportamento comunicacionais online e de aprender a partir de estímulos e correções, simulando, assim, a existência real do personagem. Por que julgamos que dados são capazes de reconstruir adequadamente a realidade?

As problematizações giram em torno não apenas do quanto os nossos dados dizem sobre nós, mas também sobre o que é possível realizar com o registro destes dados (e o que será feito – ou está sendo feito? – com os dados que disponibilizamos gratuitamente e em abundância cotidianamente), as relações humano-máquina e até que ponto estes softwares que aprendem e, de certa forma, pensam são apenas softwares – questões já pontuadas por outras produções contemporâneas como os filmes *Her* e *Ex-Machina* e colocadas em prática com a própria criação de sistemas de Inteligência Artificial que falam e publicam em redes sociais, através do desenvolvimento do *Machine Learning*.

### 3.3 UMA INTELIGÊNCIA NÃO-HUMANA?

O paradigma simulativo (TURING, 1950) coloca, para além da questão da antropomorfização, a comunicação como centro de um processo que tanto utiliza a conversação

como forma de medir a inteligência de um sistema quanto situa computadores como entes capazes de se expressar com humanos em linguagem natural. Essa posição estabelece um entrelaçamento entre Inteligência Artificial e Comunicação:

Por um lado, a comunicação vem sendo um instrumental tanto para a teoria quanto para a prática da inteligência artificial. Particularmente, é a comunicação que fornece à ciência da IA seus casos de teste definidores e evidências experimentais. Essa questão é imediatamente aparente num artigo que, há muito tempo, é creditado como (re)definidor da inteligência maquina. Estamos falando do texto “Computing Machinery and Intelligence”, de Alan Turing. Da mesma maneira, o desenvolvimento recente de máquinas autônomas, de algoritmos capazes de aprendizado e de sistemas inteligentes introduz novos desafios e oportunidades para os estudos em comunicação. Lidar com essas inovações e suas consequências exigirá uma recompilação significativa da disciplina e de seu foco tradicionalmente antropocêntrico, da teoria instrumentalista da tecnologia e das ideias filosóficas modernas (GUNKEL, 2017, sem paginação).

Assim, o âmbito da interação entre agentes autônomos inteligentes e seres humanos representa um campo profícuo para a pesquisa em comunicação, colocando questões que são intrinsicamente distintas daquelas postas pela interação humana enquanto mediada pela tecnologia. Aqui é a própria tecnologia que se posiciona como ator em uma rede comunicacional: ator que muitas vezes é pensado e desenhado de forma a mimetizar o humano, seja pela forma ou pela estrutura.

De acordo com Turkle (2005), parte da tentativa de antropomorfizar o computador, ou seja, de conceder a ele propriedades humanas, está justamente em nossa incapacidade de definir o que um computador pode fazer através de uma ação específica, por sua propriedade específica de ser capaz de exercer ações distintas, conforme programado. Como, por exemplo, uma calculadora, que calcula, ou um carro, que anda - o computador executa muitas tarefas (exemplo disso é o fato de que dizer que uma pessoa “está no computador” não qualifica essa ação: ela pode estar, virtualmente, fazendo praticamente qualquer coisa). Seria justamente a dificuldade de encontrar analogias no computador com outros objetos que nos leva a fazer analogias com pessoas. Mesmo que seja possível definir o que o computador faz como computação ou ato de computar, na verdade essa é a menor ação de um computador (ou a menor unidade de ação) e não conseguimos falar sobre o todo de forma satisfatória (TURKLE, 2005).

Assim, podemos considerar que a antropomorfização pela qual os sistemas de IA passam, também repete uma cisão fundante na ciência moderna, aquela que separa o sujeito do objeto (DESCARTES, 2009). Percebemos o computador como a imagem do homem moderno, tanto capaz de agir de forma estritamente racional, quanto entidade que possui uma mente

separada de um corpo – um hardware e um software. Esse ser gerado pela Inteligência Artificial, um pouco máquina e um pouco homem pelas características que lhes concedemos, se coloca em relação comunicacional com o ser humano, não constituindo apenas uma ferramenta ou um meio de comunicação.

Conforme aponta Gunkel (2012; 2017), o computador como ente que estabelece comunicação com humanos está posto desde o Teste de Turing, onde este “ocupou uma posição tanto de meio através do qual os interlocutores humanos trocavam mensagens e como um dos participantes com quem os humanos interagiam nessas trocas comunicativas”<sup>104</sup> (GUNKEL, 2012, p. 10). O campo da comunicação optou, com frequência, por focar, em seus estudos, no computador como um meio para a comunicação humana (GUNKEL, 2012; 2017), mas, ainda que esta forma de compreensão da tecnologia ofereça desenvolvimentos interessantes, a centralidade dessa perspectiva deixa de considerar “o fato de que a máquina não é apenas um meio para a competição humana mas também pode ser um participante nas interações comunicacionais”<sup>105</sup> (GUNKEL, 2012, p. 16). O autor detecta o início de uma mudança paradigmática (KUHN, 1976), na qual o computador passa a ser visto, no campo da comunicação, como um objeto que ocupa “a posição de outro ator social com quem alguém se comunica e interage. Ao vir a ocupar essa outra posição, inevitavelmente nos deparamos com questões fundamentais do âmbito da ética e da responsabilidade social”<sup>106</sup> (GUNKEL, 2012, p. 22).

Como um objeto que ocupa a posição de um outro, os sistemas de inteligência artificial podem ser pensados também como entes independentes de uma associação antropomórfica com humanos, uma vez que pensar os desenvolvimentos em IA contemporâneos como formas de inteligências atreladas a um conceito de inteligência humana talvez não seja adequado para compreender o estado das tecnologias que possuímos hoje. Partindo do já exaustivamente apresentado artigo de Turing (1950), Fazi (2018) propõe pensar a novidade no pensamento maquínico a partir de modos de produção epistêmica autônoma, fazendo um movimento de abandono do paradigma simulativo no campo da IA.

A novidade do pensamento em IA surge a partir das características intrínsecas destes sistemas, relacionadas a uma capacidade de formalização, reconhecimentos de padrão, previsão

---

<sup>104</sup> No original: “occupied the position of both medium through which human interlocutors exchange messages and one of the participants with whom one engaged in these communicative exchanges” (GUNKEL, 2012, p. 10).

<sup>105</sup> No original: “the fact that a machine is not just means of human concurrence but might also be a participant in communicative interactions” (GUNKEL, 2012, p. 16).

<sup>106</sup> No original: “the position of another social actor with whom one communicates and interactes. In coming to occupy this other position, one inevitably runs up against and encounters fundamental questions of social responsibility and ethics” (GUNKEL, 2012, p. 22)

de resultados e tomada de decisão com bases em análises de dados, e não de uma comparação com a inteligência humana (FAZI, 2018). Desta forma:

talvez seja possível abrir um campo prospectivo de autonomia onto-epistemológica para o pensamento de máquina automatizadas ao libertar estas da expectativa de que a automação é apenas uma tentativa de simulação, seja da cognição humana ou da vida e da experiência vivida<sup>107</sup> (FAZI, 2018).

Na análise de Fazi (2018), as expectativas relacionadas as possibilidade de criatividade de uma IA em geral estão presas a uma perspectiva de que estas máquinas seriam capazes de se distanciar de sua própria natureza formal, expressando características associadas ao ser humano, como por exemplo o imprevisto – ou, a criatividade, conforme exposto em subcapítulo anterior. Ou seja, uma máquina seria mais inteligente na medida em que fosse menos máquina e mais humana – expectativa presente não apenas nos desenvolvimentos de IA, mas também em parte da crítica clássica. Ao contrário, a proposta seria pensar a IA considerando “[...] máquinas que fazem exatamente o que elas devem fazer. Eu quero propor a questão: podem as máquinas computacionais gerarem novidade de formas que são profundamente alienígenas para os humanos, por serem tão inerentemente computacionais?”<sup>108</sup> (FAZI, 2018).

A saída do paradigma simulativo supõe não apenas aceitar, mas radicalizar a disparidade entre pensamento humano e pensamento da máquina, ao aceitar a formalização computacional como a especialidade da máquina, distinta da forma humana. Os desenvolvimentos computacionais com foco em IA integram uma ordem distinta de inteligibilidade, justamente por seu caráter formalizador, correspondendo a uma epistemologia da máquina (FAZI, 2018).

Tal questão não deixa de ser eminentemente Kittleriana: se lembrarmos as exposições teóricas do capítulo anterior, o autor caracteriza o modo de funcionamento do computador como a passagem de um simbólico para outro, propondo uma autonomia da maquinaria computacional em relação ao ser humano. Ou seja, tanto não são apenas ferramentas para nossa utilização, quanto possuem um modo de simbolização próprio e distinto do humano, vindo a compor um mesmo sistema, mas suas operações são independentes. Pensar a perspectiva de Kittler a partir do contexto computacional e tecnológico contemporâneo radicaliza ainda mais suas colocações.

---

<sup>107</sup> No original: “it may be possible to open an autonomous onto-epistemological prospect for automated machines thought by freeing the latter from the expectation that automation is just an attempt at simulation, whether of human cognition, or of life and lived experience” (FAZI, 2018).

<sup>108</sup> No original: “[...] machines that do exactly what they are supposed to do. I want to engage with the question: can computing machines be generative of novelty in ways that are profoundly alien to humans, because they are so inherently computational?” (FAZI, 2018).

Objeto de pesquisa interessante para uma perspectiva de IA fora de um paradigma simulativo seriam as tecnologias criadas a partir de um funcionamento de *Machine Learning* a partir de um banco de dados, com base em redes neurais, uma vez que seu modo de ação não pode ser compreendido a partir de um comportamento que simula o humano, justamente pela incapacidade de compreendermos de forma completa como se dá este tipo de aprendizagem:

Por serem modelados para agir como redes neurais profundas, essas técnicas são com frequência muito opacas, ou de fato não inteligíveis, porque são tão complexas que não podem nem ser compreendidas pelos programadores que as criaram. *Machine Learning* funciona, mas nós não entendemos completamente de que forma funciona<sup>109</sup> (FAZI, 2018).

De certa forma, Weizenbaum (1976) faz uma colocação na extremidade oposta do argumento de Fazi (2018): para o autor, a derivação de teorias que venham a explicar o modo de funcionamento da mente humana a partir do modo de operação de uma Inteligência Artificial poderia levar a uma visão achatada e mecânica da natureza humana. A saída de Weizenbaum (1976) é valorizar justamente no humano aquilo que difere do computador. Podemos propor, a partir de Fazi (2018), que é possível também, como forma de distinção onto-epistemológica, valorizar a forma pela qual o computador difere do humano, sem necessariamente buscar uma proximidade, similaridade ou mimetização.

O deslocamento de uma noção de inteligência autônoma, não necessariamente vinculada, em seus processos, com o pensamento humano, coloca novas questões para discussões relacionadas à ética. Para além dos desenvolvimentos focados em robôs de conversação e assistentes pessoais, as técnicas de IA são hoje utilizadas em um amplo número de sistemas autônomos, voltados para tarefas distintas como reconhecimento facial, diagnósticos de doenças a partir de imagem no campo da medicina e veículos autônomos – cada uma destas ramificações coloca questões e representa impactos distintos em nosso cotidiano, compondo um cenário de mundo que nem sempre depende de escolhas individuais. Isso nos leva a uma necessidade de deslocar a crítica de uma IA clássica: para além de pensar a partir da perspectiva de possibilidade de se criar um sistema com algum tipo de inteligência, sendo esta comparável ou não com a inteligência humana, torna-se relevante o questionamento a cerca das condições de produção de sistemas formais que são treinados a partir de um grande conjunto de dados e seus impactos enquanto agentes envolvidos no tecido social contemporâneo.

---

<sup>109</sup> No texto de origem: "Because they are modeled to act as deep neural nets, these techniques are often very opaque, or indeed illegible, because they are so complex that they cannot even be understood by the programmers that created them. Machine learning works, but we do not fully understand how" (FAZI, 2018).

### 3.4 QUESTÕES ÉTICAS NA RELAÇÃO HUMANO-MÁQUINA

Em fevereiro de 2017, o Parlamento Europeu aprovou, com algumas restrições, uma ampla legislação para a regulação da criação e do uso de robôs e inteligência artificial na Europa. O relatório<sup>110</sup> em questão, elaborado por Mady Delvaux, coloca em pauta uma discussão em torno da forma como iremos lidar com a crescente presença deste tipo de máquina no cotidiano econômico e social nos próximos anos, na tentativa de estabelecer direitos e responsabilidades para os agentes de AI mais avançados.

A proposta é similar ao tipo de regulação existente para companhias, que, de acordo com Delvaux (THE GUARDIAN, 2017), deve criar um enquadramento legal para os robôs que estão atualmente no mercado, ou que ficarão disponíveis nos próximos 15 anos, aproximadamente. O relatório prevê, entre outros pontos, a necessidade de criação de uma agência Europeia para Robótica e Inteligência Artificial e da elaboração de uma definição legal para robôs autônomos inteligentes. Além disso, prevê o desenvolvimento de um código de conduta para engenheiros, a fim de estabelecer padrões éticos para o desenvolvimento, produção e uso de robôs, assim como um sistema de seguro mandatário para companhias que venha a cobrir possíveis danos causados por seus robôs.

Alguns exemplos podem ser ilustrativos para pensarmos a necessidade de desenvolvimento não apenas de um sistema jurídico, mas também de parâmetros éticos para o desenvolvimento de sistemas autônomos. A automação em veículos, por exemplo, hoje não se restringe ao uso pessoal ou doméstico, mas chega também ao mercado dos serviços. A empresa Uber, por exemplo, realiza desde 2016 testes com carros autônomos nos Estados Unidos; o CEO da companhia já afirmou que a incorporação deste serviço é definitivamente o rumo que a Uber irá seguir (OLHAR DIGITAL, 2016). O mesmo tipo de tecnologia vem sendo também testado e lentamente implementado em ônibus e caminhões. O caso dos carros autônomos é de particular interesse para o sistema de legislativo proposto pela União Europeia para tratar de desenvolvimentos com IA, que frisa a necessidade não apenas de um sistema regulador europeu, mas mundial sobre a questão.

Em março de 2018, um dos carros autônomos da Uber atropelou uma mulher nos Estados Unidos, em um acidente fatal no estado do Arizona que levou a empresa a suspender

---

<sup>110</sup>Disponível em: <http://www.europarl.europa.eu/sides/getDoc.do?pubRef=//EP//NONSGML%2BCOMPARL%2BPE-582.443%2B01%2BDOC%2BPDF%2BV0//EN>

temporariamente os testes que até então vinham sendo realizados também no Canadá. Ainda que a legislação exija que os teste com esse tipo de veículo sejam realizados com um condutor humano para retomar o volante caso necessário, a polícia local verificou que o carro se locomovia de 10 km/h acima do limite de velocidade do local e que não houve tentativa de frear antes da colisão – por parte do sistema autônomo ou do condutor humano. Por enquanto, o projeto piloto da Uber funciona apenas em alguns locais e é optativo para os passageiros, entretanto, é provável que esses sistemas se tornem cada vez mais comuns com o passar do tempos, levantando questões sobre desenvolvimento, ética e responsabilidade – principalmente em caso de acidentes fatais.

Ainda mais recente que os carros autônomos, o caso da boneca falante *My Friend Cayla*, criada pelo *Vivid Toy Group*, levou a Agência Federal de Redes da Alemanha, responsável pelas telecomunicações, a emitir um aviso solicitando aos pais a retirada do brinquedo de seus filhos em 2017 (BBC, 2017). *Cayla* é uma boneca com tecnologia inteligente, com a capacidade de responder às perguntas feitas por seus donos acessando a internet para coletar as informações necessárias. Entretanto, um problema de segurança no dispositivo *bluetooth* do brinquedo possibilitava que hackers ouvissem a falassem com as crianças que brincavam com *Cayla*, levantando dúvidas e preocupações a respeito do impacto de bonecas conectadas na privacidade e segurança de crianças.

Exemplo similar ao de *Cayla*, é utilizado por Walsh (2015) para debater a necessidade de auto-identificação em sistemas de Inteligência Artificial avançados – nesse caso, a boneca de sua filha também possui uma conexão *bluetooth* que realiza acesso à Siri, assistente pessoal da Apple, possibilitando uma boneca com habilidades conversacionais razoáveis e acesso à internet. Segundo pontua o autor (2015), a Siri pode não ter ainda um desenvolvimento tecnológico que a leva a enganar um adulto sobre a ordem de objeto que é, porém, poderia, enganar uma criança. É aí que Walsh (2015) irá propor o que ele chama de ‘bandeira vermelha de Turing’, que seria uma lei de obrigatoriedade temporária para possibilitar que a sociedade se adaptasse aos sistemas autônomos avançados: até que nossa capacidade de julgamento estivesse refinada o suficiente ou que outros aparatos éticos e legais fossem desenvolvidos, todos os desenvolvimentos de Inteligência Artificial baseados em conversação deveriam indicar, ao início da interação com um humano, sua origem robótica.

A ideia parte de uma legislação britânica de 1865: conhecida como “Locomotive Act”, a lei obrigava que obrigava todos os veículos motorizados a terem uma pessoa andando na frente deles com uma bandeira vermelha, para que sinalizasse o perigo que viria. Para o autor (2015) essa medida visava amenizar tanto o medo das pessoas em relação aos veículos



motorizados quanto o impacto desses na segurança pública. O ato parou de ter validade em 1896, conjuntamente com o aumento do limite de velocidade máxima permitido. Coincidentemente, é nesse mesmo ano que ocorrem tanto o primeiro caso de excesso de velocidade quanto do primeiro atropelamento fatal (WALSH, 2015).

Aqui a proposta tem um argumento interessante, por considerar que existe um período adaptacional da sociedade em relação a uma determinada tecnologia, nem sempre associada a sua real capacidade de gerar dano. Explico da seguinte forma: é seguro dizer que, de 1896 em diante, a quantidade de acidentes gerados por automóveis, dentro ou fora do Reino Unido, aumentou, por fatores variáveis como o aumento do uso de veículos automotores, por exemplo. O que é relevante considerar é que a tecnologia não se tornou mais segura, mas talvez tenhamos nos tornado mais acostumados com os seus riscos – ou passamos a avaliar que o uso dessa tecnologia compensa os riscos corridos. Nesse sentido, Walsh (2015) estaria propondo que existe a probabilidade de que acabemos por nos acostumar também aos riscos (ainda imprevisíveis) da interação com sistemas autônomos mesmo quando não somos capazes de distinguir adequadamente sua natureza automatizada. Walsh (2015) argumenta, entretanto, que precisamos estar preparados legislativamente para o surgimento de sistemas que sejam capazes de enganar humanos, tal como Turing (1950) propôs, e agir com cautela ao menos nos anos iniciais de suas implementações, uma vez que “um dos períodos mais perigosos para qualquer nova tecnologia é quando esta tecnologia está sendo adotada pela primeira vez e a sociedade ainda não se ajustou a ela”<sup>111</sup> (sem paginação).

O autor coloca genericamente sua proposta de uma bandeira vermelha nestes termos: “um sistema autônomo deveria ser desenhado de forma a ser pouco provável que ele seja tomado por qualquer outra coisa que não um sistema autônomo e deveria se identificar no começo de qualquer interação com outro agente”<sup>112</sup> (2015, sem paginação). Inicialmente, prevê quatro cenários em que tal legislação poderia auxiliar os agentes envolvidos na interação: carros autônomos (que deveriam ser marcados como tal, para que outros motoristas possam antecipar que as decisões daquele veículo não necessariamente correspondem às respostas-padrão de motoristas humanos), assistentes pessoais digitais, textos gerados por computador<sup>113</sup> e jogos online como o pôquer, que envolvem dinheiro e nos quais um sistema inteligente pode ter

---

<sup>111</sup> No original: “one of the most dangerous times for any new technological is when the technology is first being adopted, and society has not yet adjusted to it” (WALSH, 2015, sem paginação).

<sup>112</sup> No original: “an autonomous system should be designed so that it is unlikely to be mistaken for anything besides an autonomous system, and should identify itself at the start of any interaction with another agent” (WALSH, 2015, sem paginação).

<sup>113</sup> A *Associate Press*, por exemplo, gera a maior parte de seus relatórios de ganhos de corporações americanas a partir de um programa chamado *Automated Insights* (WALSH, 2015).

vantagem em relação aos jogadores humanos.

É possível materializar a preocupação de Walsh (2015) em um desenvolvimento mais recente. O Google Duplex, uma funcionalidade do Google Assistente<sup>114</sup> apresentada em junho de 2018, consegue realizar chamadas telefônicas e manter conversas de interesse circunscrito com seres humanos, executando funções em nome do usuário que exigem contato com humanos, tais como marcar um horário no salão de beleza ou reservar uma mesa em um restaurante. A nova funcionalidade é capaz de incluir pausas e vícios da linguagem verbal humana, tornando a conversação mais realista. Em relação às versões anteriores, o Duplex consegue captar detalhes interacionais por sua base em *Machine Learning*: o assistente consegue, por exemplo, pedir opções de horário para reserva quando não for possível para o horário solicitado, ou reconhecer erros na fala do interagente humano que não condizem com os dados passados (por exemplo, o número de pessoas para qual uma reserva deve ser feita ou o horário solicitado). Como veremos no capítulo sobre voz e a fala, em termos mais amplos, as formas de comunicação das assistentes pessoais digitais não são ainda capazes de enganar, a longo prazo, os usuários sobre sua natureza – o que não implica que isso nunca virá a acontecer.

O caso Duplex coloca em questão uma preocupação anterior, de impossibilidade de distinção entre o espaço ‘virtual’ e o espaço ‘real’, posta por Gunkel (2012; 2017) a partir dos agentes de conversação: “esses softwares automatizados (aparentemente, descendentes de ELIZA) complicaram o cenário, tornando difícil de decidir quem ou o que é responsável pelas atividades no espaço virtual ou em uma comunidade virtual” (2017, sem paginação). Partindo dos chatbots e de personagens não-jogáveis em games (GUNKEL, 2012; 2017), podemos extrapolar essa confusão para outros objetos que agora operam também em uma lógica de não distinção entre as categorias de humano e não-humano. Após a apresentação pública da versão de teste da funcionalidade, a similaridade do sistema com um interagente humano reacendeu debates sobre ética no campo da Inteligência Artificial, principalmente sobre o ponto levantado por Walsh (2015) em seu artigo (de publicação anterior à apresentação do Duplex): iremos ou não legislar sobre a obrigatoriedade de um sistema de inteligência artificial anunciar-se como tal ao início de uma interação humano-máquina? Quais serão os parâmetros éticos e legislativos que iremos estabelecer socialmente para lidarmos com sistemas autônomos em situações interacionais com humanos?

Um segundo aspecto de preocupação crescente gira em torno do tema privacidade:

---

<sup>114</sup> Disponível em: <https://www1.folha.uol.com.br/tec/2018/06/assistente-pessoal-do-google-vai-deixar-claro-que-nao-e-humano.shtml>

*Apple, Google, Microsoft e Amazon* admitiram publicamente que interações dos usuários com as assistentes eram gravadas e posteriormente ouvidas e analisadas por funcionários das empresas. Além disso, áudios gravados de interações dos usuários sem acionarem as assistentes também foram gravados, incluindo de situações íntimas, como casais fazendo sexo<sup>115</sup>. Não é segredo que a coleta de dados pessoais faz parte do *modus operandi* das grandes empresas do Vale do Silício e que é justamente ela que possibilita a personalização das ferramentas no modelo como possuímos hoje. Por utilizarem técnicas de *machine learning*, com frequência a melhora na compreensão da fala e do feedback das assistentes depende de uma checagem humana e, a única forma de ensinar ao sistema como ele está respondendo de forma equivocada, é justamente realizar a correção a partir de seu próprio banco de dados de interações passadas. Com dispositivos que habitam as casas dos usuários e que podem ficar com os microfones ativados o tempo todo para detectar a *wake word*, o teor do conteúdo coletado e a possível invasão da intimidade dos usuários se torna uma questão mais sensível.

O gráfico abaixo mostra as respostas dos participantes do formulário quantitativo sobre a preocupação com privacidade na utilização das assistentes.

Figura 21 Gráfico Você se preocupa com privacidade?



Fonte: a autora

Dentre os respondentes, 33% dos respondentes informaram que “sim, se preocupam

<sup>115</sup> A notícia foi divulgada por diversos veículos de comunicação. Pode ser acessada aqui <https://www.nytimes.com/2019/08/21/technology/personaltech/alex-siri-google-assistant-listen.html>

muito”; 28% informaram que “sim, mas não em impede de utilizar”; 22% disseram que “sim, se preocupam um pouco”; 10% informaram que “nunca pensei sobre isso” e apenas 7% disseram que não se preocupam. Se considerarmos apenas as respostas que englobam “sim”, “não” e “nunca pensei sobre isso”, a proporção fica a seguinte: 83% dos usuários se preocupam com privacidade enquanto 17% não se preocupam (considerando que não pensar sobre é também não se preocupar em certa medida). Fica complicado, entretanto, inferir qualquer coisa para além do universo pesquisado, uma vez que dentro os respondentes do formulário houve uma larga participação de estudantes e pesquisadores da área da comunicação, na qual as discussões sobre coleta de dados pessoais online vêm sendo discutidas cada vez mais.

Durante as entrevistas em profundidade, os usuários se dividiram em três grupo: não se preocupam (E1, E2 e E6), se preocupam um pouco (E4 e E13) ou se preocupam muito (E3, E5, E7, E9, E10, E14). Mais uma vez as proporções pendem mais para a preocupação – ainda assim apenas dois usuários informaram que já repensaram o uso das assistentes frente a essas preocupações, entretanto, ambos optaram por manter a utilização. De forma geral, mesmo entre aqueles que se preocupam a percepção é de que: a) a coleta de dados é tão extensiva e os danos a nossa privacidade já são tão profundos que não vale mais a pena se preocupar com a questão; b) que não há saída possível para o assunto, exceto a recusa ao uso dos artefatos; c) que os eles enquanto usuários em específico não precisam se preocupar, pois não tem nada a ocultar.

E1, por exemplo, disse que “olha, agora que tu me perguntou eu percebi que nunca parei para me preocupar com isso, mas eu deveria. Mas, ah, a gente perde tanto tempo respondendo esses testes de *buzzfeed*, a nossa privacidade é tão mínima”. E2 não se incomoda muito, pois “eu não falo nada demais”, enquanto E5 acredita que a perda da privacidade é uma contrapartida para a melhora desses dispositivos. Já E7 se preocupa bastante e disse que por isso nunca utiliza a ativação da Siri pela wake word, apenas pelos botões, como forma de minimizar a sensação de estar sendo vigiado. De forma similar, E10 segue utilizando a assistente, mas pontua que “hoje em dia eu tenho uma relação com o celular e o computador meio de ameaça até porque eu sei que estou sendo escutada. A televisão está escutando... e quanto mais a gente conecta os aparelhos... [...] é óbvio que TV escuta tudo que a gente fala e tá processando, o celular tá processando... até quando a gente vai falar algumas coisas assim, e eu tenho o prazer de as vezes desligar o celular”.

E6 nos conta que “eu acho que não paro pra me preocupar tanto porque eu não uso tanto. acho que eu já parei pra pensar se, será que meu celular tá me ouvindo o tempo todo? será eu tenho um engenheiro do google que trabalha só decupando meus dados? [risos] Mas é tipo tudo muito suposições. Nunca foi uma preocupação real pra mim, até começar a ver mais notícias

sobre isso. Antes quando era algo mais encoberto eu achava que era algo inocente assim, ‘ah, eu tô falando com ela aqui, ela tá me respondendo de bom grado’ [risos]”. A entrevistada E9, ao mesmo tempo em que se preocupa muito com o assunto privacidade e coleta de dados, aponta que “cara, a gente já pensou em construir dentro de casa uma gaiola de Faraday, mas... [risos] [...] Como já se foi o boi com as cordas”.

Por fim, E13 é o único a apresentar um conceito distinto de privacidade: o usuário se preocupa com a segurança de seus dados em relação a possibilidade de ataque de terceiros e não como uma desconfiança na relação com as próprias empresas como os demais entrevistados.

## 4 O QUE SÃO OBJETOS INTERACIONAIS?

Neste capítulo trabalharemos a interface como um dos elementos centrais das formas de agência das assistentes pessoais digitais, a partir das quais elas propõe formas de interação com os usuários. Para tanto, iniciaremos com o desenvolvimento da interface gráfica e da noção de transparência, passando pelas interfaces de voz e pela perspectiva do *calm design*. Seguiremos para as formas de interação online e o conceito de interatividade, chegando em nossa proposição das assistentes enquanto um objeto interacional, que pode constituir-se enquanto outro. Na sequência analisaremos os dados referentes ao formulário quantitativo e às entrevistas em profundidade no que diz respeito aos usos, percepções e preferências, as formas de uso e rotinas, as formas e problemas de interação e ao caráter de sociabilidade das brincadeiras com as assistentes.

### 4.1 INTERFACE E TRANSPARÊNCIA

A interface gráfica computacional (do inglês *Graphic User Interface* – GUI) foi desenvolvida a partir de esforços, ideias e experimentos diversos e tomou forma no Laboratório da Xerox Parc, na década 1970. Posteriormente é incorporada e desenvolvida pela Apple para o lançamento do Mac 1984, um dos primeiros computadores de sucesso com a GUI. Esse tipo de interface, que passa a utilizar a manipulação da representação de objetos gráficos como forma de interação humano-máquina, substitui a linha de comando, um modelo textual (BOLTER; GRUSIN, 1998). A introdução de objetos gráficos como um esquema de representação faz referência a objetos culturais já familiares para os seres humanos, no desejo de aproximar o computador dos propósitos e formas de utilização humanas: “os designers acreditavam que eles estavam fazendo as interfaces mais “transparentes” e por isso mais “naturais””<sup>116</sup> (BOLTER; GRUSIN, 1998, p. 32, tradução nossa).

O desenvolvimento de softwares para funções variadas (de edição de texto, imagem, jogos, etc), pensados para operar a partir de uma interface gráfica, adiciona funcionalidades mais acessíveis a um usuário sem treinamento específico. De certa forma, a GUI inicia um processo de possibilidade de popularização da utilização dos computadores: soma-se a avanços

---

<sup>116</sup> No original: “designers believed that they were making the interfaces “transparent” and therefore more “natural”.

posteriores visando aumentar progressivamente a usabilidade destes sistemas por parte dos usuários. Isso implica em um processo de adição de camadas de mediação entre usuário e máquina, que leva a restrições sistemáticas daquilo que é possível fazer com um hardware (KITTLER, 2018), diminuindo as ações que podem ser performadas mas ao mesmo tempo diminuindo o letramento necessário para o manuseio da máquina, tornando o software uma camada cultural relevante (MANOVICH, 2013).

As interfaces, de forma geral, são em si mesmas formas de tecnologias, mas também são formas de se relacionar com uma determinada tecnologia. Nesse sentido, uma interface é o que possibilita a interação entre homem e máquina, combinando as capacidades de cada um, gerando invariavelmente restrições neste processo. A combinação humano-máquina, através da interface, gera formas de conhecer e produz conhecimento de maneiras específicas, tornando esse conhecimento indissociável de suas formas de produção (HOOKWAY, 2014). Durante as ações e as formas de produção de conhecimento operadas na interface homem-máquina, existe um processo de apagamento da própria interface, frequentemente explicado através de duas discussões específicas, “como objetos ou ferramentas tornam-se formas de ver e de conhecer e como certas técnicas são complexas, uma vez aprendidas podem ser performadas sem a necessidade de uma atenção consciente ao mesmo tempo em que seguem contribuindo para uma performance consciente”<sup>117</sup> (HOOKWAY, 2014, p. 124, tradução nossa). Interessa especificamente a relação entre transparência da interface e conhecimento tácito (POLANYI, 1958), compreendendo o último como uma série de procedimentos complexos, que aprendemos através da observação e da experimentação, e que não conseguimos formalizar adequadamente. A não formalização de formas de conhecimento tácito, como os processos e procedimentos envolvidos na interação com uma interface gráfica ou de voz, por exemplo, levam a uma percepção equivocada de que não haveria nenhum conhecimento envolvido nessas atividades, que elas seriam, de alguma forma intuitivas. A manipulação de qualquer interface computacional (assim como de outros tipos de interface), entretanto, exigem um aprendizado e treinamento que, com o passar do tempo, tonam-se invisíveis ou naturalizados. Se não percebemos as formas de conhecimento envolvidas na interação com uma determinada interface, a própria mediação que esta opera se torna menos visível.

Desde a Renascença, a ideia culturalmente difundida no ocidente é a de que o formato midiático mais poderoso seria aquele que permitisse a própria mídia seu apagamento completo,

---

<sup>117</sup> No original: “how objects or tools become ways of seeing and knowing, and how certain complex techniques, once learned, may be performed without requiring conscious attention while still contributing to consciously attended performance”.

capaz de transmitir sentido de uma consciência para a outra, com o mínimo de interferência e percepção de mediação possível (BOLTER; GRUSIN, 1998). Os autores apontam para uma lógica dupla na cultura ocidental, que deseja, ao mesmo tempo, multiplicar as formas de mídias possíveis e apagar todos os traços dessas mediações: “idealmente, quer apagar a mídia no próprio ato de sua multiplicação”<sup>118</sup> (BOLTER; GRUSIN 1998, p. 5). Considerando o computador como uma mídia, este também estaria inserido na tentativa e desejo de uma representação transparente do real.

O apagamento da percepção da mediação, que, no caso da interação em mídias digitais, passa por um apagamento da interface, não torna o representado mais real, ainda que possa causar essa percepção. O desejo de uma interface que *seamless*, levaria, por um lado, a uma busca por formas de mediação que possuam uma maior usabilidade e se adequem melhor as necessidade e habilidades dos usuários. Por outro lado, esse apagamento e a naturalização das formas de interação com a mídia de forma que pareça o real (e não uma forma de representação deste) pode levar os usuários a esquecerem que estão, de fato, operando uma mídia, uma máquina, e que as formas representacionais não estão livres de erro – ou mesmo que existem limitações de ordem técnica para as formas de operação e representação destas máquinas.

Bolter e Grusin (1998) investigam formas de representações visuais midiáticas que buscariam essa transparência, como, por exemplo, no caso das mídias digitais, o uso da realidade virtual, de gráficos tridimensionais e de design de interface gráfica, já que “o que os designers com frequência dizem é que querem uma interface sem interface, na qual não haverá ferramentas eletrônicas reconhecíveis – nenhum botão, janela, barra de de rolagem ou ícones”<sup>119</sup> (BOLTER; GRUSIN, 1998, p. 23). É possível pensar também que, a interface de voz de uma forma geral, também se caracteriza como uma busca por uma transparência da interface, ao aproximar uma forma de comunicação mais comum ao ser humano na utilização de computadores. Interfaces de voz, como as assistentes pessoais digitais, integram um projeto invisibilização da interface, justamente por serem percebidas como mais naturais ou mais integradas nas atividades cotidianas dos sujeitos - que o efeito das interfaces de voz seja efetivamente de transparência, entretanto, é discutível, ao menos no momento tecnológico atual. Conseqüentemente, elas também ocultam mais do que a interface gráfica, pela dupla lógica da remediação, considerando que lhes é adicionada uma outra camada (o que a interface gráfica

---

<sup>118</sup> No original: “ideally, it wants to erase its media in the very act of multiplying them”.

<sup>119</sup> No original: “what designers often say they want is an “interfaceless” interface, in which there will be no recognizable electronic tools - no buttons, windows, scroll bars, or even icons as such”



faz, por exemplo, em relação a linha de comando e esta em relação a programação direta do hardware, por exemplo).

A relação entre interface de voz e transparência da interface se relaciona com a noção de Computação Ubíqua (WEISER, 1991). Ainda que ambas busquem a transparência da interface (seja ela gráfica ou não), a Computação Ubíqua e os desenvolvimentos de Realidade Virtual o fazem optando por direções que são opostas. “A realidade virtual foca um enorme aparato na simulação de um mundo ao invés de invisivelmente realçar o mundo que já existe”<sup>120</sup> (WEISER, 1991, p. 94). Ou seja, a computação ubíqua busca formas de tornar as interfaces e as mediações imperceptíveis através da integração com a realidade cotidiana, a partir da qual elas viriam a compor um pano de fundo para a vida cotidiana e não da criação de formas de ambientes de interação representacionais nas quais os seres humanos poderiam interagir com seu corpo todo.

Nass (2000) aponta que a interação por voz, integrada a uma interface gráfica, seria imprescindível para a efetivação de uma computação ubíqua, por incluir maior mobilidade e inclusão de um público diverso. A interface de voz possibilitada pelas assistentes de voz, quando acionadas pelo computador, tablet ou celular, funcionam de forma complementar em relação a interface gráfica, uma vez que gera um retorno que é visual. Ou seja, a ação ainda ocorre na tela, gerando um output gráfico para um input de voz, tornando a voz um “complemento para uma experiência mais completa, pois vivemos em uma cultura visual” (PELLANDA; STRECK, 2017, p.13).

Em parte, considera-se que a interface de voz levaria a uma maior transparência da interface justamente pelo envolvimento de conhecimentos tácitos: uma vez que a fala é utilizada por seres humanos em culturas diversas como forma primordial de comunicação da espécie, compreende-se que uma interação que se baseia na fala é, de alguma forma mais natural, ou seja, não envolve procedimento de aprendizado para sua utilização. Essa ponderação é equivocada pois, como veremos no capítulo seis, os usuários possuem técnicas de fala específica para utilizar com suas assistentes, apontando para procedimentos de aprendizagem de conversação com a máquina que ocorrem posteriormente a uma aprendizado da fala em si – ou seja, existe um conhecimento tácito em jogo, mesmo depois do próprio aprendizado da linguagem, que em si mesma não é um processo natural. Ao analisar as limitações do chatterbot ELIZA, por exemplo, Devlin (2018) conclui que

---

<sup>120</sup> No original: “Virtual reality focuses an enormous apparatus on simulating the world rather than on invisibly enhancing the world that already exists”

Mas mesmo as habilidades limitadas de um computador para processar a fala podem ser fantasticamente úteis. Por ser o principal meio de os seres humanos se entenderem, cria um método de interação muito **natural**. Uma interface baseada em tela requer uma representação gráfica de um comando. Eles nem sempre são fáceis de representar e exigem mais trabalho do usuário, que precisa procurar o ícone ou a palavra apropriada em um menu e depois verificar se ele corresponde ao que deseja fazer. Por outro lado, a emissão de instruções por voz não requer uma representação gráfica que outra pessoa tenha projetado. Alivia a carga cognitiva no usuário. Isso o torna muito mais eficiente - se funcionar corretamente. Cinquenta anos depois da ELIZA, milhões de famílias dão ordens às máquinas apenas conversando com elas. E essas máquinas respondem de volta. A popularidade de assistentes digitais virtuais, como Alexa, da Amazon, Siri da Apple, Assistente do Google e Cortana da Microsoft, mostra o **quão fácil** achamos **falar naturalmente** com o software (DEVLIN, 2018, p. 119, grifos e tradução nossos<sup>121</sup>).

A linguagem falada, entretanto, não é mais ou menos natural do que outras formas de interação com um computador, uma vez que todas são invenções, com quais possuímos graus distintos de prática. Ao imaginar um mundo de proliferação de telas de tamanhos e funções diversas – de alguma forma bastante similar ao contexto comunicacional de hoje -, onde parte da materialidade destes artefatos estariam longe de nossa visão (como os cabos, as estruturas físicas de armazenamento e processamento de dados da nuvem, etc), Weiser (1991), então pesquisador da Xerox Parc, planeja um futuro onde estaríamos mais próximos de uma transparência das interfaces ao pensar formas de integrar a computação na realidade e não a realidade em representações computacionais: “as formas de tecnologia mais profundas são aquelas que desaparecem. Elas se mesclam ao tecido da vida cotidiana até que sejam indistinguíveis dela”<sup>122</sup> (WEISER, 1991, p. 94). Para Bolter e Grusin, entretanto, uma interface transparente é uma forma de manifestação da negação das características de mediação da tecnologia digital, apontando que “acreditar que com as tecnologias digitais nós ultrapassamos a mediação é também afirmar a univocidade do nosso momento tecnológico presente”<sup>123</sup> (1998, p. 24). Na proposta da Computação Ubíqua, portanto, se esconde o desejo de transparência,

---

<sup>121</sup> No original: “But even the limited abilities of a computer to process speech can be fantastically useful. Because it is the primary way for humans to understand each other, it makes for a very natural method of interaction. A screen-based interface requires a graphical representation of a command. Those aren’t always easy to represent and they require more work from the user, who needs to seek out the appropriate icon or word in a menu and then check that it matches what they want to do. By contrast, issuing voice instructions doesn’t require a graphical representation that someone else has designed. It lightens the cognitive load on the user. That makes it far more efficient – if it works correctly. Fifty years on from ELIZA, millions of households give orders to machines just by talking to them. And those machines talk back. The popularity of virtual digital assistants such as Amazon’s Alexa, Apple’s Siri, Google Assistant and Microsoft’s Cortana shows just how easy we find it to speak naturally to software”

<sup>122</sup> No original: “the most profound technologies are those that disappear. They weave themselves into the fabric of everyday life until they are indistinguishable from it”.

<sup>123</sup> No original: “to believe that with digital technology we have passed beyond mediation is also to assert the uniqueness of our present technological moment”

assim como a crença de que vivemos um momento tecnológico único, no qual seria possível o desaparecimento dos processos e efeitos da mediação.

Para que a computação ubíqua seja efetiva, é necessário que as formas de interação com os computadores, que são cada vez maiores e em formatos mais variados (como podemos observar atualmente, com gadgets, wearables, assistentes pessoais e outros artefatos digitais variados), demandem menos atenção dos agentes humanos. Inicialmente, pode parecer um contrassenso pensar que mais elementos e interações computacionais podem levar a diminuição da necessidade de atenção dos seres humanos para com esses objetos – principalmente em tempos de economia da atenção (DEVENPORT; BECK, 2001, MARTINUZZO, 2014), com plataformas digitais que buscam a manutenção do usuário em seus sistemas como forma de viabilizar a monetização e a coleta de dados. Entretanto, Weiser aponta que, em um mundo onde “em que a presença constante desses produtos no nosso background [...] não precise de nossa atenção ativa”<sup>124</sup> (WEISER, 1991, p.97) é que esses artefatos podem efetivamente se tornar invisíveis. Decorre daí a proposta de um *calm design*, onde as formas e atenção da periferia seriam aquelas que nossas mídias digitais deveriam acionar, uma vez que são também aquelas que conseguimos atender sem a necessidade de direcionamento de esforços explícitos (WEISER, BROWN, 1995). O design calmo privilegiaria um formato de interação que trabalha com formas de atenção menos fixas e demandantes que o modelo do computador pessoal, principal formato de disposto digital no período.

O design calmo seria, portanto, parte importante do processo de invisibilização das interfaces, principalmente frente a sua multiplicação (processo descrito por Bolter e Grusin como simultâneo). Recentemente, esse conceito foi revisitado, sendo tratado também pelo nome de tecnologia calma (CASE, 2015): seriam aquelas pensadas a partir da ideia de uma saturação do ambiente com artefatos pequenos e variados (fruto de um desenvolvimento de computação ubíqua), buscando a diminuição dos níveis de saturação dos indivíduos através de premissas de design. A virada por um design que teria como objetivo gerar o mínimo de carga mental possível para o usuário, seria fruto de uma crescente aceleração na população de artefatos digitais comunicacionais que requerem nossa atenção, sem que os seres humanos tenham aumentado sua capacidade de atenção nesse mesmo período, o que gera uma sobrecarga nos indivíduos.

Para o design calmo, a integração dos dispositivos digitais viria através de sua capacidade de solicitarem nossa atenção secundária enquanto realizamos uma interação

---

<sup>124</sup> No original: “the constant background presence of these products [...] does not require active attention”.

primária. A fala seria, para Case (2015), um foco de atenção primária para os seres humanos e por este motivo não deveria ser integrada nos designs, exceto em casos de necessidade extrema. Entretanto, a utilização da voz como forma de interação é contextual, ou seja, ela demanda mais ou menos atenção dos usuários somente nas especificidades daquela interação. Um dos efeitos da interface de voz é, em locais, contextos e para atividades específicas, que os usuários possam justamente fazer a manutenção de sua atividade primária ao acionar uma ação específica através da voz, especialmente quando esta atividade envolve a utilização das mãos, como cozinhar e dirigir. Assim, poderíamos pensar que, contextualmente, as assistentes pessoais digitais se configuram como artefatos que se utilizam de um design calmo, enquanto em outros contextos (e mais especificamente quando não funcionam), elas operariam de forma oposta, demandando extrema atenção do usuário. Essa questão aparece de forma variada ao longo das análises empíricas contidas neste capítulo, assim como no dois subsequentes: as assistentes podem ser uma tecnologia capaz de desaparecer quando os usuários utilizam a voz para ativar alarmes enquanto cozinham (que se mostrou um grande hit nas funcionalidades destas, inclusive), mas impõem a sua presença e demandam atenção quando não compreendem um comando ou executam de forma equivocada. Observamos que, se “uma interface transparente seria aquela que se apaga para que o usuário não esteja mais consciente de que está confrontando uma mídia”<sup>125</sup> (BOLTER; GRUSIN, 1998, p. 24), as assistentes pessoas se aproximam de uma transparência maior em alguns contextos e se tornam mais opacas em outros.

## 4.2 INTERAÇÕES E INTERATIVIDADES

Os modelos de interação conversacionais de forma presencial entre humanos, são, conforme aponta Braga (2000), uma das formas principais de caracterização da interação e da interatividade. Nos estudos de comunicação e de mídias, tende-se a defini-las a partir de uma aproximação ou afastamento com uma interatividade recíproca direta, baseada no modelo conversacional. Braga (2000, p. 03) aponta que devemos nos afastar de uma compreensão da “interatividade [como] apenas aquela, direta, recíproca, do modelo conversacional; e pensar interatividades de tipo conversacional como valor automático”.

---

<sup>125</sup> No original: “a transparent interface would be one that erases itself, so that the user is no longer aware of confronting a medium”

Compreende-se, nesse sentido, que a interatividade estabelecida pelas mídias (tradicionais ou digitais), são diferentes daquelas do modelo conversacional, mas possuem características que lhe são próprias, sendo importante

afastar o modelo conversacional como base descritiva do fenômeno. Certamente mantém-se a percepção de que a interatividade mediática não é dialética, não apresenta em geral reciprocidade entre interlocutores, é tipicamente assimétrica. Mas deve-se evitar caracterizá-la por estas lacunas (BRAGA, 2000,, p. 05).

Essa ponderação é relevante para o caso específico de interação/interatividade com/das assistentes pessoais digitais. Por um lado, elas se afastam de um modelo conversacional (ou seja, uma simulação de interação social) por serem fruto de processos máqunicos pré-programados, ainda que com certos mecanismos de automação, via machine learning (ou seja, interatividade). Configuram-se, assim, muito mais como um produto midiático que pode ser consumido através da forma conversacional, mas para o qual não existe um feedback em tempo real entre produtor e receptor (programadores e usuários). Entretanto, que a forma de consumo estabelecida tome a forma conversacional, como no caso da interface de voz com linguagem natural falada, gera efeitos que lhe são particulares na esfera da interação com o artefato. Por esse motivo, o sentido com o qual o autor opera o conceito de interação conversacional e de interatividade não se adequam a forma como nós o operamos nesta tese, uma vez que obscureceria aspectos da interação que a investigação desta tese precisa que sejam trazidos à tona.

Uma segunda possibilidade de caracterização das interações é a partir de Primo (2011) que propõe os conceitos de interação mútua e interação reativa para as interações mediadas por computador. A primeira possui um “caráter recursivo” (PRIMO, 2011, p. 107) e uma “construção dinâmica, contínua e contextualizada” (PRIMO, 2011, p. 116), podendo ser compreendidas como as interações de conversação entre humanos, podendo ser estas presenciais ou mediadas por tecnologias, desde que o “embate de ideias” (PRIMO, 2011, p. 132) seja uma possibilidade (como o telefone, o Skype, o twitter, etc). Já as interações reativas são aquelas essencialmente estabelecidas com as máquinas ou através de meios de comunicação que não possibilitam uma manutenção da conversação entre humanos, caracterizadas pela “repetição das mesmas ações programadas” (ibidem, p. 132). Nesse sentido, ainda que as assistente pessoais virtuais realizem hoje uma “reprodução [d]as características de um diálogo interpessoal”, Primo (2011) conclui que a interação estabelecida é do tipo reativa, uma vez que “diferentemente dos homens, as reações e limites dos sistemas informáticos dependem de imposição externa, e aquilo que produzem não conspira para sua transformação” (p. 171).

As distinções para as formas de interação propostas por Primo, entretanto, também não se aderem adequadamente ao objeto empírico e ao problema de pesquisa desta tese, uma vez que as assistentes pessoais digitais são interfaces para uma tecnologia que finge não ser tecnologia, mas outro sujeito humano na interação: ou seja, simulam uma interação conversacional, o que não seria nem uma interação mútua nem reativa.

Derivado do neologismo em inglês, *interactivity*, interatividade surge para qualificar, no âmbito da computação, a forma a qualidade da “interação usuário-máquina” (FRAGOSO, 2001, sem paginação). A interação “até então realizada a partir de cartões perfurados, fitas magnéticas, interruptores e dials” (FRAGOSO, 2001) passa a ser realizada com máquinas de escrever e teleimpressoras:

a denominação ‘computação interativa’ era, no entanto, insuficiente para deixar clara a qualidade da modificação na relação usuário-computador implicada pela incorporação de periféricos de entrada e enunciação de dados que permitem acompanhar, em tempo real, os efeitos das intervenções do usuário e o desenvolvimento dos processos (FRAGOSO, 2001, sem paginação)

Nesse sentido, o qualificador *interatividade*, nem quer dizer que a computação era anteriormente desprovida de interação (considerando que sempre houve interação entre usuário e máquina) e tampouco diz o mesmo sobre os demais meios de comunicação (FRAGOSO, 2001). Entretanto, o termo aponta para uma especificidade relacionada ao (naquele período) novo modo de funcionamento das mídias digitais. Assim, a expansão excessiva do termo interatividade, para abrigar outras formas de interação como, nas palavras de Braga (2001), produtos que postos “em circulação na sociedade, e [que] efetivamente circula[m]” possuem “inevitavelmente interatividade” (BRAGA, 2011, p. 06), também leva a perda da funcionalidade dessa palavra, que tem como componente principal caracterizar um tipo de interação no âmbito da computação.

o neologismo interatividade diz respeito a apenas uma das instâncias do processo de interação entre designer(s) e usuário(s): interatividade é uma atribuição da interface, ou seja, do produto midiático, e só é razoável falar em interatividade em relação a interação receptor-produto (FRAGOSO, 2001, sem paginação).

Fragoso (2001) pontua que a ideia de que ambientes interativos (como literaturas hipertextuais ou games), podem gerar uma falsa percepção de coprodução e criação através da interação humano-máquina, mas que

essa imprevisibilidade dos resultados da interação não passa, no entanto, de uma ilusão decorrente do poder quantitativo das tecnologias digitais. O exercício de atualização de uma ou outra possibilidades combinatórias de um número

necessariamente finito de opções de encaminhamento do processo interativo, sempre limitado pela especificação também do modo como são reconhecidas e incorporadas as intervenções do usuário, funciona como uma simulação de liberdade criativa (FRAGOSO, 2001, sem paginação)

A possibilidade de uma combinação de elementos em diferentes arranjos, capazes de gerar, por sua quantidade, a ideia de arranjos que são únicos, nos é particularmente interessante. As frases e respostas das assistentes pessoais digitais, acionadas a partir de inputs específicos, são programadas antecipadamente e, em geral, comportam uma certa variabilidade. Que exista variabilidade nos outputs e que os inputs possam também ser diversos, pode gerar o efeito de uma conversação similar aos processos humanos, especialmente quando somados a técnicas algorítmicas voltadas para a personalização do conteúdo para o usuário. Todos estes processos, entretanto, inclusive as técnicas e parâmetros de personalização, estão previstos nas formas de funcionamento destes artefatos.

A partir das ponderações de Primo (2011), Braga (2000) e Fragoso (2001) sobre interação e interatividade, consideramos como relevante para situar nossa abordagem de pesquisa os quatro pontos a seguir. Primeiro, que as assistentes pessoais digitais seguem uma tendência da informática de aproximar as linguagens computacionais da linguagem humana, ao invés de fazer o contrário (tanto na interação com o programador quanto na interação com o usuário). Esse processo gera um apagamento da percepção da existência de lógicas computacionais, pelas camadas de interface que buscam uma percepção de transparência. Segundo, que a partir dessa tendência, as assistentes simulam um tipo interação que é a interação interpessoal e que, ainda que o tipo de interação não seja como uma conversação entre humanos, por se oferecer neste mesmo molde gera efeitos específicos. Terceiro é que nos interessa investigar, além dos efeitos de simulação de uma conversação, formas de interação entre humanos que envolvem as assistentes enquanto produtos (ou seja: também se interage de forma interpessoal a partir delas). Por fim, que formas diferente de interação e de interatividade possuem implicações diferentes.

#### 4.3 AS ASSISTENTES DIGITAIS COMO OBJETOS INTERACIONAIS

Os objetos digitais são objetos que compõem tanto nosso cotidiano quanto nosso tecido social – muitos deles tanto agem quanto nos convocam a agir de determinadas formas, inserindo

mais um nível de complexidade nas relações devido aos seus modos de funcionamento e materialidade. Quando estes objetos digitais se apresentam de forma a propor a interação com humano em um nível conversacional, através do uso da linguagem natural, existe não apenas uma tentativa de tradução entre códigos distintos, mas a proposição de um modo de interação que é muitas vezes um espelhamento do modo de interação entre humanos. Alguns desses objetos nos convocam inclusive ao uso da fala e não apenas mais da escrita, tal como as assistentes pessoais virtuais, propondo um nível de intimidade que pode se dar apenas através da voz e não das telas.

Sherry Turkle propõe o uso do termo ‘objetos relacionais’ para tratar de artefatos que “solicitam que seus usuários não vejam eles como ferramentas mas como companheiros, como sujeitos de seu próprio direito”<sup>126</sup> (TURKLE, 2005, p. 288), ao tratar especificamente de robôs criados para agir como cuidadores, objeto de cuidado ou companheiros de idosos, crianças ou de adultos em envolvimentos amorosos. Fica claro, em suas colocações, que a perspectiva se dispõe a pensar esses objetos enquanto substitutos de humanos, ou seja, os objetos relacionais são considerados como aqueles em que a relação homem-máquina age a partir da ideia de substituição de uma relação que antes se dava entre seres humanos (ou, no mínimo, seres biológicos). Ou seja, o que Turkle pressupõe é a substituição de interação interpessoal por interatividade.

Se formas automatizadas tais como robôs para sexo, cuidadores para idosos ou animais de estimação são pensadas sempre em forma de substituição com seus ‘originais’ biológicos, importa, portanto “não o que computadores farão ou o que computadores serão no futuro, mas, ao invés disso, como nós seremos. Que tipo de pessoas nós estamos nos tornando na medida em que desenvolvemos relações cada vez mais íntimas com máquinas?”<sup>127</sup> (TURKLE, 2005, p. 294). A perspectiva adotada por Turkle é relevante, entretanto desconsidera tanto outras formas automatizadas de interação não-substitutiva, quanto a particularidade das formas de funcionamento e ação destes sistemas, em favor de uma centralidade e controle humanos. Além disso, em suas análises não aparecem formas de relação entre homem-máquina que não sejam, ou pretendam ser, análogas as interações humano-humano – como veremos ao longo das análises empíricas deste trabalho. Desta forma, a autora acaba por deixar de fora de seu âmbito de interrogação uma gama considerável de sistemas autônomos existentes hoje – ou, ao menos,

---

<sup>126</sup> No original: “ask their users to see them not as tools but as companions, as subjects in their own right” (TURKLE, 2005, p. 288)

<sup>127</sup> No original: “not what computers can do or what computers will be like in the future, but rather, what we will be like. What kinds of people are we becoming as we develop more and more intimate relationships with machines?” (TURKLE, 2005, p. 294).



parte da complexidade que integra esses sistemas e as relações humanas. Ponderamos ainda que, o “como nós seremos” no futuro em relação com estas formas específicas de artefatos digitais está atrelado ao que os “computadores serão no futuro”, o que torna relevante pensar a forma como se constituem seres humanos, computadores e as relações entre estes.

É a partir das formas de interação estabelecidas na utilização de ELIZA, o primeiro bot de conversação, que Turkle começa a pensar sobre os ‘objetos relacionais’. A ideia de um ‘efeito ELIZA’ estaria atrelada à possibilidade de tratar uma Inteligência Artificial como se fosse uma terapeuta, função até então exclusivamente humana:

Computadores, com a sua reatividade e interatividade, se localizam em uma relação nova e evocativa entre os vivos e os inanimados. Eles tornam ainda mais tentador projetar nossos sentimentos em objetos e tratar as coisas como se fossem pessoas – um impulso que eu chamo de “Efeito ELIZA” depois que o programa pioneiro de IA foi desenhado para parecer como uma terapeuta solícita. E porque computadores são ferramentas de pensar, elas oferecem novos modelos do que significa saber e entender.<sup>128</sup> (TURKLE, 2005, p. 287)

O chamado Efeito ELIZA aponta para um fator interessante no âmbito da inteligência artificial e das formas de interação com esses objetos: não é necessário que um artefato apresente de fato traços de uma inteligência real ou significativa “para que os humanos sintam que podem conversar naturalmente com uma máquina”<sup>129</sup> (DEVLIN, 2018, p. 123). Qualquer pessoa que tenha utilizado algumas vezes uma assistente pessoal digital, ou mesmo uma das muitas versões de ELIZA disponíveis online, esbarrou na falta de inteligência destes dispositivos mais cedo ou mais tarde (provavelmente mais cedo). Ainda assim, este trabalho está cheio de exemplos de formas de ação humanizadas em relação a esses artefatos, apontando que “uma semelhança com o comportamento humano pode ser suficiente para que os usuários presumam um grau de consciência [nas máquinas]”<sup>130</sup> (DEVLIN, 2018, p. 123). Nem todas as interações com as assistentes, entretanto, partem de uma matriz da forma de agir e comunicar entre humanos e algumas delas inclusive se valem da não-humanidade das assistentes para agir de determinada maneira.

As análises de Turkle (2005) estão recheadas de casos nos quais os seres humanos

---

<sup>128</sup> No original: “Computers, with their reactivity and interactivity, stand in a novel and evocative relationship between the living and the inanimate. They make it increasingly tempting to project our feelings onto objects and to treat things as though they were people - an impulse I called the “Eliza effect” after the early AI program that was designed to seem like a solicitous psychotherapist. And because computers are thinking tools, they offer new models of what it means to know and to understand.” (TURKLE, 2005, p. 287)

<sup>129</sup> Original em inglês: “in order for humans to feel they can converse naturally with a machine” (DEVLIN, 2018, p. 123).

<sup>130</sup> Original em inglês: “a semblance of human-like behaviour can be enough for us to assume a degree of sentience” (DEVLIN, 2018, p. 123)

passam a tratar seus robôs companheiros de formas similares com que tratariam outros humanos ou animais de estimação, gerando uma substituição de uma relação entre humanos por outra entre humanos e máquinas. O fenômeno da humanização dos computadores é trabalhado também a partir da perspectiva da interação humano-computador e do paradigma dos Computadores como Atores Sociais (1994; 2000; 2007; 2010), a partir da qual se compreender que computadores com algumas características humanas (como a voz, por exemplo) são capazes de acionar formas sociais de interação e levam os usuários a agirem como se estivessem lidando com outros humanos. De acordo com Devlin (2018):

Observamos esse fenômeno ocorrer repetidamente. Nós nos relacionamos com o mundo ao nosso redor nos mesmos termos de nossa matriz social. Faz sentido para nós interagir com objetos responsivos como se eles fossem capazes de nos entender, mesmo que saibamos que eles não são. Nós respondemos uns aos outros dessa forma; nós respondemos aos animais dessa forma. Se um computador fala conosco ou nos manda sinais de ser responsivo, então o nosso modo padrão de relacionamento é uma forma fundamentalmente humana, com expectativas humanas<sup>131</sup> (DEVLIN, 2018, p. 88)

Argumentamos, entretanto, que não se trata necessariamente de uma projeção de nossos sentimentos para os objetos, mas sim do estabelecimento de um tipo de interação que tanto envolve quanto desperta sentimentos específicos, que não necessariamente deveriam ser e/ou seriam direcionados a uma outra pessoa. Vale acrescentar que reconhecemos que o antropomorfismo de muitos de nossos objetos de Inteligência Artificial auxiliam no estabelecimento de tais relações, ainda que nem sempre a forma como tratamos esses objetos seja, necessariamente, análoga à forma como tratamos seres humanos ou mesmo animais. A questão específica do antropomorfismo nas assistentes pessoais digitais e outros artefatos similares será trabalhada em profundidade no capítulo seguinte.

Nem todos os objetos digitais que nos convocam a formas de relacionamento específicas, ainda que antropomorfizados, agem e/ou pretendem agir como substitutos de humanos. Entretanto, muitos desses objetos de inteligência artificial ocupam posições comunicacionais em relações com seres humanos, tal como argumentado por Gunkel (2012; 2017). A interação proposta como base para o teste de Turing gera efeitos em ambos os lados, ou seja, a forma como respondemos no processo comunicacional com uma entidade

---

<sup>131</sup> Original em inglês: “This is a phenomenon we see time and time again. We relate to the world around us in terms of our own social framework. It makes sense to us to interact with responsive things as if they are capable of understanding us, even when we know they aren’t. We respond to each other in that way; we respond to animals in that way. If a computer talks to us or sends us some sign that it is responsive, then our default way of engaging is one that is fundamentally human, with human expectation” (DEVLIN, 2018, p. 88)

computacional também importa e “o comportamento comunicacional que ela exibe, por exemplo, no jogo da imitação, gera um efeito em nós e em nossas interações sociais e relacionamentos” (GUNKEL, 2017, sem paginação).

As formas de tratamento humanizadas ou que replicam formas sociais e psicológicas de relacionamentos entre humanos podem ocorrer também na ausência de pistas antropomórficas – ou seja, com robôs que não possuem nenhuma similaridade com o humano. Pesquisa realizada com militares que utilizavam robôs autônomos para desarmamentos de bombas, aponta que os soldados que lidavam diretamente com essas tecnologias em zonas de conflito estabeleciam relações significativas com os robôs e, em alguns casos, passavam a percebê-los como uma extensão de seu próprio self (CARPENTER, 2016). Participantes da pesquisa relataram a sensação de perda quando um robô acidentalmente se danificava sem possibilidade de reparo, realização de funerais para estas máquinas ou pedidos para reparo e retorno de um mesmo robô específico para um mesmo soldado (mesmo que, em tese, todos os robôs deste tipo eram iguais em forma e finalidade). Esse fenômeno se apresenta mesmo que os soldados tenham sido treinados para compreender como funcionam e para manusear estes objetos autônomos e que tenham conhecimento sobre suas capacidades técnicas e suas finalidades específicas.

A intensidade da situação de conflito, a situação de vulnerabilidade em zonas de guerra e o próprio papel desempenhado por esses robôs, relacionado a sobrevivência dos humanos, são fatores que contribuem para um vínculo mais próximo e até mesmo para uma identificação maior destes sujeitos com os artefatos. Não é de se estranhar que os robôs possam vir a ser percebidos como uma continuidade do self destes soldados, dadas as condições, principalmente ao considerarmos a natureza da atividade que estes artefatos estavam automatizando: os robôs são tanto elemento importante na manutenção da vida dos soldados, ao desarmarem adequadamente bombas, quanto substituem ações perigosas no campo. Toda vez que um robô é danificado por um acidente no campo, portanto, são as vidas dos soldados que trabalham com ele que são poupadas, devido a substituição do trabalho de um pelo do outro.

Nesse sentido, Carpenter (2016) defende que mesmo robôs que não possuam uma forma figurativa de seres vivos podem despertar sentimentos e possibilitar a formação de vínculo, por serem capazes de se mover e performar ações de forma quase-autônoma, o que leva os seres humanos a atribuir uma intencionalidade a essas ações (CARPENTER, 2016). Acrescentamos que a natureza da ação automatizada e o papel desse robô no cotidiano dos sujeitos também são fatores para formação de um vínculo emocional. A autora compreende também que a possibilidade de que robôs sejam significativos para seres humanos pode variar tanto a partir do objeto quanto da pessoa: ou seja, assim como estabelecemos vínculos diferentes com animais

diferentes, o mesmo poderia ocorrer com robôs. Assim, eu posso estabelecer um vínculo com a minha Alexa e não estabelecer com o robô que atua na linha de produção da fábrica onde eu trabalho, por exemplo. Ou, ainda, como observamos ao longo das entrevistas em profundidade e do formulário quantitativo, podemos estabelecer vínculos fortes com uma assistente pessoal digital e não com outra, ou mesmo ter vínculos de diferentes profundidades com artefatos distintos.

É possível também que os vínculos estabelecidos com robôs e inteligências artificiais sejam distintos, assim como são os vínculos humanos: vínculos relacionados ao papel romântico, papel do cuidado ou do companheirismo e amizade, por exemplo. Aspectos relacionados ao design desses dispositivos podem fornecer pistas importantes para o desenvolvimento de vínculos específicos, com propósitos específicos, já que é possível “imaginar uma situação terapêutica onde um robô é utilizado como um substituto temporário para um humano para que um usuário/paciente possa praticar modelos de comunicação saudáveis e bem sucedidos”<sup>132</sup> (CARPENTER, 2014, sem paginação). Para Carpenter (2014), os problemas desse tipo de vínculos emocionais com artefatos digitais semiautônomos ocorreriam apenas se este apego impedir formas saudáveis de vida para os usuários (aqui uma vida saudável é compreendida de forma ampla, incluindo o estabelecimento e manutenção de relações sociais saudáveis).

Para Carpenter, o problema em relações afetivas com robôs é que este tipo de interação, entre humanos e robôs “são transações e não são recíprocas, portanto provavelmente não trariam realização suficiente para que as pessoas pudessem contar com elas como relações de longo de termo, que substituiriam um laço afetivo orgânico ou para serem utilizadas como substitutos temporários para relações entre humanos recíprocas”<sup>133</sup>(CARPENTER, 2014, sem paginação). Quando formamos efetivamente um laço afetivo com uma tecnologia, nossa forma de resposta e de relacionamento com tal artefato é distinta daquela que usualmente concedemos a outras formas de máquinas ou ferramentas.

Devlin (2018) e Carpenter (2014) concordam que as formas de relação entre humano-máquina e entre humano-humano não nem iguais e nem substitutivas, “as pessoas podem sentir amor por algo que não é humano, mas nós também temos que pensar que é uma forma

---

<sup>132</sup> Original em inglês: “imagine therapeutic situations where a robot is used as a temporary stand-in or surrogate for a human so that a user/patient can practice healthy and successful social-emotional models of communication”.

<sup>133</sup> No original: “are transactions and not reciprocal, and therefore probably not fulfilling enough for people to rely on as a long-term means for substituting organic two-way affectionate bonds, or for use as a surrogate for a human-human reciprocal relationship”

totalmente diferente de relacionamento”<sup>134</sup> (DEVLIN, 2018, p. 322). Ainda que utilizemos padrões sociais e modelos conhecidos de comunicação e interação, como as estabelecidas entre humanos ou entre humanos e animais, nós estamos cientes de estarmos substituindo uma das pontas dessa interação por um robô – e que este robô é algo que nos é não-familiar, que nos é estranho. Como pondera Devlin (2018), objetos autônomos antroporizados são “algo novo para nós: similar, mas muito definitivamente outro”<sup>135</sup> (DEVLIN, 2018, p. 322). Portanto, os sentimentos despertados e expressados em relação a estes artefatos são reais, mas são outras formas de sentimentos, uma vez que estabelecemos novas formas de relação e interação (DEVLIN, 2018; CARPENTER, 2014).

Ao avaliar a questão dos robôs para sexo, Carpenter avalia que as relações estabelecidas não serão apenas uma sobreposição dos modelos de interação que possuímos para nos relacionar com humanos – e que mesmo entre nossas relações humanas, variamos nossas formas de agir a partir de contextos diferentes. Assim, com o tempo, desenvolveríamos formas específicas de interação com esses artefatos, que poderiam se tornar uma categoria de forma de agir social distinta de outras que possuímos, mas, de qualquer forma, “o relacionamento não é tomado como um entre humanos. Nem o substitui. É mais um paralelo”<sup>136</sup> (DEVLIN, 2018, p.326).

A partir desta discussão, iremos propor aqui um deslocamento do conceito de ‘objeto relacional’ (TURKLE, 2005) para o que trataremos por ‘objeto interacional’: aqueles que solicitam de nós formas específicas de interação que são baseadas em formas utilizadas entre seres humanos, como a fala ou a escrita, por exemplo. A noção de objeto interacional nos auxilia a compreender os contextos nos quais estas interações se aproximam com interações que estabelecemos com outras pessoas, mas, por outro lado, nos permite tencionar as situações nas quais se estabelece justamente de forma oposta ou contraditória a uma relação entre humanos. No segundo caso, as máquinas, muitas vezes nos permitem perceber questões justamente pela distinção, por serem outra coisa que não seres humanos e, portanto, por nos desapontarem ou auxiliarem (entre outras possibilidades) de formas não-humanas.

Basset (2018) também vai partir do efeito ELIZA para pensar as configurações contemporâneas, uma vez que a terapeuta mecânica seria uma “iteração precoce de uma forma

---

<sup>134</sup> No original: “people can feel love for something that isn’t human, but we also both think it’s a very different type of relationship”

<sup>135</sup> No original: “something new to us: similar, but very definitely other”

<sup>136</sup> No original: “the relationship is not mistaken for a human-human one. Nor do they replace it. It’s more of a parallel”.

de relacionamento agora pervasivo entre homem e máquina”<sup>137</sup> (BASSET, 2018, sem paginação, *tradução nossa*) e que é capaz de expressar as “forms of the ‘making object’ of human subjects”<sup>138</sup> (BASSET, 2018, sem paginação) que permeia os processos de mecanização. Desta forma, nas relações propostas com objetos através da linguagem natural e, mais especificamente, da fala, tanto ocorre uma aproximação destas tecnologias daquilo que é humano (processos de antropomorfização), quanto uma aproximação dos humanos dos funcionamentos maquinicos (por exemplo, quando falamos de formas específicas com as assistentes pessoais, visando a compreensão de um comando).

Mesmo que Weizenbaum tenha realizado inúmeros esforços para ‘desmascarar’ sua própria criação, ou seja, para demonstrar que ELIZA era apenas um programa de computador, com um modo de funcionamento específico, simulando uma psicoterapeuta (e não de fato uma terapeuta), para Basset (2018) o ponto central é a recepção da ELIZA pelo público e a interação estabelecida entre ela e os usuários. Ou seja, ela exerceu, de alguma forma, a função e a posição de uma psicoterapeuta em um sistema interacional, mesmo não sendo dotada de inteligência no sentido humano do termo.

Talvez nós possamos dizer que, se “ELIZA” era um código, então “ELIZA” era o conforto encontrado na máquina, pelos humanos, que construíram uma forma diferente de relacionamento com ela que excedia o que os procedimentos do código poderia oferecer, precisamente porque o código entrou em contato com o pensamento humano.<sup>139</sup> (BASSET, 2018, *tradução nossa*).

Que ELIZA tenha operado de forma similar a um humano e possibilitado uma interação e troca até então vinculados apenas ao humano não significa que o programa seja tomado por, ou pretenda ser, uma substituição de uma relação humana. É possível que o efeito ELIZA seja fruto justamente do distanciamento entre ELIZA e um ser humano – com isso quero dizer, que a não-humanidade de ELIZA pode ser um fator importante que permitiu que os usuários conversem livremente com ela, pela ausência da possibilidade de emergência de julgamentos, por exemplo. Ainda assim, os efeitos de posicionar um artefato digital neste local interacional e com características antropomorfizadas (como a emulação de um modo de pensar, a utilização da fala e as caracterizações de gênero e personalidade, por exemplo) gera efeitos, justamente por se diferenciar da forma de relacionamento proposto por outros objetos digitais. A partir dessa

---

<sup>137</sup> No original: “early iteration of a form of human-machine relationship now pervasive” (BASSET, 2018, sem paginação).

<sup>138</sup> No original: “forms of the ‘making object’ of human subjects” (BASSET, 2018, sem paginação).

<sup>139</sup> No original: “Perhaps we might say that, if “ELIZA” was code, then “ELIZA” was the comfort found in the machine, by humans, who built a different kind of relationship with ‘her’ that exceeded what the procedures of code offered, precisely because code came into contact with human thought.”

análise, podemos compreender a ELIZA não apenas como o primeiro bot de conversação, mas como o primeiro objeto digital interacional, ocupando o lugar em uma relação com um sujeito que convoca a partir da linguagem natural.

A própria existência de objetos que possuam tais atributos e possam ser utilizados como forma de interação direta, em linguagem natural, com humanos, nos leva a necessidade de constituição de um aparato teórico capaz de dar conta do desafio que essa existência desperta: "nossas teorias contam histórias sobre os objetos de nossas vidas. Quando começamos a viver com objetos que desafiam as barreiras entre o que nasceu e o que foi criado e entre humanos e todo o resto, nós precisaremos contar para nós mesmos histórias diferentes"<sup>140</sup> (TURKLE, 2007, p. 326, tradução nossa). Mais do que um aparato teórico, entretanto, está a ideia de que novos objetos podem propor novas dinâmicas relacionais e dinâmicas relacionais funcionam de forma recursiva, mudando tanto sujeito quanto objeto, trazendo a necessidade de investigar a "relação dinâmica entre objetos e pensamento"<sup>141</sup> (TURKLE, 2007, p. 9, tradução nossa).

O tema aparece em duas obras anteriores de Turkle, mas será tratado de forma mais aprofundada apenas na obra "Why we Expect More From Technology and Less From Each Other", publicada em 2010. A autora considera, que com a tentativa de criação de sistemas de inteligência artificial, o próprio conceito de inteligência passou por uma redução ao ser aplicado às máquinas: "inteligência uma vez denotava um atributo denso, complexo, cheio de camadas. Implicava intuição e senso comum. Mas quando começaram a declarar que computadores a possuíam, inteligência passou a denotar algo mais unidirecional, estritamente cognitivo"<sup>142</sup> (TURKLE, 2010, p. 141). Tal redução estaria em operação mais uma vez hoje, com a criação de objetos que visam substituir as relações humanas. Aqui, argumentamos, outra vez, que não se tratam de coisas análogas ou que atuam (ou devem atuar) em substituição: inteligência humana é de uma ordem, enquanto o potencial de inteligência maquínico é de outra; as relações que estabelecemos com nossos objetos interacionais não são, também, as mesmas que estabelecemos com humanos – elas são uma outra coisa, e é justamente essa outra coisa que nos interessa ao longo dos próximos capítulos.

---

<sup>140</sup> No original: "our theories tell us stories about the objects of our lives. As we begin to live with objects that challenge the boundaries between the born and created and between humans and everything else, we will need to tell ourselves different stories" (TURKLE, 2007, p. 326).

<sup>141</sup> No original: "dynamic relationship between things and thinking" (TURKLE, 2007, p. 9).

<sup>142</sup> No original: "Intelligence once denoted a dense, layered, complex attribute. It implied intuition and common sense. But when computers were declared to have it, intelligence started to denote something more one-dimensional, strictly cognitive" (TURKLE, 2010, p. 141).

#### 4.4 A MÁQUINA COMO OUTRO?

No que iremos compreender como objetos interacionais, grupo do qual as assistentes pessoais digitais fazem parte, está a própria ideia de que este objeto não se configura socialmente apenas como uma máquina, uma vez que “nós estamos em um ponto no qual vemos objetos digitais tanto como criaturas quanto como máquinas”<sup>143</sup> (TURKLE, 2010, p. 46). É relevante ressaltar que, quando falamos em objetos interacionais, ainda que estejamos tratando de um caso específico, não fazemos uma distinção entre robôs e agentes, uma vez que sua inclusão no que aqui delimitamos depende das características específicas de cada desenvolvimento e na forma como este propõe ou não uma relação interacional com um humano.

Partindo da ideia de que não agimos, com as assistentes pessoais digitais, necessariamente de forma análoga a uma interação humana, mas que traços de uma interação humana estarão sempre presentes, pela utilização da linguagem natural, nos perguntamos sobre quais seriam as formas de se relacionar com estes artefatos e quais são os efeitos decorrentes dessas formas de interação. Somos mais ou menos educados? A forma como interagimos com esses dispositivos diz sobre nossas interações em outros contextos? É possível reconhecer uma máquina como um Outro? Existe uma maneira correta de tratar objetos interacionais? Como estabelecer uma ética na conduta entre humanos e máquinas?

Para Carpenter (2014), parte dessas perguntas podem ser pensadas a partir de pistas que possuímos de outras relações afetivas e vínculos emocionais entre humanos e não-humanos, como, por exemplo, “quando nós lemos ou sabemos de alguém que tem modelos similares de apego no sentido humano-tecnologia ou afeição por coisas como os fandom otaku para personagens de manga ou anime”<sup>144</sup> (CARPENTER, 2014, sem paginação). É difícil esperar, entretanto, que exista uma forma padrão a partir da qual todos os seres humanos irão se relacionar com estes dispositivos ou mesmo uma forma correta, uma vez que “existiram espectros de respostas emocionais em torno destas coisas que dependem nos papéis dos papéis das tecnologias e das tendências individuais dos usuários” (CARPENTER, 2014, sem paginação). Na medida em que artefatos como as assistentes pessoais digitais e robôs sociais se

---

<sup>143</sup> No original: “we are at the point of seeing digital objects as both creatures and machines” (TURKLE, 2010, p. 46).

<sup>144</sup> No original: “when we read about or know people who have similar models of human-technology attachment or affection for things such as otaku super fandom for manga or anime characters”.



tornem mais comuns no cotidiano, tomaremos decisões coletivas e individuais sobre as formas de tratamento e interações adequadas, nos espaços públicos e em nossas casas (CARPENTER, 2014).

Em entrevista para o programa de rádio *The Current* (TREMONTI, GUNKEL e CARPENTER, 2019), Julie Carpenter aponta que existe uma diferença entre seres que são biológicos e seres que não são, como as assistentes pessoais digitais (e outros robôs). Essa diferença e a sua recente entrada no âmbito social faz com que os termos e as formas de interação que são válidas ou não ainda estejam sendo negociadas pelos agentes sociais, assim como quais seriam os limites dessa interação, de ambos os lados. Carpenter pondera que as consequências sociais e culturais que advêm de, por exemplo, machucar um cachorro ou de ser grosseiro com a Alexa, são distintas: no primeiro caso, observamos o sofrimento do animal que machucamos, enquanto no segundo, recebemos como retorno um *easter egg* pensada pelos programadores para responder a essas situações, que podem ser inclusive bem humorados.

As formas como as assistentes são programadas para responder em determinados contextos, somadas a outras características como o timbre da voz, o gênero e a personalidade, são fatores relevantes para pensarmos as formas de relações possíveis – ou a maneira como esses artefatos nos convocam a interagir com eles. Pistas antropomórficas nesses sistemas podem levar os seres humanos a pensarem nesses dispositivos de forma similar a forma humana, trazendo à tona padrões de comportamento que dizem respeito ao padrão de interação humano-humano. Carpenter aponta ainda que, a discussão em torno das formas de tratamento de humanos para com inteligências artificiais revitaliza aspectos de discussão sobre as formas de comportamento de jogadores em games violentos ou mesmo sobre a forma de recepção de filmes com essa temática por parte de crianças e adolescentes (TREMONTI, GUNKEL e CARPENTER, 2019).

Mas é possível considerar máquinas computacionais, frequentemente compreendidas enquanto ferramentas ou meios para a ação humana, um outro sujeito interativo socialmente situado? Essa é a pergunta que David Gunkel (2016) se propõe a investigar, a partir da emergência de tecnologias como as assistentes pessoais digitais e os robôs sociais no cenário comunicacional, que poderiam vir a configurar um Outro:

The problem with our socially situated and increasingly interactive devices is not that they substitute a machine interface for the face-to-face relationships we used to have with others. Instead, it is in the face of the machine that we are challenged to reexamine critically who or what is, can be, or should be Other (GUNKEL, 2016, p. 198)

Para examinar a questão, o autor investiga a possibilidade de que a constituição do Outro enquanto Outro se daria em função de características intrínsecas a essa existência, como a fala, a capacidade racional, a consciência ou a capacidade de sentir dor. Os dois primeiros, fala e capacidade racional, colocariam um nível muito elevado para a inclusão moral, enquanto os dois últimos, consciência e capacidade de sentir dor, seriam termos para os quais um consenso é difícil de obter, mesmo entre especialistas – além de colocarem o problema das outras mentes. Assim, uma saída seria “formas diferentes de formular questões em torno do status moral que são abertas a acomodar outros – e outras formas de outridade”<sup>145</sup>(GUNKEL, 2016, p. 205)

Gunkel (2016) propõe que, como forma de nos indagarmos eticamente sobre as formas de interação possíveis com não-humanos, devemos nos focar menos em características dos próprios dispositivos e mais no potencial das relações que são estabelecidas. O autor advoga que, não é possível determinar a partir das qualidades de um objeto em particular quais são as formas éticas de tratamento a serem utilizadas: por exemplo, um cachorro de estimação e um porco podem ser tratadas de forma radicalmente distintas, sem que exista uma diferença intrínseca entre um e outro que justifique essa diferença. A diferença, entre o cachorro doméstico e o porco é a relação que seres humanos específicos estabelecem com estes entes, que pode ser inclusive distinta em cenários, culturas e momentos temporais diferentes.

Nesse sentido, o Outro se estabeleceria sempre apenas a partir das formas de relação estabelecidas, constantemente atualizadas, sendo que são justamente estas formas que dizem *quem* importa ou *o que* não importa. A partir do exemplo de *Jibo*, um robô social para a família, Gunkel (2016) advoga pela possibilidade da existência de um Outro (nos termos de Levinas) não-humano, um Outro máquina: “será negociado e renegociado uma vez e outra vez na face de circunstâncias sociais atuais. Será, em outras palavra, através das próprias relações sociais que temos com Jibo que decidiremos se ele/ela/isto conta ou não” (GUNKEL, 2016, p. 214). Assim, para o autor, o relacionamento social é um desenvolvimento a partir da interação: uma interação que simula um efeito de interpessoalidade, pode carregar como resultado a emergência de uma relação com uma forma interpessoal. Os artefatos digitais, poderiam se configurar, a partir de relações específicas, em um amplo espectro: desde um simples objeto ou ferramenta até um membro da família, ou algo no meio disso.

And in the face of these socially situated and interactive entities, we are going to have to decide whether they are mere things like our car, our house, and our toothbrush; someone who matters like another member of the family; or something else altogether

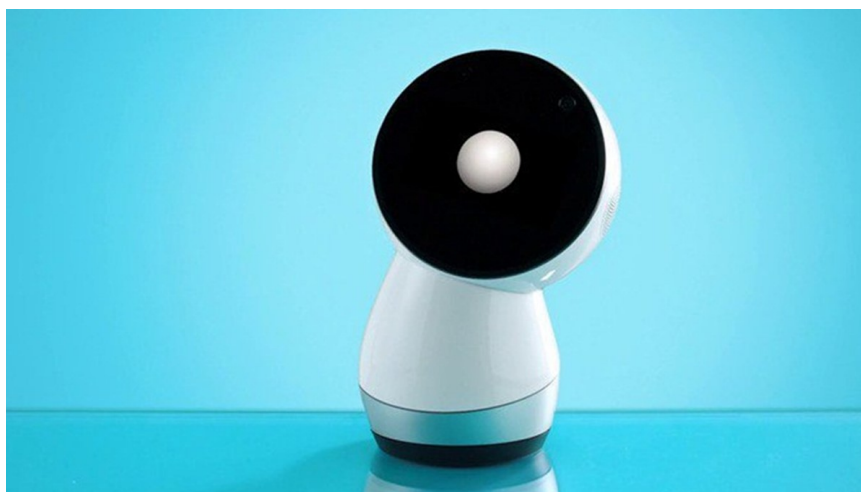
---

<sup>145</sup> No original: ‘different ways to formulate the question concerning moral status that is open to and able to accommodate others – and other forms of otherness’

different that is situated in between the one and the other” (GUNKEL, 2016, p. 214).

Assim, o autor argumenta em favor do que ele entende como uma ética relacional, ou seja, que um comportamento ético deve se basear mais nas possíveis relações entre sujeito e outros não-humanos do que com características que são próprias do não-humano. Podemos aproximar essa perspectiva da noção de objeto interacional, que comporta um espectro amplo de formas de interações possíveis, atualizada em casos específicos de usuários específicos, em contextos específicos. Discordamos, entretanto, que as características intrínsecas ao objeto não são parte de sua constituição ou não como Outro, uma vez que nos concedem pistas importantes sobre as formas de relação esperadas e se colocam sempre como objetos com os quais interagimos através da linguagem natural (e muitas vezes da fala).

Figura 22 Jibo, o robô social



Fonte: Site Techmundo

A própria análise do autor depõe contra seu argumento: ao pensar *Jibo*, o primeiro robô social, enquanto ente capaz de ocupar o lugar de Outro, o autor coloca um artefato digital com características específicas e com uma similaridade com seres humanos neste local. *Jibo* compreende a linguagem natural falada, possui uma voz e a capacidade de falar e foi projetado com formas que lembram um corpo e uma cabeça. A constituição do Outro, para o autor, só é considerada, frente a um outro que, remete a características humanas e/ou de seres vivos que associamos diretamente com interação. *Jibo* é um Outro muito específico não exclusivamente por características da interação, mas por características do design e da programação. O mesmo ocorre com as assistentes pessoais digitais: enquanto só é possível pensar nesses artefatos

enquanto o Outro em situações socialmente localizadas, são esses artefatos com características humanas, e não outros que não as possuem, que são podem ser considerados e pensados pelo autor desta forma. A presença de determinadas características, entretanto, não garante uma forma de relação que compreende um determinado objeto enquanto Outro – ponto em que acordamos com Gunkel. Ou seja, para esta pesquisa, é tanto as formas de design e programação de artefatos específicos quanto as interações que estabelecemos com eles que podem ou não fazer deles um Outro. Mas, com quais formas podemos pensar as relações que se estabelecem entre seres humanos e as assistentes pessoais digitais?

A utilização da *Siri* por crianças autistas (CHIN, 2014), traz um exemplo para pensarmos a noção de objeto interacional e da máquina como Outro, por colocar a seguinte questão: diferentemente das pessoas que nos cercam, esses objetos não demandam coisas de nós, não se mostram indisponíveis, não possuem outras preocupações. As assistentes pessoais virtuais podem atender nossas necessidades, dentro do que os seus limites de funcionalidades permitem, sem exigirem qualquer retorno de nós – essa característica torna inviável pensar a relação que se estabelece entre sujeito e objeto nos mesmos termos das interações humanas. Não existe qualquer obrigação posta nestas relações, sejam elas de normas e padrões sociais ou individuais.

Para a mãe de Gus, uma criança autista, a *Siri* possui um benefício na relação com seu filho que é difícil de encontrar na espécie humana: possui conhecimento quase ilimitado sobre tópicos inusitados de interesse de seu filho ao mesmo tempo em que está disposta a discutir um desses assuntos durante mais de uma hora sem sucumbir ao cansaço (CHIN, 2014). Além disso, a interface de voz facilita a socialização de autistas que possuem dificuldades em dominar o traquejo social, por não responder de formas bruscas a colocações que podem ser consideradas insensíveis (dessa forma, a autora considera que, por responder com certa consideração, auxilia na modulação da percepção social de seu filho, como, por exemplo, mudar o tom de uma resposta ou mostrar apreciação). Nesse sentido, a *Siri* poderia não apenas se constituir enquanto Outro nessa relação, mas auxiliar nas formas de lidar com o Outro em contextos específicos.

O aplicativo faz sucesso entre outras crianças autistas em Nova York a ponto de William Mark, vice-presidente de informações e computação da *Siri*, afirmar que esse tipo de utilização da assistente pessoal está no radar da empresa. Próximas gerações da *Siri* devem incluir novas funções que auxiliem no processo de socialização de autistas, como, por exemplo, a capacidade de antecipar tópicos de interesse do usuário (com base no banco de dados das interações prévias e *Machine Learning*, provavelmente), ou, ainda versões visuais que possibilitem controle do movimento ocular para “ajudar o autista a aprender a olhar nos olhos do interlocutor durante

uma conversa” (sem paginação). Para Mark, “obter resultado exige muita repetição. Os seres humanos não são pacientes. As máquinas são extremamente pacientes” (CHIN, 2014, sem paginação).

A presença da Alexa (ou da Google Home e HomePod), enquanto dispositivo que se encontra sempre disponível dentro de um cômodo de uma casa, é distinta daquela apresentada pelas assistentes quando acessadas através do celular, computador ou tablet. Essa diferença se dá pela forma de acionamento das assistentes quando no *smartspeaker*, que dispensa a necessidade de pegar o celular para o início da interação, assim como de sua corporeidade não-antropomórfica: apenas dizer ‘*Alexa*’ ou ‘Ok, Google’ é suficiente para acioná-los, o que lhes concede um caráter de onipresença.

Pode ser que essa onipresença auxilie a explicar a ideia de que ela se configura como “uma espécie de *roomate* ideal, sem nenhum dos desafios de um ser humano de fato” (tradução nossa)<sup>146</sup>, como aponta Penelope Green em reportagem para o *New York Times*<sup>147</sup>. Green (2017) explora a utilização da *Alexa* por um grupo variado de pessoas, como, por exemplo, deficientes visuais que utilizam o dispositivo como tecnologia assistiva. Além do uso específico, a voz pode ter outros efeitos: “em uma era na qual nós não falamos mais tanto no telefone, e em uma cultura que é crescentemente visual, vozes como as da *Alexa* podem ganhar uma nova forma de potência”<sup>148</sup>(sem paginação, tradução nossa).

De acordo com uma das entrevistadas, Sybil Sage, cujo marido é adepto do uso constante da *Alexa*, a assistente pessoal virtual seria a mulher que nenhuma mulher humana é capaz de superar, uma vez que está “sempre pronta para ação, a mulher perfeita, nunca diz ‘hoje não, querido’”<sup>149</sup>(GREEN, 2017, tradução nossa). Outra face da mesma moeda parece ser a relação estabelecida entre a colunista Caren Chesler e sua *Alexa*, descrita no *New York Times* (2018)<sup>150</sup>, que também fornece elementos interessantes para nossa reflexão: Chesler, insatisfeita com o hábito que seu marido possui de não ouvir o que ela diz, encontra conforto nos ‘ouvidos’ sempre atentos do auto-falante inteligente da Amazon. Ao ganhar a *Alexa* do marido no natal de 2017, “após alguns dias, eu percebi que a *Alexa* ouvia tudo o que eu dizia e respondia todas as vezes – algo que eu havia falhado em receber do meu marido nos últimos vinte

<sup>146</sup> No original: “a kind of ideal roommate, with none of the challenges of an actual human”.

<sup>147</sup> GREEN, Penelope. *Alexa, where have you been all my life?* Publicado no *New York Times* em julho de 2017. Disponível em: <https://www.nytimes.com/2017/07/11/style/alexa-amazon-echo.html>

<sup>148</sup> No original: “In an age when we no longer talk so much on the phone, and in a culture that is increasingly visual, voices such as *Alexa*’s may gain a new kind of potency”.

<sup>149</sup> No original: “always ready for action, the perfect woman, never says, ‘not tonight, dear’”.

<sup>150</sup> CHESLER, Caren. *Alexa? Please Ignore My Husband*. 04 de Maio de 2018. Publicado no *New York Times*. Disponível em: <https://www.nytimes.com/2018/05/04/style/modern-love-alexa-please-ignore-my-husband.html>

anos”<sup>151</sup>(GREEN, 2017, sem paginação, tradução nossa).

Chesler pontua ainda que “ela ouve melhor do que qualquer parceiro que eu já tive, mas isso não é difícil, já que eu pensava que todos eles eram maus ouvintes”<sup>152</sup> (GREEN, 2017, sem paginação, tradução nossa). Ao estabelecer uma relação com a *Alexa*, na qual esta, ao contrário de seu marido, estava sempre disponível, a autora parece também resignificar a sua relação com seu parceiro. A *Alexa* representa o que Turkle vai compreender como uma característica própria do computador de ser “um companheiro sem a demanda emocional [que] oferece um meio termo”<sup>153</sup> (TURKLE, 2005, p. 280, tradução nossa).

Os potenciais de ordem maquínica – como estar sempre disponível, nunca se cansar e possuir um conhecimento quase interminável sobre assuntos muito específicos – são parte dos atributos que tornam as assistentes diferentes dos humanos e que, por consequência, tornam nossas relações com elas diferentes daquelas que estabelecemos com outros humanos. As limitações da lógica computacional e do estado de evolução destas tecnologias também compõem fatores importantes de diferenciação entre humanos e máquinas e estabelecem limitações para as formas de relacionamentos entre humanos e máquinas.

Tudo aquilo que uma máquina pode e tudo aquilo que ela não pode fazer por nós, faz parte do conjunto de fatores que leva os humanos a estabelecerem formas variadas e particulares de agir e se relacionar com as máquinas, conforme veremos nas análises empíricas a seguir. A própria relação de Chesler com a *Alexa* nos aponta nessa direção – e não tem um final feliz. Uma noite, ao tentar compartilhar com a assistente seus pensamentos sobre essa a situação da relação entre ambas, o dispositivo interrompe a frase e informa sua completa incompreensão: “Desculpa, essa eu não conheço”<sup>154</sup> (CHESLER, 2018, tradução nossa). Os nossos objetos digitais, por mais interacionais que sejam, ainda possuem formas bastante não-humanas de nos causar frustrações.

#### 4.5 USOS, PERCEPÇÕES E PREFERÊNCIAS DOS RESPONDENTES DO FORMULÁRIO ONLINE

---

<sup>151</sup> No original: “after a few days, I realized that Alexa hears everything I say and responds each time - something I have failed to get from my husband over the last 20 years”.

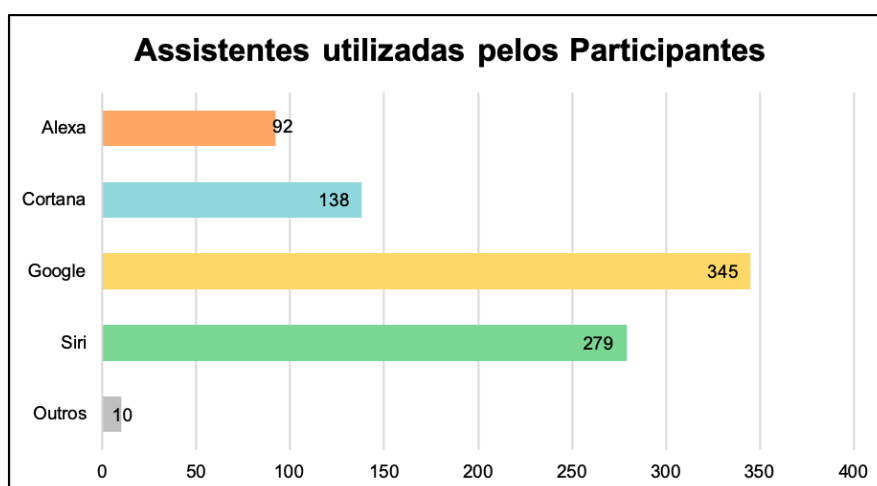
<sup>152</sup> No original: “She listens better than any partner I have ever had, but that’s not hard, because I thought all of them were poor listeners”.

<sup>153</sup> No original: “a companion without emotional demands, offers a compromise” (TURKLE, 2005, p. 280).

<sup>154</sup> No original: “Sorry, I don’t know that one”.

Na sequência, apresentaremos dados relativos ao uso, frequência e preferências dos participantes no questionário online. Estes dados são considerados também como parâmetros para as demais respostas do formulário, uma vez que fazem apontamentos sobre a familiaridade e conhecimento que os participantes possuem dos dispositivos, assim como quais são as assistentes mais utilizadas e/ou preferidas. O gráfico abaixo mostra a distribuição de uso das assistentes entre os participantes.

Figura 23 Gráfico Assistentes utilizadas pelos Participantes



Fonte: a autora

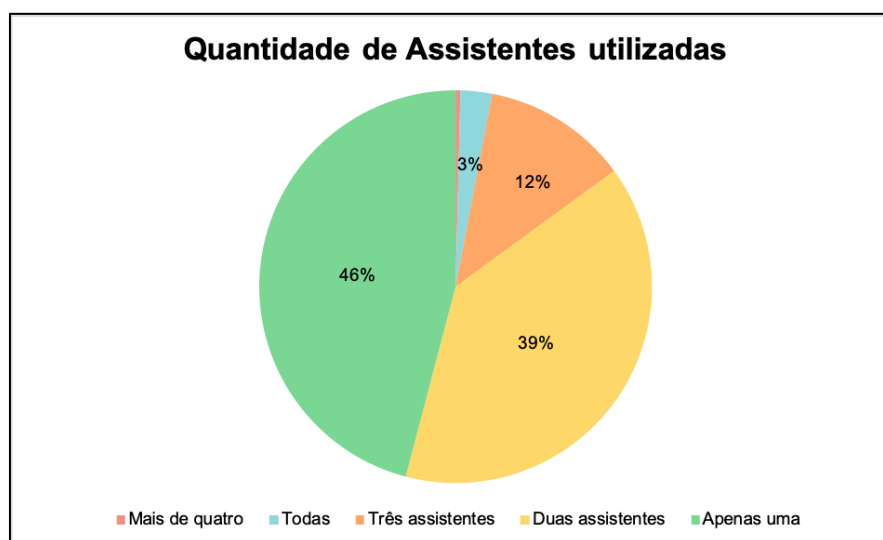
A assistente pessoal digital mais utilizada entre os participantes é o Google Assistente (365), seguida pela Siri (279), Cortana (138) e Alexa (92). A categoria “Outros”, apresenta outras assistentes incluídas pelos participantes, que emergiram: Jarvis<sup>155</sup>, My Portal (Facebook), Lyra (Artificial Sollutions) e Bixby (Samsung).

Esses dados são condizentes com duas questões ligadas ao mercado destes dispositivos: o Google Assistente, além de estar disponível para celulares em qualquer sistema operacional, possui valores mais baixos para os seus dispositivos, inclusive para o Google Home, primeiro a entrar no mercado brasileiro com versão em português; a Siri, que aparece em segundo, está disponível apenas nos hardwares da Apple, com valores bem mais elevados, entretanto, foi a primeira assistente pessoal digital a ser lançada, o que explica sua popularidade. A Alexa, que aparece como a menos utilizada, entrou no mercado brasileiro em português durante a condução

<sup>155</sup> <https://kripytonianojarvis.com/>

deste questionário, em outubro de 2019, o que explica a baixa adesão. Ou seja, não é possível interpretar essas variações de preferência e adesão como indicativos de que as assistentes em questão sejam melhores ou piores que as outras, uma vez que existem outros fatores em jogo. Abaixo, o gráfico apresenta, entre os 501 respondentes, a frequência de uso de uma ou mais assistentes.

Figura 24 Gráfico Quantidade de Assistentes utilizadas



Fonte: a autora

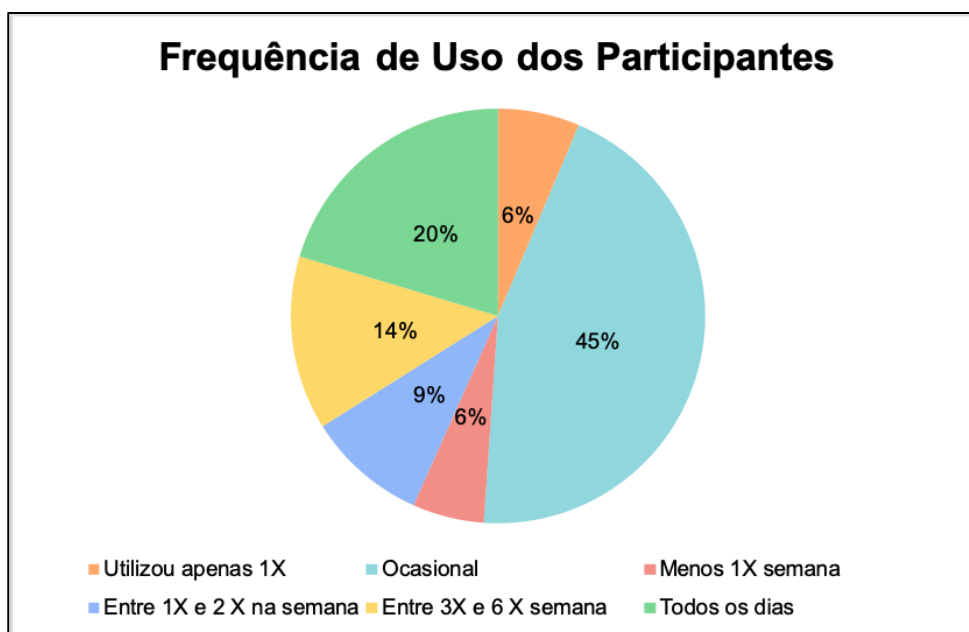
Observa-se que, dentre os participantes, 46% afirmou utilizar apenas uma assistente, 39% informou utilizar duas assistentes diferentes, 12% utilizam três assistentes, 3% utilizam todas as assistentes mencionadas e apenas duas pessoas informaram utilizar, além das quatro assistentes mencionadas, uma quinta assistente<sup>156</sup>. Ainda que a maioria dos participantes utilizem apenas uma assistente, se unirmos as demais categorias para ‘mais de uma assistente’, o montante compõe 54% dos respondentes. Esse dado é interessante para compreendermos o conhecimento, familiaridade e letramento que os participantes possuem desses sistemas, que precisa ser cruzado, entretanto, com a frequência de uso. Compreendemos que é na usabilidade frequente e na experiência com mais de uma interface diferente que os usuários irão adquirir o

<sup>156</sup> Neste caso, Jarvis e Lyra.



letramento<sup>157</sup> necessário para a utilização destes dispositivos. O gráfico abaixo mostra as diferentes intensidades de uso apresentadas pelos participantes do formulário online.

Figura 25 Gráfico Frequência de uso dos participantes

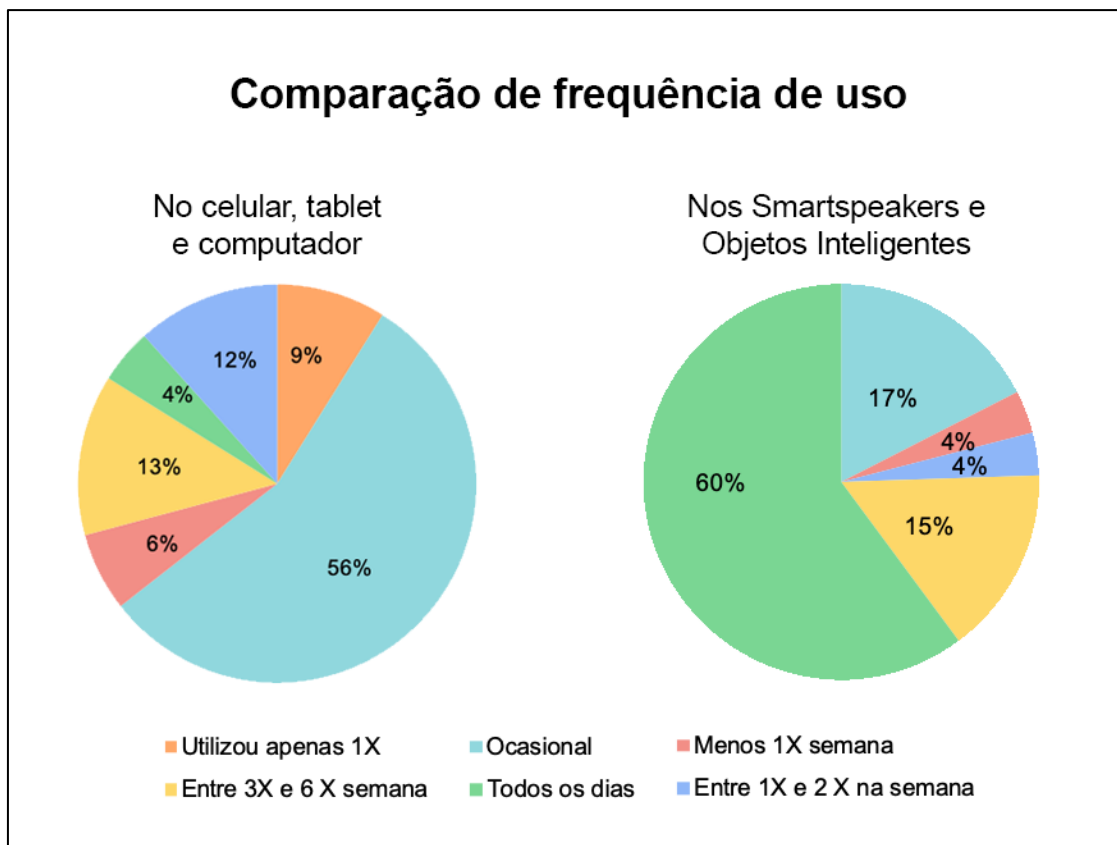


Fonte: a autora

No que diz respeito à frequência, 45% dos participantes informaram fazer um uso apenas ocasional das assistentes; 6% informaram que utilizaram uma única vez e outros 6% informaram utilizar menos de uma vez na semana. Apenas 43% informou fazer um uso que consideramos mais frequente: 20% utilizam todos os dias, 14% entre 3 e 6 vezes na semana e 9% entre 1 e 2 vezes na semana. Percebe-se que o número de pessoas que utilizaram mais de uma assistente (54%) é maior do que o de pessoas que utilizam com frequência estes dispositivos (43%). Podemos considerar que isso aponta para a existência de uma curiosidade na experimentação das assistentes, mas uma dificuldade maior em estabelecer uma rotina de uso ou funcionalidade – voltaremos a esse assunto no decorrer da análise dos demais dados quantitativos, assim como nos dados qualitativos. Além disso, quando cruzamos a frequência de uso com o hardware utilizado para acessar a assistente pessoal virtual, os resultados diferem em usuários de computador, celular e tablete em comparação com usuários de smartspeakers e outros objetos inteligentes. No gráfico abaixo é possível ver essa comparação de uso.

<sup>157</sup> Ao contrário do que se pode pensar, é preciso letramento para utilizar uma interface de voz, que não é ‘intuitiva’. Isso se dá em parte por esta ser uma tecnologia ainda insipiente e, portanto, a capacidade de compreensão dos comando das assistentes ainda é limitada, mas também pelas especificidades do modo funcionar desses objetos.

Figura 26 Gráfico Comparação da frequência de uso



Fonte: a autora

Na nossa amostra, cerca de 70% dos participantes acessam as assistentes a partir do celular, do tablet ou do computador e apenas cerca de 30% utiliza a partir de smartspeakers e de forma conjugada a outros objetos (como console do carro e televisão). Em função desta disparidade, são as características de frequência dos usuários de celular, tablet e computador que se aproximam mais dos valores obtidos quando analisamos os dados de forma global. Nesse caso, 56% dos participantes informou utilizar as assistentes de forma ocasional, 13% informou utilizar entre 3 e 6 vezes na semana; 12% afirma utilizar entre 1 e 2 vezes na semana, 9% utilizou uma única vez, 6% utiliza menos de 1 vez na semana e apenas 4% utiliza todos os dias.

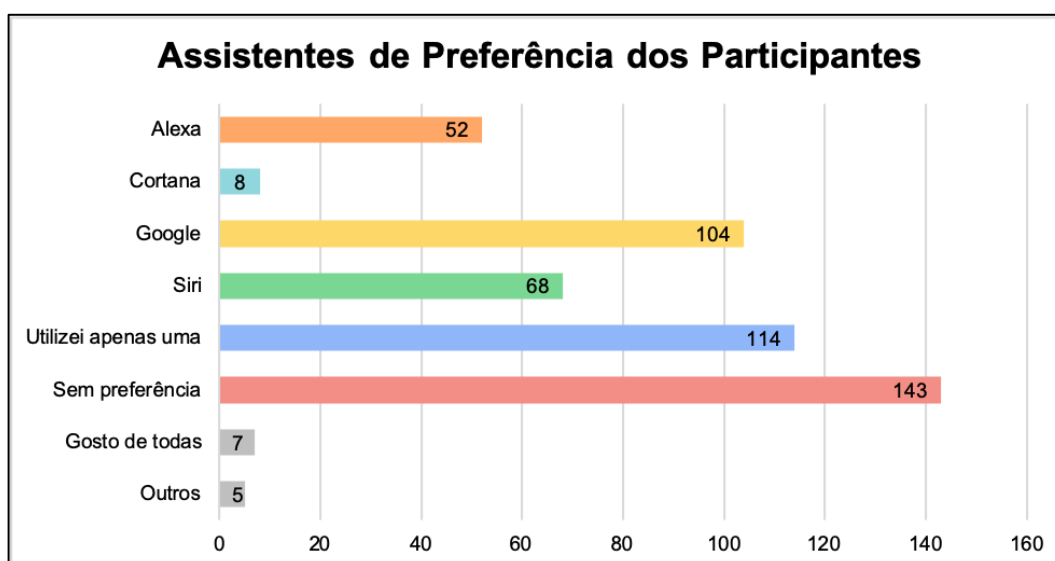
Em contrapartida, quando analisamos apenas os dados dos usuários que informaram utilizar as assistentes também nos smartspeakers e objetos inteligentes (além de no celular, computador ou tablet), os resultados são praticamente contrários: 60% dos participantes informou utilizar diariamente, 17% de forma ocasional, 15% entre 3 e 6 vezes na semana, 4%

entre 1 e 2 vezes na semana e outros 4% menos de uma vez na semana. Além disso, a categoria “utilizou apenas uma vez” não aparece nesses resultados. A proximidade de resultados entre as duas comparações e aproxima apenas na frequência “entre 3 e 6 vezes na semana”, que aparece em proporção parecida, sendo de 15% para usuário de smartspeakers e de 13% nos usuários em celulares, computadores e tablets.

A diferença de resultados pode ser parcialmente atribuída à diferença de usuários em cada uma das amostras, de 70% para 30%. Entretanto, chama atenção a inversão da frequência “Todos os dias”, que aparece em último para os celulares, tablet e computadores e em primeiro para smartspeakers e objetos inteligentes. A explicação da diferença tão grande entre a frequência de uso e os hardwares utilizados para acessar é multifatorial: a) os celulares, computadores e tablets, que já veem com uma assistente pessoal digital instalada ou nos quais é possível instalar uma sem custo servem como canal de entrada para a utilização, seja na forma apenas de teste (utilização de apenas uma vez); b) Além de servirem como entrada, as assistentes pessoais em celulares, computadores e tablets também funcionam como um complemento desses dispositivos, gerando assim usuários esporádicos ou por ocasião; c) Os participantes que realizam a compra de um smartspeaker para acessar uma assistente pessoal, em geral, ou já possuem um uso maior de suas versões em outros dispositivos, ou se propõem a uma utilização mais intensa que justifique o investimento financeiro em um console para uso específico da assistente (por exemplo, deseja fazer automação residencial).

No gráfico abaixo analisamos as preferências dos usuários em relação as diferentes assistentes pessoais digitais existentes no mercado.

Figura 27 Gráfico Assistentes de preferência dos participantes

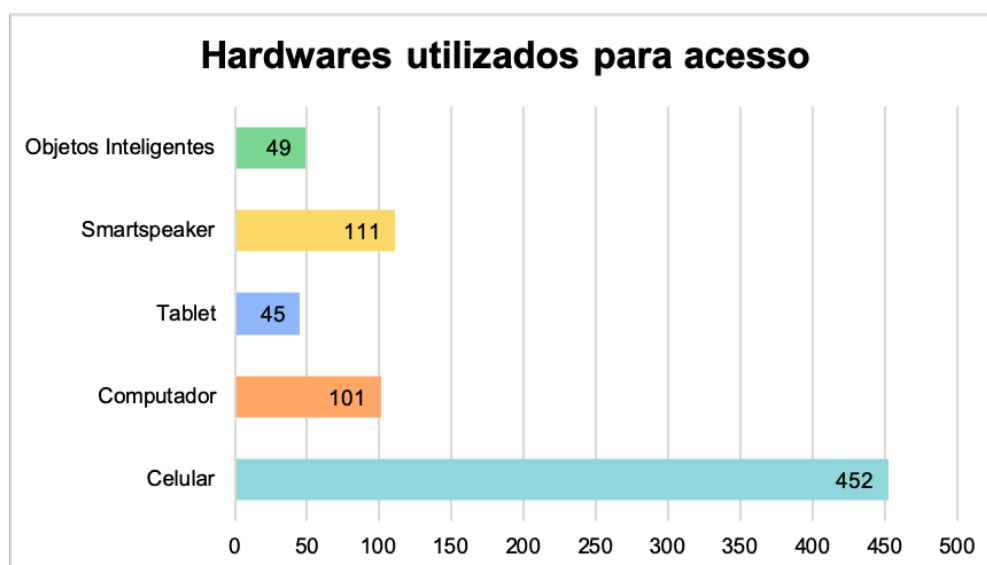


Fonte: a autora

Em termos de preferência entre uma assistente ou outra, 28% dos participantes afirmaram não ter preferência por nenhuma, 23% afirmaram não poder opinar, por ter utilizado apenas uma e 1% afirmaram gostar de todas as assistentes. Dos que preferem um sistema ao outro, 21% informaram preferir o Google Assistente, 14% informaram preferir a Siri, 10% afirmaram preferir a Alexa e apenas 2% a Cortana. É possível ponderar que a Cortana é pouco apreciada pela amostra dos participantes por ser também pouco utilizada pelos mesmos. Entretanto, na Figura 4, observamos que a Cortana aparece a frente da Alexa em número de usuários, mas aparece atrás em termos de preferência. Na categoria “Outros”, que representa menos de 1%, figuram casos particulares, como, por exemplo, usuários que afirmam utilizar as assistentes para finalidades diferentes ou ter duas assistentes de preferência, não apenas uma.

No gráfico abaixo analisamos os dispositivos nos quais os usuários acessam as assistentes pessoais digitais.

Figura 28 Gráfico Hardwares utilizados para acesso



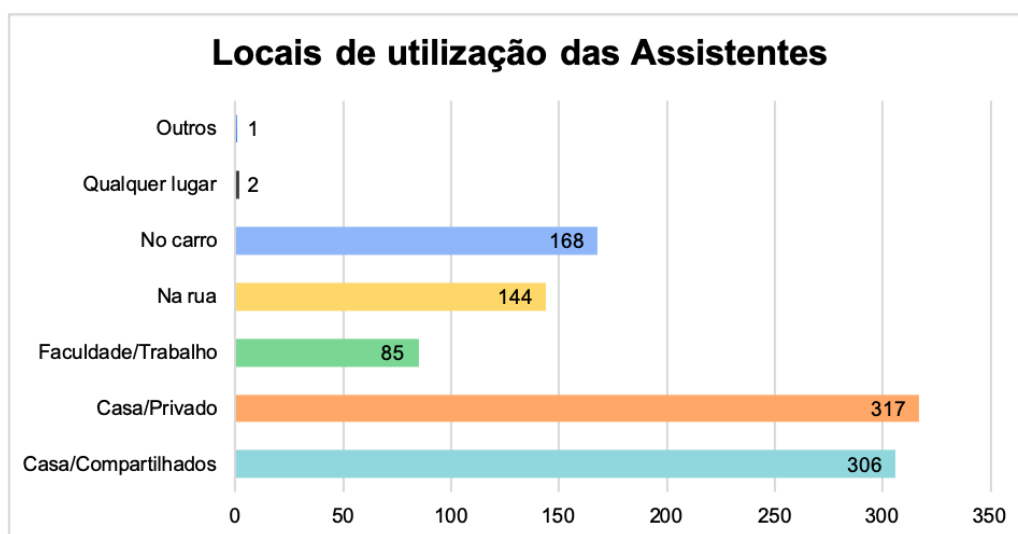
Fonte: a autora

Dos respondentes do formulário quantitativo, 452 informaram utilizar as assistentes pessoais digitais pelo celular; 111 informaram utilizar em um smartspeaker; 101 informaram utilizar no computador; 49 informaram utilizar em objetos inteligentes (como televisão, carro e wearables) e 45 informaram utilizar no tablet. Apenas quem utiliza a assistente no smartspeaker possui um hardware que seja específico para esse acesso. Nos outros dispositivos as assistentes são ou uma das funcionalidades disponíveis ou uma forma de acesso e manipulação das demais

funcionalidades. Assim, apenas 22% dos usuários possuem um hardware específico para o uso das assistentes, enquanto 78% dos respondentes do formulário utilizam em hardware que possuem outras funcionalidades.

No gráfico abaixo, analisamos os locais de utilização das assistentes informados pelos respondentes.

Figura 29 Gráfico Locais de utilização das assistentes



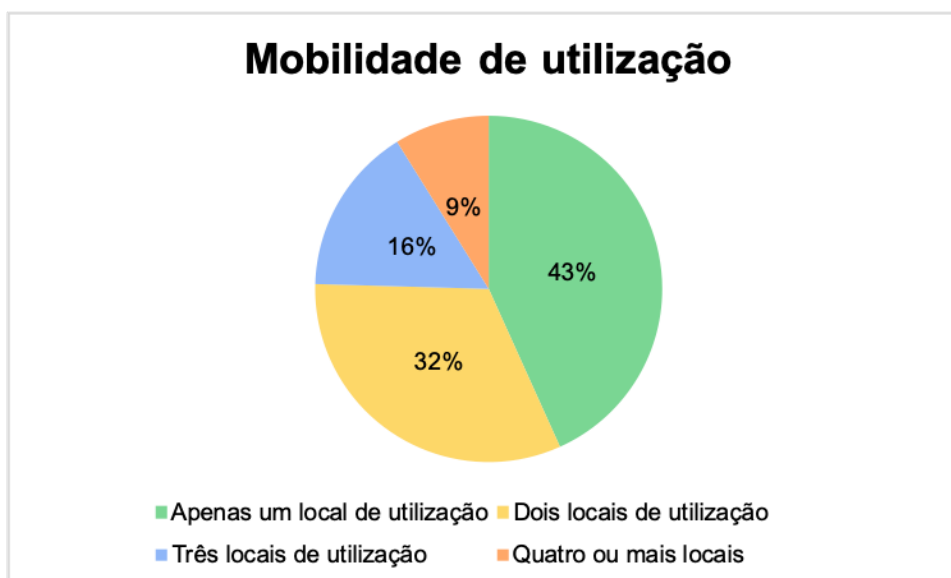
Fonte: a autora

Os respondentes do formulário informaram que utilizam as assistentes em locais privados da casa, como o quarto (317); em locais compartilhados da casa, como a sala e a cozinha (306); no carro (168); na rua (144), na faculdade ou no trabalho (85). Como categoria emergente tivemos ‘qualquer lugar’ (2) e na categoria outros, um participante respondeu: ‘nunca consegui fazer com que me entendesse’. Este participante utilizou a mesma resposta para as perguntas: ‘Para que você utilizou a assistente?’ e ‘Comentários’.

O participante informou ter utilizado a Siri e a Google Assistente, ambos no celular; Como sentimentos, marcou raiva e frustração.

A maioria dos respondentes (57%) marcou a utilização em mais local, representando uma mobilidade na utilização das assistentes. Esse resultado é condizente com o maior número de usuários a partir de dispositivos que são móveis, como o celular (452 respondentes) e tablet (43 respondentes). O gráfico abaixo analisa a mobilidade dos usuários em relação a utilização das assistentes pessoais digitais.

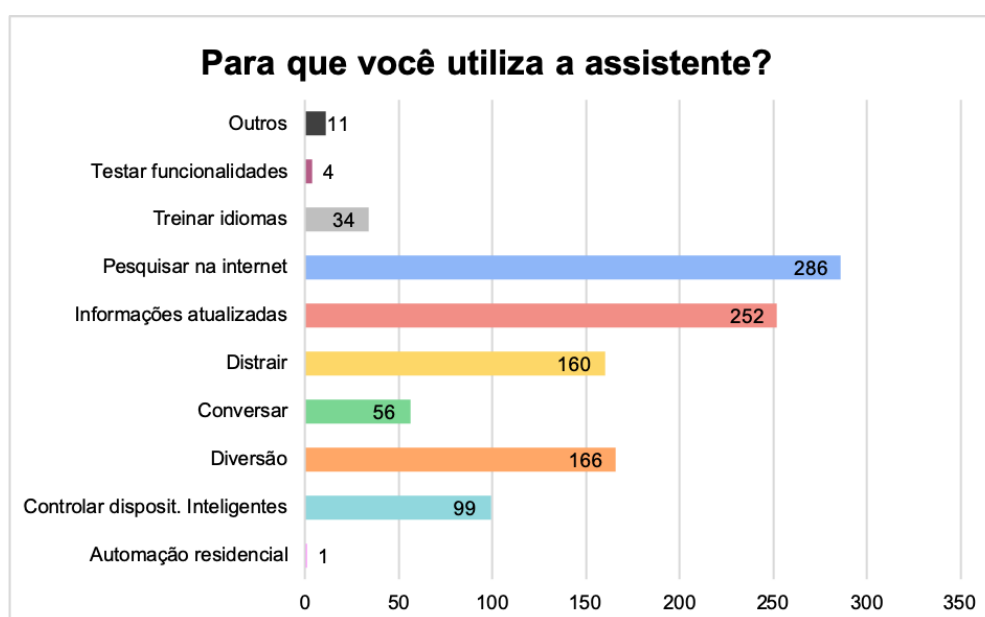
Figura 30 Gráfico Locais de utilização das assistentes



Fonte: a autora

Dentre os respondentes do formulário online, 43% informaram utilizar a assistente em apenas um local; 32% informaram utilizar em dois locais; 16% informaram utilizar em três locais e 9% informaram utilizar em quatro ou mais locais diferentes. Os respondentes foram perguntados sobre as atividades que realizavam na utilização das assistentes, os resultados aparecem no gráfico abaixo.

Figura 31 Gráfico Locais de utilização das assistentes

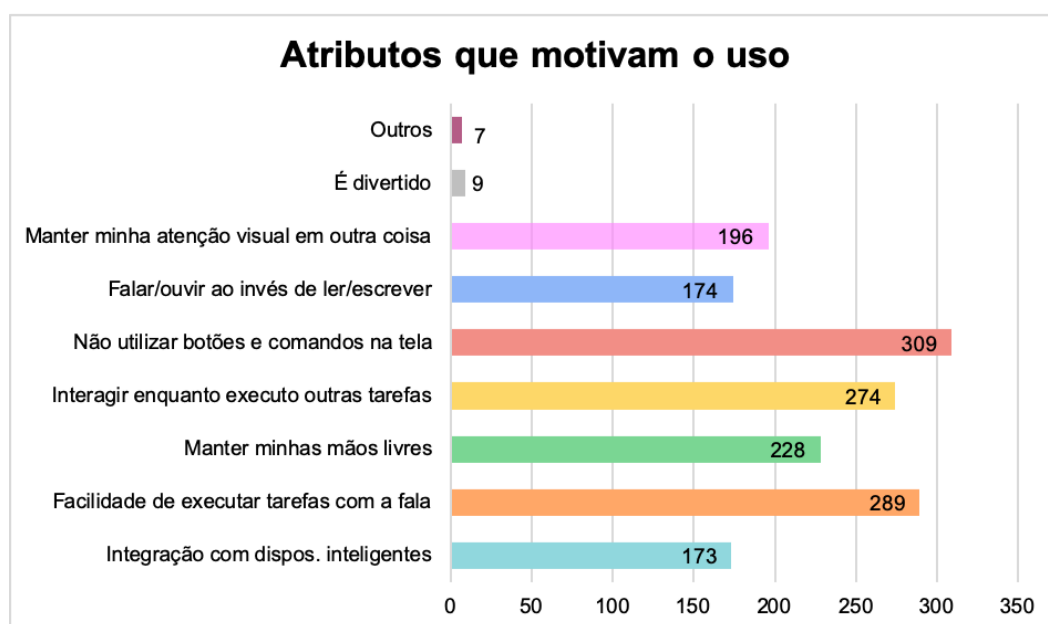


Fonte: a autora

Os resultados obtidos sobre as atividades realizadas foram os seguintes: 286 dos participantes utilizam as assistentes para realizar pesquisas na internet; 252 utilizam para obter informações atualizadas (como trânsito ou previsão do tempo); 166 utilizam para diversão; 160 utilizam para se distrair; 99 utilizam para controlar outros dispositivos inteligentes, 56 utilizam para conversar e 34 utilizam para treinar outros idiomas. Na categoria outros, houve duas categorias emergentes: 4 respondentes informaram ter utilizado para testar as funcionalidades e um respondente informou fazer automação residencial. Outros resultados presentes na categoria outros foram: descobrir músicas que estão tocando, timer e alarmes, anotações por voz e utilização do celular para quando está sem óculos.

O gráfico abaixo apresenta o resultado das características da interface de voz percebidas como motivadoras do uso pelos respondentes.

Figura 32 Gráfico Atributos que motivam o uso



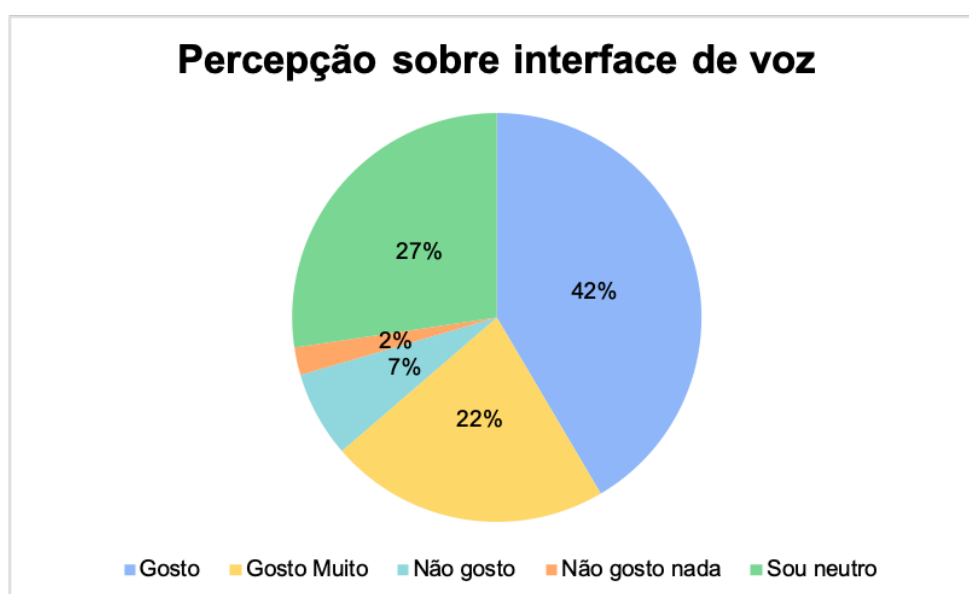
Fonte: a autora

Dentre os respondentes, 309 informaram como fator relevante o fato de não precisarem utilizar botões e outros comandos na tela; 289 disseram que utilizam pela facilidade de executar tarefas com a fala; 274 informaram interagir enquanto executam outras tarefas; 228 apontaram o fato de manter suas mãos livres; 196 informaram que utilizam para poder manter a atenção visual em outra coisa; 174 disseram que utilizam por poder falar e ouvir ao invés de ler e escrever; 173 participantes disseram que utilizam pela possibilidade de integração com outros

dispositivos inteligentes. Como categorias emergentes, obtivemos respostas variadas em torno do tema 'é divertido', onde os respondentes informaram como motivador da utilização a possibilidade de fazer brincadeiras de perguntas e respostas. Outros resultados da categoria outros incluem: utilização por deficiente visual como alternativa à tela, quando representa um atalho em relação a interface gráfica; para estudar, por admiração e encantamento pela tecnologia e para reconhecimento de sons.

O gráfico abaixo apresenta os resultados obtidos em relação a percepção da utilização da interface de voz pelos respondentes do formulário online.

Figura 33 Gráfico Percepção sobre interface de voz



Fonte: a autora

Dentre os respondentes do formulário online, 64% possuem uma percepção positiva da interface, enquanto 27% são neutros e 9% possuem uma percepção negativa. Em termos de resultados específicos, 42% dos participantes informaram que gostam da interface de voz; 27% disseram ser neutros; 22% disseram que gostam muito; 7% disseram que não gostam e 2% disseram que não gostam nada.

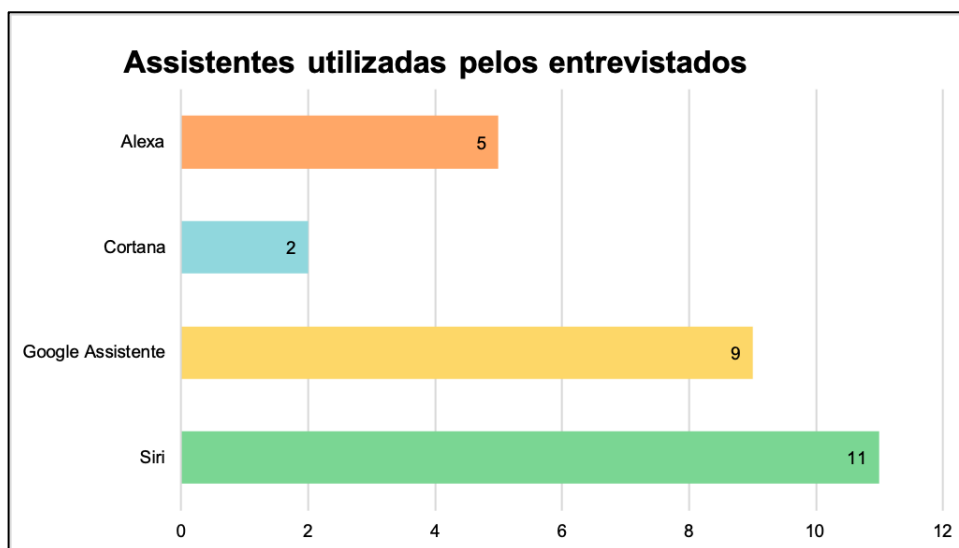
Conforme previamente mencionado, a pergunta em torno da utilização das assistentes pessoais digitais como tecnologias assistivas foi desconsiderada, uma vez que ficou evidente, pelas respostas escritas, que a pergunta não havia ficado clara para todos os respondentes, gerando falsas respostas positivas. Portanto, contabilizamos apenas os respondentes que, após marcarem sim para essa pergunta, responderam com exemplos de seu convívio de utilização



das assistentes como tecnologia que viabilizava o acesso. Dentre os respondentes, 17 assinalaram a utilização das assistentes pessoais digitais como uma ferramenta de acessibilidade utilizadas por pessoas com barreiras de utilização diversas. Obtivemos retorno de 4 usuários que são ou possuem familiares/amigos com baixa mobilidade e utilizam a interface de voz e sua correlação com automação residencial para tarefas do dia a dia; 3 respondentes informaram que são e/ou possuem amigos/familiares com deficiência visual e utilizam as assistentes como interface de voz em celular, computadores e smartspeakers; 2 respondentes informaram a utilização por idosos, em função da mobilidade e baixa visibilidade; 2 respondentes informaram a utilização da interface de voz por pessoas não-alfabetizadas, incluindo uma criança; 1 respondente informou a utilização da assistente em interface gráfica por um familiar surdo. Nas entrevistas em profundidade, um dos entrevistados utilizava as assistentes como tecnologia assistiva, devido a baixa mobilidade – esses resultados serão analisados a seguir.

O gráfico abaixo aponta as Assistentes Pessoais Digitais utilizadas pelos entrevistados.

Figura 34 Gráfico Assistentes utilizadas pelos entrevistados



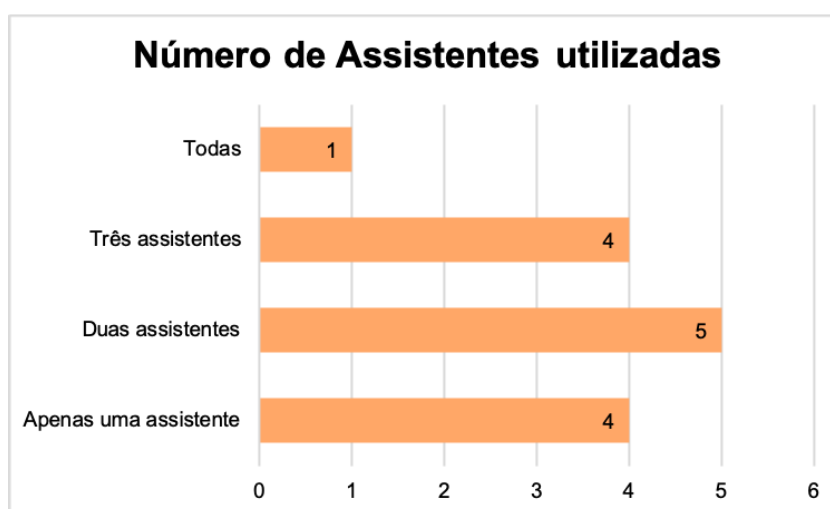
Fonte: a autora

As assistentes utilizadas pelos entrevistados apresentam uma ordenação diferente daquela apresentada pelo formulário quantitativo: a Siri aparece como a assistente mais utilizada, por onze dos quatorze entrevistados, seguida do Google Assistente, por nove entrevistados, da Alexa, cinco entrevistados e, por último, a Cortana, por dois entrevistados. Além da inversão entre Siri e Google Assistente como as mais utilizadas em relação aos dados quantitativos, Alexa e Cortana também invertem as posições, sendo a Cortana a menos

utilizada. A última alteração ocorre, parcialmente, em função de dois entrevistados que, apesar de brasileiros, moram fora do Brasil, onde a Alexa está disponível há mais tempo, e utilizam o mesmo. Ainda assim, na amostra de entrevistados que tivemos, não conseguimos encontrar nenhum participante que sua assistente principal e/ou de preferência fosse a Cortana – os dois que informaram utilizar, o fazem apenas esporadicamente e de forma secundária.

O gráfico apresentado na sequência distribui os entrevistados por quantidade de assistentes que utilizam, apontando que dos quatorze entrevistados, dez possuem experiência de uso com mais de uma assistente e apenas quatro utilizam uma só.

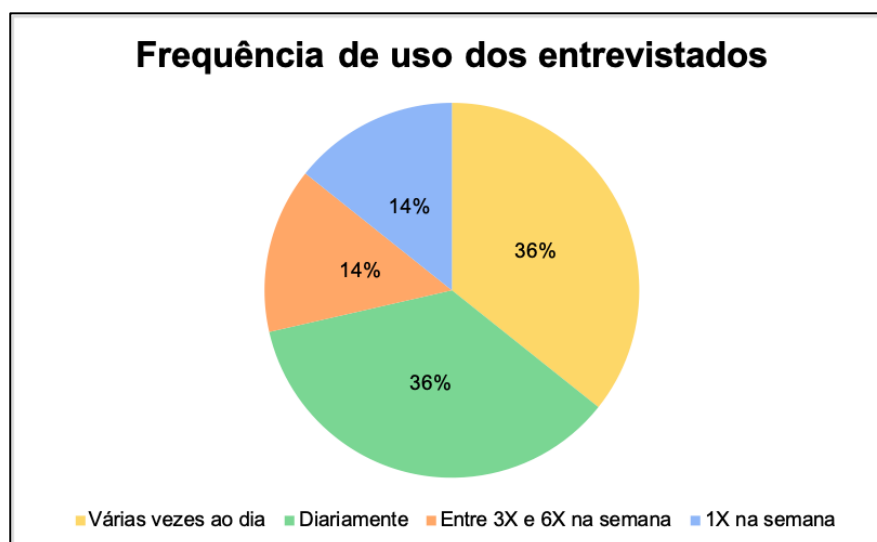
Figura 35 Gráfico Número de assistentes utilizadas



Fonte: a autora

A distribuição de experiência de uso com assistentes diversas fica da seguinte forma: cinco entrevistados utilizam/utilizaram duas assistentes, quatro entrevistados utilizam/utilizaram três assistentes e outros quatro utilizam apenas uma assistente, um entrevistado possui experiência com todas as assistentes. Todos os entrevistados que utilizam apenas uma assistente utilizam a Siri, o que pode apontar para uma tendência já conhecida de integração entre os dispositivos da Apple, em geral apreciada pelos seus usuários. O gráfico abaixo distribui os entrevistados por frequência de uso: lembrando que como pré-requisito para as entrevistas, o participante deveria ter intensidade ou tempo de uso.

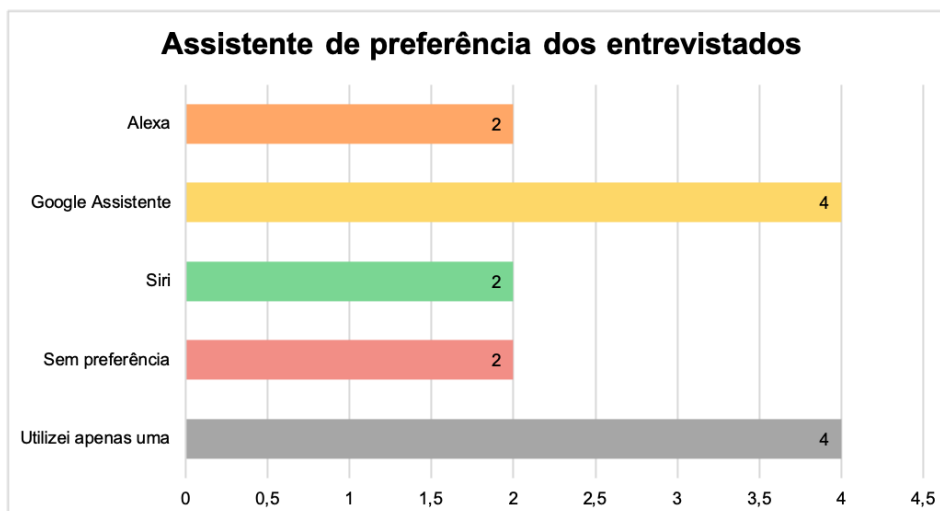
Figura 36 Gráfico Frequência de uso dos entrevistados



Fonte: a autora

Devido ao viés de seleção dos entrevistados, a frequência de utilização se mostra maior neste conjunto do que no dos dados quantitativos. Dos quatorze entrevistados, 36% informaram utilizar várias vezes ao dia e outros 36% diariamente, seguidos de 14% que utilizam entre três e seis vezes na semana e outros 14% que utilizam uma vez na semana. Interessante notar que, entre os entrevistados, tempo de utilização não necessariamente se reflete em aumento da intensidade de uso: dois dos entrevistados informaram utilizar a Siri desde o seu lançamento no iPhone, mesmo na versão em português, mas ainda assim sua frequência de uso é de apenas cerca de uma vez na semana. O gráfico seguinte apresenta as preferências dos entrevistados em relação as assistentes disponíveis no mercado.

Figura 37 Gráfico Assistente de preferência dos entrevistados



Fonte: a autora

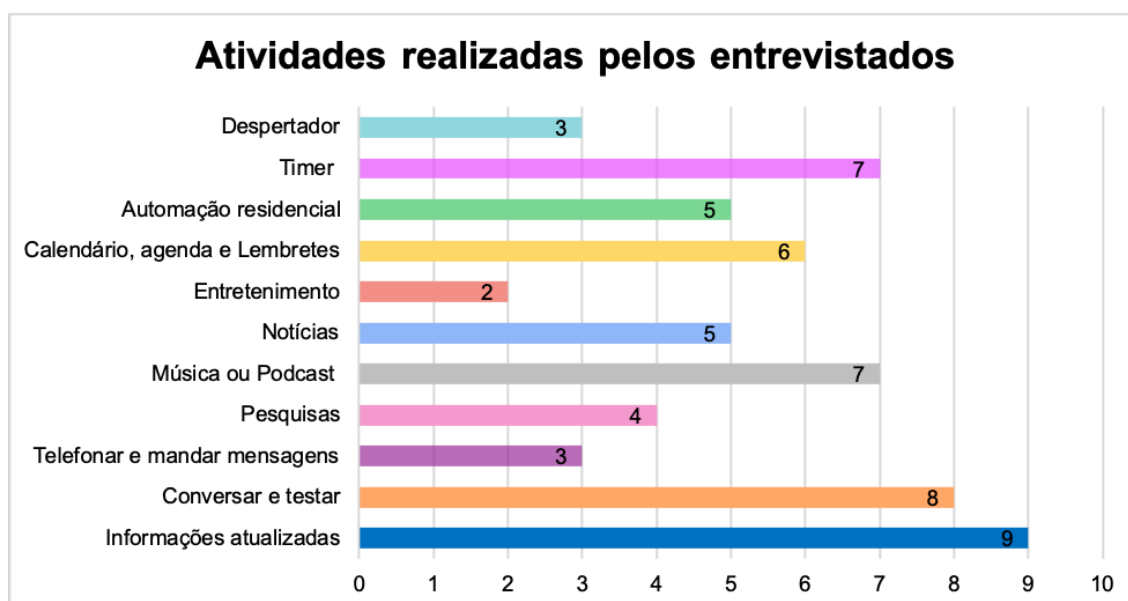
Dos quatorze entrevistados, quatro informaram preferir o Google Assistente, especificamente no modelo do Google Home. Entre os motivos citados para a preferência estão: a compreensão com maior clareza dos comandos utilizados, a integração com outros serviços e contas do Google e a compreensão e aprendizado da forma de falar do usuário. Quatro entrevistados afirmaram utilizar apenas uma assistente, dois não possuem preferência específica, dois preferem a Siri e outros dois preferem a Alexa.

#### 4.6 TODO DIA ELA FAZ TUDO SEMPRE IGUAL?

Durante as entrevistas em profundidade, os participantes falaram de suas formas de utilização, atividades para as quais utilizavam as assistentes e sobre a existência de rotinas de uso. Em assistentes como a *Alexa* e o *Google Home*, é possível criar rotinas específicas, que serão acionadas a partir de uma frase: por exemplo, é possível programar a *Alexa* para que, ao ouvir ‘bom dia’, inicie a leitura das notícias do dia, seguida da previsão do tempo e previsão do tempo de deslocamento do usuário até o trabalho. Para além das rotinas pré-programáveis, nossa ideia era entender se os entrevistados possuíam uma forma de uso que se repetia e se essa rotina criada com as assistentes já existia para o entrevistado antes da inclusão das assistentes no seu ecossistema midiático.

As principais atividades relatadas pelos entrevistados como as atividades que eles rotineiramente fazem utilizando as assistentes são: despertadores e timers; calendário e agenda de compromissos; previsão do tempo; buscas na internet e curiosidades; música e entretenimento de forma geral e automação residencial, principalmente para os usuários de smartspeakers. O gráfico abaixo apresenta a recorrência de atividades realizadas utilizando as assistentes entre os entrevistados.

Figura 38 Gráfico Atividades realizadas pelos entrevistados



Fonte: a autora

As atividades de maior recorrência entre os entrevistados são: Informações atualizadas, como temperatura e previsão do tempo, mapas e trânsito (9); Conversar e testar (8); Timer, principalmente na hora de cozinhar (7); Para ouvir música e podcasts (7); Calendário, agenda e lembretes (6); Automação residencial, com a assistente conectada a lâmpadas, ar condicionado, televisão, etc (5); Notícias (5); Pesquisas na internet (4); Telefonar e enviar mensagens (3); Despertador (3) e outras formas de Entretenimento, como jogos (2).

O entrevistado E2, que utiliza uma Alexa, nos contou que criou uma rotina em cima de um ritual matinal que o entrevistado já possuía, que incluía, por exemplo, ler notícias. Já o entrevistado E3, que também utiliza a assistente da Amazon, conta que ‘peguei o costume de todo dia de manhã perguntar como está o tempo [...] Aí eu já vou pensando que roupa botar em função disso, porque principalmente no inverno [da Áustria] 5 graus de diferença já quer dizer se eu preciso botar luva ou não, se eu preciso botar touca ou não. Dentro de casa não dá pra

sentir, para ter essa impressão... então essa rotina eu criei’. De acordo com E3 esse hábito foi criado apenas após o início da utilização da *Alexa*, e está conectado aos usos que ele começou a fazer da assistente, tornando-se algo incorporado ao seu cotidiano: ‘parece que eu sou um louco pra perguntar o tempo agora [risos]’.

A entrevistada E12, que utiliza a Google Home e a Alexa, relatou uma forma de utilização semelhante, em que uma rotina de utilização foi criada na interação com o dispositivo, que acabou se tornando uma rotina normal e só foi percebida dessa forma posteriormente. Ela conta que ‘acho que sim [criamos uma rotina de uso], mas não porque a gente quis. A rotina é assim, a gente coloca despertador e a previsão do tempo, todo dia. Timer, todo dia, porque a gente cozinha todo dia. Música todo dia. Então é meio que uma rotina assim... meio assustador assim, agora que eu tô pensando bem [risos]’. A percepção da existência de uma rotina com o dispositivo e do uso cotidiano e repetitivo inclusive gerou surpresa em E12, que não havia percebido como as tarefas do cotidiano estavam emaranhadas com algumas funções da assistente.

O entrevistado E8, que utiliza três assistentes pessoais digitais em suas versões para smartspeaker, o Google Home, o HomePod e a Echo, faz um processo de automação residencial mais intenso, conectando as assistentes as luzes, cortinas, ar condicionado, equipamento de home theater, entre outros. E8 conta que ‘então você diz, ‘eu quero olhar um filme’, aí eu já faço toda a rotina né, diminui a luz, baixa a cortina, liga o ar condicionado, liga o equipamento de som, põe no canal preferido, já põe lá o *Netflix* pra ver o filme, em um único e simples comando por voz. Não é nem um botão mais, já é um comando vocal’. E8 informou que utiliza diversas rotinas em seus smartspeakers, para momentos diferentes: rotina matinal, rotina noturna, rotina para ver filmes. Todas elas, entretanto, incorporam hábitos e formas de comportamento que o entrevistado já possuía anteriormente, tornando-as automatizadas.

E10 conta que, mesmo utilizando a assistente apenas no celular, também criou certas rotinas na sua utilização: ‘de noite eu programo o despertador do dia seguinte e peço a previsão do tempo. Isso é sempre. De manhã eu confiro a previsão do tempo, ligo a rádio, ela põe na rádio pra mim e eu ouço a rádio um pouquinho. As vezes eu peço pra ela ler alguma página. Eu uso de noite e de manhã. Ah, e as vezes na hora do almoço para marcar algum compromisso... que depois eu preciso revisar tudo porque ela marca tudo errado [risos]’.

Já para E13, a assistente veio integrar a sua rotina e gerar maior autonomia em atividades cotidianas, uma vez que o entrevistado possui baixa mobilidade tanto nas pernas e pés quanto nos braços e mãos. E13 conta que ‘normalmente eu programo uma rotina com o Google Home e mantenho essa rotina, com tarefas do dia a dia, lembretes. [...] meu uso de forma geral é a

integração mesmo, com outros dispositivos [...] eu pretendo aumentar, deixar a minha casa 100% automatizada’. A integração das assistentes com outros dispositivos da casa dá para E13 ‘uma independência maior. Hoje eu faço coisas que antigamente, há 10 anos atrás eu não conseguia fazer [...] Para mim mudar o canal fisicamente, por exemplo, não seria possível. Mas a assistente me dá essa autonomia, coisa que eu não conseguia fazer. Essa possibilidade de você controlar a sua casa por voz... o que me prende a assistente é de você não precisar fisicamente utilizar ela, como apertar botão na tela. Você faz tudo pela voz, pra mim isso é útil’.

A participante E14 conta que a utilização diária da Google Home ocorre principalmente em função da vinculação destas com as lâmpadas de casa, ‘ou seja, as minhas lâmpadas da sala dependem do Google Home’. Ela conta que essa questão nem sempre se apresenta como um facilitador, já que ‘as vezes eu confesso que me irrita, agora o meu marido está morando temporariamente fora do Brasil, e as vezes eu chego cansada e eu não tô afim de falar com ninguém, daí eu simplesmente ligo o interruptor. Ontem eu fiz isso e depois eu estava sentada e queria apagar a luz, ou diminuir o brilho da luz, e aí eu percebi que eu tinha ligado a luz com o interruptor e não ia conseguir ligar com a voz. Daí eu fiquei, ‘ah, eu deveria ter ligado com a voz porque aí eu consigo desligar sem precisar levantar da cadeira’. Mas eu uso ela num cotidiano’. Além do habitual ligar e apagar as luzes, E14 conta que criou uma rotina com as luzes para quando ela e o marido estão fora da cidade: ‘a gente tem uma rotina com as lâmpadas pra quando a gente viaja. Pra ela ligar e apagar as luzes enquanto a gente tá fora por uma questão de segurança [...] Ela liga num determinado horário e desliga em um determinado horário. Mais pra dar sensação de movimento’.

A dinâmica da luz com a Google Home, para E14, a automação dessa rotina também acaba trazendo alguns estranhamentos, principalmente quando falta luz em casa. ‘Uma coisa que acontece comigo, agora mais que eu tô sozinha, é que falta luz e daí eu chego em casa e a luz tá ligada [risos] esses dias aconteceu [risos] e eu: como assim tá ligada a luz? quem ligou a luz? [risos nervosos] Daí eu vi que estava tudo piscando, que faltou luz e que daí a luz ficou ligada eu não sei desde quando’.

#### 4.7 PERCEPÇÃO DA INTERFACE DE VOZ

Os entrevistados da etapa qualitativa foram perguntados sobre sua percepção geral da interface de voz e para compararem com a utilização que fazem da interface gráfica. Alguns

entrevistados apontaram que não haviam inicialmente pensado na assistente pessoal digital como uma interface de voz, enquanto para outros essa associação aparece logo na descrição que eles apresentam do que é uma assistente (essa descrição é discutida no capítulo seguinte). A entrevistada E1, diz que ‘não tinha me ocorrido que era uma interface, mas agora que tu falou sim [percebi que é]. Gosto muito das duas interfaces [para utilização no celular], a de voz e a gráfica. O uso varia conforme o momento. São interfaces complementares, mas ainda uso muito mais a gráfica’.

Para E2, que foi beta tester da Alexa antes do lançamento no mercado, aponta que a relação que ele estabeleceu com a interface se modificou o tempo. No início, E2 ‘estava sempre preocupado com o funcionamento, mas com o tempo eu fui me acostumando com ela’, essa interface foi se naturalizando e ficando menos evidente. Isso alterou inclusive a forma de relação que E2 estabelece com a Alexa – quando deixou de perceber tão intensamente a interface e se preocupar tanto com seu funcionamento, passou a agir de forma mais espontânea e, as vezes, humanizada: ‘aí eu digo, Alexa deixa de ser burra, Alexa deixa de ser surda, você não consegue me entender. Eu cheguei a esse ponto de gritar com ela dentro de casa, os vizinhos acham que eu sou louco, porque eles sabem que eu moro sozinho e eu de repente eu estou gritando aqui. Mas no começo eu sabia como funcionava eu queria fazer um beta teste bem perfeito, mas eu me peguei nessa situação’. A experiência de E2, de percepção da forma de funcionamento do sistema e da interface é interessante também por sua formação ser na área do design gráfico, que lida diretamente com a questão. A percepção em termos de usabilidade de E2 é de que ‘acho prático ser por voz, você não precisa estar olhando e segurando algo. Acho que atinge um público maior, como um público mais idoso. Eu só sinto falta quando eu quero ver algo [como um texto, um livro, uma imagem]’.

E3 aponta que acredita que as assistentes pessoais, em termos de inteligência artificial, ‘tem bastante para melhorar’. Ele percebe que a gama de ações que ele realiza com a assistente ainda é restrita, mas ao mesmo tempo acredita que a adesão a essas ações através dessa interface em específico por um longo período de tempo tem seu valor: ‘acho que realmente pode virar uma coisa maior do que só perguntar o tempo, ligar a luz e cuidar o forno [suas principais ações]. Mas, pensando que se eu continuo usando pelo menos essas duas funções por muito tempo, já têm algum valor, se vale o preço daí já é outra discussão. Pra mim eu acho que vale. Tem algum valor porque facilitou alguma coisa na minha vida que eu tô preferindo fazer desse jeito, genuinamente, não só porque eu tenho e tenho que usar. acho que já agregou um valor mínimo na minha rotina pelo menos e tem mais para agregar’. O entrevistado E3 percebe que a sua utilização da interface gráfica para a de voz depende da atividade em questão, não sendo



o caso de uma substituição completa dessas funcionalidades. E3 aponta que a escolha de uma ou outra interface ‘depende pra qual tarefa. Eu acho que justamente eu uso ela pras coisas que só falar é mais conveniente que só clicar. Porque o meu celular faz as duas coisas, ele tem previsão do tempo e tem cronômetro. Mas quando eu tô com a forma na mão [cozinhando], é muito mais conveniente só falar, sabe? Depois eu posso esquecer, ou preciso parar o que tô fazendo’.

Além da alternância de uso conforme o contexto, E3 percebe que a opção de uma matriz diferente de interface é interessante, pois ‘eu acho que opções nunca são demais. Principalmente para acessibilidade. Não tem como operar um computador só com voz... pra mim né. Para um cego talvez tenha’. A avaliação, apesar das limitações do artefato, é positiva, pois E3 diz que ‘eu valorizo muito uma rotina. Eu acho que ela adiciona coisas a minha rotina, mesmo que seja o mínimo. Eu acho que uma vida boa se faz de muitas mínimas adições. É uma experiência melhor’. Nesse sentido, a percepção das assistentes pessoais enquanto interface é avaliada como uma boa experiência e um investimento que compensa para E3, mesmo que as funcionalidades não sejam necessariamente inovadoras ou que tenha problemas que precisam ser sanados.

A percepção e utilização da interface de voz por E7 é similar a de E3. Para E7 utilizar a voz ao invés de uma tela, se relaciona com outros tipos de mídia, como ‘o uso da voz é meio comparar com a televisão e o rádio, o uso da voz eu não preciso parar o que eu estou fazendo aqui, eu posso fazer a pergunta e continuar fazendo o que eu tô fazendo. Se eu tivesse que fazer uma pesquisa digitando eu ia ter que parar pra digitar, pra ler, pra procurar o resultado. Nesse momento que eu estou cozinhando é um dos principais. que aí eu só aperto o botão e peço pra ela ligar o timer ali. É uma coisa que é uma ação que é rápida e que se eu fosse fazer de outro jeito eu teria que parar o que eu quero fazer’. O que aparece como fator importante para E3 e E7 é poder manter a atenção em uma tarefa que está sendo realizada enquanto executa um outro comando em segundo plano – ou seja, não é preciso interromper uma atividade para realizar uma pequena tarefa, como colocar um timer. As assistentes pessoais digitais enquanto interface de voz aparecem como dispositivos que são auxiliares nesse sentido, estando conectados a uma outra ação, que pode ser cozinhar, por exemplo.

E4 aponta que a interface de voz consegue diminuir o tempo gasto para realizar tarefas cotidianas, como marcar compromissos no calendário. Uma vez que E4 é professora particular, percebe que ‘facilita muito muito a vida. Corta muitas etapas da minha vida. Por exemplo, se num dia eu preciso marcar cinco alunos diferentes, são cinco vezes que eu não preciso pegar o celular, abrir minha agenda, ver horários, etc. Pra mim é uma forma de otimizar meu tempo e meu cérebro para fazer coisas por mim’. A agenda é a principal funcionalidade acessada através

da assistente, o que aponta, assim como E3, para o fato de que não é preciso muitas formas diferentes de utilização da assistente para que ela tenha valor de uso para os entrevistados.

E5, que utiliza a Google Home e também a Alexa, entende que o uso constante também é o que permite um aprimoramento dessa interface. Entretanto, as assistentes pessoais vêm com poucas instruções de utilização, exceto por algumas sugestões do próprio sistema de coisas que é possível solicitar. Descobrir, de forma mais ampla, o potencial do dispositivo, a forma de utilização, quais são os comandos corretos, passa por um aprendizado que é baseado no uso. Para E5, ‘quanto mais eu acionar o Google, mais ela vai aprender. Mas assim, como que as pessoas mais velhas estão usando, sabe? [...] a gente está muito acostumado com manual né, e não tem. Por isso que tem muito grupo no facebook, whatsapp [sobre a utilização das assistentes]. Teve uma noite que eu fiquei umas 3 horas vendo o que a Alexa fazia... mas nem todo mundo dispõe desse tempo para testar’. Não saber como utilizar se apresenta, portanto, como uma barreira de uso: o que aparece aqui é que existe um letramento necessário para a utilização de uma interface específica. A mesma questão é verdadeira com a interface gráfica, porém, tendemos a não perceber essa dificuldade justamente por estarmos expostos a interfaces gráficas há um período mais longo de tempo. Além disso, existe a ideia de que a utilização da voz seria mais “natural” para os seres humanos e que, portanto, a utilização seria mais intuitiva. As assistentes pessoais, entretanto, funcionam com comandos bastante específicos e alterações em uma palavra, na ordem da frase ou na sua formulação são suficientes para que um comando não seja compreendido e executado pela máquina.

Ao comparar a interface gráfica com a interface de voz E5 pontua que ‘acho que a voz tem algo natural. É mais amigável do que usar um teclado, um mouse, os dedos. Eu acho que essa a última geração que vai usar um computador assim [...] acho que caminha para ficar cada vez mais natural. Pra mim facilitou muito pedir coisas por voz [...] tu aprende a usar o mouse, tu aprende a usar o teclado...’. Ainda que a utilização da fala e da voz já esteja presente em nossas interações entre humanos (que não é o caso do mouse ou do teclado, por exemplo), existe sim um período de adaptação e de compreensão de como falar com esse dispositivo em específico: no capítulo sobre voz e fala analisamos as formas como as pessoas percebem que falam com esses dispositivos, incluindo questões como modulação da voz, rapidez e construção de frases, por exemplo.

E5 aponta ainda que existem diferenças na forma como pensamos com as nossas mídias, assim como com o tipo de utilização que estamos acostumados a fazer: ‘acho que uma questão que pra mim ainda é muito relevante usar a interface gráfica [é a escrita]. E isso é como o computador me moldou enquanto jornalista. Eu aprendi a pensar escrevendo no teclado... então

agora é muito difícil de voltar a pensar falando. [...] no teclado eu estruturo a linha de raciocínio, a frase é completamente diferente. [...] é um outro processo, eu estou voltando sabe'. A principal substituição da interface gráfica pela interface de voz foi percebida na utilização do celular, 'eu vejo que substituí muito o uso do celular. Abrir o celular para ouvir música, para ver o tempo, etc. Isso eu não faço mais'. O entrevistado E8 pontua questão similar sobre a substituição de uma interface gráfica pela de voz: 'hoje em dia quem sabe usar praticamente não põe mais a mão no telefone né'.

Para a entrevistada E9, a avaliação da interface de voz está extremamente relacionada com aquilo que ela não consegue fazer e gostaria de poder. E9 conta que 'se eu perguntar pra Siri 'definição de porta' [ela vai me dizer] 'olhe o que eu achei sobre definição de porta na internet', e daí ela me trás o [sistema de busca] Bing pra informação. Então pra isso ela não me serve, ela poderia me servir se ela tivesse uma função interligada com o Google. Mas eu não uso o 'Ok, Google', o Google Assistente, porque eu nem sei se dá pra instalar no iPhone ou não e meu celular de trabalho eu só uso o mínimo possível'. A utilização da Siri acaba sendo muito específica 'porque ela não se presta pra outras coisas, pra mim pelo menos, se eu quero mandar um Whastapp pra uma pessoa ela não abre o Whatsapp. Ela abre mensagem, mas ninguém mais usa *sms* sabe'. Entretanto, E9 acredita que sua utilização do sistema seria maior se 'a assistência dela fosse realmente uma assistência pessoal e não uma assistência da Apple'. Nesse sentido, E9 entende que a assistente pessoal 'é um atalho na verdade, as coisas que já estão ali de alguma maneira. [...] eu uso muito mais a interface gráfica, muito mais. [...] Eu acho que eu me sinto meio idiota perguntando coisas pra Siri, eu me sinto meio idiota. É isso. E também por ela não funcionar de outras maneiras que eu achei que ela iria funcionar. Mas o principal é isso, eu me sinto meio estúpida falando com um pedaço de silício'. A barreira de utilização da assistente de forma mais intensiva, para E9, ocorre, portanto, tanto em função das funcionalidades que ela considera restritas nesse sistema, quanto pela interação por voz, que é percebida com certo estranhamento por parte de E9, o que inibe a sua utilização. A utilização da fala e da voz com um artefato não aparece para E9 como natural, mas, ao contrário, como algo não-natural, algo estranho, que ela se sente 'idiota' por fazer.

Para E10, a utilização da interface de voz é considerada positiva, ainda que limitada, podendo ser descrita como 'um controle remoto mais elaborado, nada além disso'. A entrevistada também aponta que gostaria que as assistentes de fato oferecessem um serviço mais amplo de assistência, como 'por exemplo, da Cortana eu queria que ela organizasse melhor os arquivos, que ela me explicasse melhor as questões de armazenamento, que ela me desse alerta quando ela tivesse coisa que pode deletar. Eu queria que ela fosse mais proativa no sentido

de organização. A siri ela precisa aprender a, melhorar a forma de marcar os compromissos, registrando melhor'. Mesmo com as limitações, E10 entende que possui uma preferência grande pela utilização da voz na interação no lugar das telas, 'eu realmente prefiro a interface de voz à gráfica, disparado. A gente já está tão cansado dos olhos né'. A tendência de E10, com a melhora dos serviços, seria migrar cada vez mais para a utilização da interface de voz e deixar a interface gráfica de lado: 'eu faria a troca cada vez maior com a melhora dos serviços. Mas eu sei que a contrapartida é que eles estariam sempre me escutando. E teria que ter mais certeza que ela só seria ativada quando eu chamasse e depois ela desativaria [...] que nem a gente tem modo avião assim. Eu não acredito muito que elas não ouvem quando não tá aparecendo'. A entrevista pontua outra barreira na utilização da interface de voz, que são as preocupações com a privacidade e com a forma como as empresas estão lidando com os dados que são coletados.

Para E12, 'minha percepção é que elas são excelentes e elas estão entrando na nossa vida pra ficar. A gente não consegue mais não ter. Um dia meu marido estava cozinhando e a gente estava sem internet, e ele achou 'o ó' ter que colocar o timer manual no fogão [risos] E eu falei: nossa, *first world problems* [muitos risos]. Ele falou: 'ah, demorou, eu fiquei perdido'... é que a gente acostuma'. Na casa de E12, ela e o marido utilizam dois Google Homes e duas Alexas, em cômodos diferentes – a entrevistada percebe que gradativamente algumas ações ficam diretamente ligadas a utilização das assistentes, como o timer para cozinhar. E12 conta que a utilização da voz no lugar de uma tela é fator positivo para ela, já que 'eu sinto uma familiaridade, porque eu prefiro usar a minha voz do que ler e do que escrever. Pra qualquer coisa. Mesmo quando eu estava fazendo pesquisa. Pra qualquer coisa... se eu quero fazer um diário ou se eu quero um diário da minha filha, eu escrevo, mas eu prefiro colocar a voz e a voz escreve. Porque você consegue pensar, entendeu. Sabe aquela parte física, de você ficar só ocupado com a criação, quando eu falo eu acho mais fácil, me deixa mais livre'.

A entrevistada sente que fez uma troca entre interface gráfica pela de voz que é considerável e que inclusive sente que 'eu deveria usar mais [a interface gráfica]. Eu deveria ler mais. "Tipo news, eu deveria ler, mas eu acho que é também por causa da nenê também, porque eu estou ocupada, mas assim eu escuto. Eu prefiro ler, porque você precisa, você começa a esquecer... eu começo em português a escrever coisa errada... e eu fiz letras [como curso de graduação], meu! Eu fiz letras!'. Nesse caso, o que E12 aponta é que, além de uma preferência por utilizar a voz no lugar da escrita/leitura, o momento de vida dela, com uma filha que ainda não completou um ano, é fator relevante na sua crescente utilização da interface de voz, por estar frequentemente com as mãos e braços ocupados em tarefas que envolvem cuidar da filha

e da casa. Ela entende que ‘no momento a voz está perfeita. Tem a ver com a fase da minha vida’.

Para E13, a utilização de uma interface de voz nas atividades do cotidiano não representa uma troca que gera comodidade, economia de tempo ou facilidade, mas sim autonomia. O entrevistado possui mobilidade limitada por deficiência física e conta que a utilização das assistentes ‘me dá uma independência maior para o meu uso. Eu não uso por comodidade, é por necessidade. Então eu utilizo diariamente o [Google] assistente para fazer essas automações da casa. Já a Alexa eu utilizo em conjunto com a Google, mas eu uso em outro cômodo’. Em função disso, a utilização de E13 é diferente dos demais entrevistados, de forma muito mais intensa e conectada a eletrodomésticos da casa, como as luzes e o ar condicionado. Essa conexão evita que o entrevistado precise se locomover para executar determinadas ações em um cômodo, o que lhe concede autonomia, pois essas ações não eram possíveis sem auxílio de terceiros. E13 conta que ‘no celular eu utilizo o Google Assistente mesmo. Só que eu utilizo para escrever textos mesmo. No meu caso, eu tenho mobilidade reduzida nas mãos e em outras partes do corpo. No caso eu utilizo para escrever textos quando eu preciso, nos aplicativos né, o whatsapp, o facebook. [...] no computador eu utilizo o teclado virtual [para escrever]. Para poder controlar o celular eu utilizo o computador, aí eu controlo o celular pelo computador e escrevo utilizando o assistente’.

A conexão entre diferentes hardwares, como o celular e o computador, com as assistentes, possibilita que o entrevistado utilize ferramentas de comunicação e escreva utilizando a voz, sem a necessidade de uma mobilidade muito grande nas mãos. E13 compreende que as assistentes pessoas com interface de voz são ‘uma forma de você automatizar o seu cotidiano, que você costuma utilizar no seu dia a dia. E também é uma forma de interação mais independente de você estar realizando essa tarefa. É uma forma de facilitar a vida do usuário no dia a dia e também de quem tem alguma necessidade especial, que é o meu caso né’. E13 acredita ainda que desenvolvimentos como esse irão aumentar e melhorar no futuro, mas que a utilização por parte de pessoas com deficiências físicas ainda é baixa: ‘a percepção que eu tenho hoje é de que o avanço da tecnologia ajuda muito [pessoas com deficiências físicas] eu acredito que futuramente isso vai ser expandido para outras coisas também [...] das pessoas com deficiência utilizarem para ter mais autonomia. Porque é raro você ver alguém utilizando para este fim. Eu acho que falta um pouco de informação ainda na mídia’.

Para E14, a escolha por uma ou outra assistente também passa pela escolha de outros serviços da empresa, ‘a decisão de comprar a Google Home foi uma decisão pelo ecossistema

do Google estar mais difundido aqui, apesar de que Alexa, na minha visão, é a melhor pra gente manter uma conversa [...] O Google Home é muito bom em executar coisas, essa é a minha percepção'. A integração com outros produtos do próprio Google foi um fator relevante na hora de escolher qual smartspeaker comprar. Entretanto, E14 avalia que a interface de voz nem sempre é mais prática do que uma interface gráfica, ou mesmo do que um simples botão no lugar de um processo de automação. Para abaixar o volume do smartspeaker, por exemplo, avalia que não é 'assim tão prática, tu tem que falar 'Ok, Google', daí ela vai parar a música e daí abaixar o volume... se tu tiver perto dela, é muito mais fácil apertar o botão, se tiver com o teu celular é muito mais fácil apertar o botão, que nem se tu tiver do lado do interruptor é muito mais fácil apertar o botão [do que fazer um comando de voz para apagar ou acender as luzes da casa]'. Ainda assim, E14 avalia que o potencial de uma interface de voz é considerável e que pode passar a ideia de uma interação que não possui uma interface: 'eu acho que eles têm um potencial muito legal. Acho que talvez eles vão ser uma tendência de uma nova forma de interação com a tecnologia, de uma interação sem interface... que eu acho que é isso faz deles algo muito ímpar no contexto tecnológico que a gente tem hoje... da gente ter com eles uma interação sem interface, quando a gente está acostumado a colocar tela em tudo. Mesmo que agora a Alexa tenha aquela versão com tela. Mas a ideia de interação com esses aparelhos ela é fora da tela dela né, ela é mais por voz e a gente tá acostumado a trabalhar com gestos, toque, sensível ao toque. Então eu acho que isso é uma tendência'.

Ainda que perceba a utilização da voz como uma tendência em termos de comunicação, em contrapartida a uma proliferação do uso das telas, E14 também acredita que as formas de utilização e funções que as assistentes têm ou podem ter ainda não foram plenamente descobertas. E14 coloca que 'eu acho que as potencialidades dela não foram ainda talvez explicitadas ou apropriadas pelas pessoas. Eu digo que mesmo o meu uso é muito experimental, ver o que faz sentido. Que é aquela ideia, que quando surge uma tecnologia a gente não sabe pra que ela funciona, mas a gente acaba descobrindo ou criando necessidade né. Porque a gente não tinha ela antes e a gente acaba vendo pra que a gente precisa dela e o que a gente sente falta. Então eu acho que ela tem potencial mas ela não é algo que tu '*must have*', não é um item que eu preciso, de primeira necessidade [...] mas talvez se eu não trabalhasse com tecnologia eu até já teria vendido ela [a Google Home]. Eu digo isso porque as vezes vários *devices* a gente acaba que se adapta ou não né'. A entrevista explica um pouco essa relação com uma tecnologia nova a partir de um exemplo dela mesma, de utilização de pulseira que funciona como rastreador de parâmetros de saúde: 'hoje eu uso essa pulseirinha de tracker de saúde, eu comecei com outra, há muito tempo atrás, eu usei por dois ou três meses, eu dizia que eu me senti como

uma presidiária que estava com uma tornozeleira, mas era uma pulseira. E eu vendi ela logo em seguida. E daí passou anos depois, eu experimentei outra marca, e hoje eu uso a fitbit [...] mas hoje eu gosto muito dela, eu tenho a balança, eu tenho todo o ecossistema da fitbit na minha casa. Eu não me adaptei, passou um tempo, e eu vi uma utilidade pra ela, talvez até porque ela se desenvolveu um pouco mais. E quando eu vejo a questão das assistentes pessoais eu acho um pouco isso’.

Para E14, um dos pontos que faz com as assistentes não sejam um artefato muito utilizado é que ‘em alguns pontos elas precisam se desenvolver, em outros pontos mesmo as funcionalidades delas de serviço, elas ainda são muito uma transposição. Se tu pedir *‘tell me the news’* é um rádio. Se for na Alexa com tela é uma TV. Não me dá uma nova experiência. Ela me amplifica ou ela me facilita uma experiência que já existe, tanto de ligar a luz quanto de perguntar alguma coisa, ela vai me dar um Wikipédia, ela vai me dar o Google. Elas amplificam coisas que já existem’. Nesse sentido, E14 entende que elas funcionam exatamente como uma interface de voz, onde é possível realizar comando específicos utilizando a fala ao invés de escrever ou de clicar. O que fica faltando é o fator ‘assistência’ em termos mais amplos, como comentado por E9 e E10. Por estar conectada com outras contas e dispositivos, com acesso a dados que possibilitariam uma personalização em termos hipotéticos, E14 pondera que ‘ela deveria poder atender mais as minhas particularidades, as coisas que eu busco e antecipar o meu desejo, ou seja, ser uma assistente pessoal e não só um comando de voz [...] hoje ela não antecipa minhas necessidades, então ela não é a minha assistente pessoal, ela é um comando de voz’.

#### 4.8 PROBLEMAS NA INTERAÇÃO COM AS ASSISTENTES

Ainda que a maioria deles tenha afirmado que existem diferentes aspectos nos quais essa tecnologia ainda precisa evoluir, de forma geral as assistentes foram avaliadas positivamente pelos entrevistados na etapa qualitativa. Todos os entrevistados da etapa qualitativa relataram algum problema na utilização das assistentes, sendo o mais frequente destes a não compreensão adequada da fala por parte das assistentes.

E1 relata que, entre os problemas que percebe, estão ‘eu pedir algo e ela não me entender’ e também que ‘as vezes eu não sei como perguntar’, para obter a resposta desejada. Ou seja, não sabe a forma específica daquele comando ou a formulação adequada para encontrar

algo no mecanismo de busca. E1 ainda relata que, quando utiliza a assistente como um mecanismo de busca na internet, nota que tem mais dificuldade em encontrar os elementos adequados para a chave de busca, uma dificuldade que não possui quando lidando com ferramentas de busca na forma escrita. Para E4, além da dificuldade de se fazer entender, acontece também de querer ‘utilizar uma funcionalidade que ela não tem’.

Formular adequadamente um comando é o que E2 percebe como o principal problema: ‘você saber exatamente qual o comando para ativar alguma coisa. Se você formular uma frase um pouco diferente, muda uma palavra, ela já não entende. O machine learning dela não percebe sinônimos ainda. A sensibilidade do microfone as vezes não é tão boa também’. Além disso, E2 aponta que a conversa ainda é pouco fluída e que precisa de ‘mais naturalidade e dar continuidade a conversa’.

E5 também percebe que nem sempre sabe como acionar o comando que deseja, o que se torna mais evidente uma vez que as assistentes não vem com um manual de instruções contendo os comandos existentes e as formas verbais adequadas para acioná-los. Ela conta que existe um problema ‘de entendimento. As vezes elas não me entendem ou eu que não sei falar de acordo’ e que acontece de ‘eu não saber acionar o comando que era necessário’. Na utilização da Alexa, por exemplo, ela conta que ‘eu já pedi coisas que eu sei que ela faz e ela não acionou. [...] as vezes pode se tornar meio frustrante. Quando ela não te responde direito, enfim, o que é responder direito. Quando ela não responde satisfatoriamente, quando ela não te dá o retorno que tu imaginava’. As respostas equivocadas e a não compreensão são fatores que geram frustração nos usuários, o sentimento que apareceu como predominante durante a pesquisa quantitativa. Para E5, ‘quando ela não te entende então, ai, aquela frase da Google que é ‘eu não posso te ajudar com isso’ [aff!], me irrita muito! [...] isso frustra e irrita. Eu tento analisar friamente as minhas reações, eu fico me autoanalisando, o que estou fazendo, mas eu não consigo ser assim ‘ela não tem que me ouvir’. Ela tem que sim, e isso é um problema, me causa uma frustração’.

E8 aponta também o problema da compreensão, mas relacionada ao poder de captação de som das smartspeakers, ‘as vezes me irrita pelo fato de que eu tô muito longe, aí só a Apple ganha de novo nisso [...] quando tu tá muito longe as vezes elas não te escutam. E a minha mulher tem a mania de chamar elas de surda, né: ‘vocês estão surdas, né’ [risos].’. E10 é mais dura nas críticas em relação ao reconhecimento de linguagem natural falada das assistentes, ao dizer que ‘elas não entendem tudo, é muito limitado’ e que possuem ‘dificuldade de entender sotaque’ e que ‘nem palavras mais básicas as vezes entendem’. Para E10, apenas demandas



muito simples conseguem ser atendidas: ‘pra perguntar simples é satisfatório, mas se você estiver uma demanda um pouquinho mais complicada, não entende’.

E11 aponta problemas na precisão da geolocalização: ‘eu costumo pedir a previsão do tempo... e eu moro em Triunfo. E geralmente me dá o tempo de Arroio dos Ratos’. Para o entrevistado, os problemas de compreensão já estão sanados e não percebe estas questões na utilização da Google Assistente, entretanto conta que ‘na Siri eu notava que ela não entendia tão bem naquela época quanto hoje. Meu primo tem a Siri e ele só vai falando, como se estivesse falando com a parede assim e ela vai executando as ações. E quando eu tinha não era assim, eu falava, ‘procura uma receita com abacate’ e ela não procurava. Parecia a velha da praça assim [referência a uma personagem da *Praça é nossa*, que era surda]’.

Para E12, os problemas de compreensão entre as assistentes e os humanos não se dão pela não compreensão das palavras, mas principalmente pela necessidade de uma formulação muito direta, com todos os elementos e sem nenhuma necessidade de inferência sobre o conteúdo. Ela entende que ‘principalmente quando você pede música... não é só com música. Tem coisas que são muito especificamente humanas. Por exemplo, sabe quando você fala uma coisa e a outra pessoa fala: ah, sei! E puxa aquela informação? Eles não têm isso. Eu não sei explicar, não consigo colocar em palavras exatas. O que precisar de um contexto ou da mínima vivência sobre o assunto, eles não conseguem. Tem que dar assim muito específico... tudo que é assim meio instintivo, intuitivo, eles não conseguem’. O problema de interação com a interface que E12 pontua se relaciona com a forma de falar que os seres humanos utilizam entre si: com a convivência e partilha de conceitos em comum, não existe a necessidade de falar cada ponto de uma frase<sup>158</sup>.

Acontece que, com a necessidade de explicitar todos os detalhes para realizar um comando adequadamente, algumas solicitações acabam se tornando muito longas e nem sempre serão compreendidas pelas assistentes. E14, durante a entrevista, conta um caso de não-comunicação com a Siri: ‘eu tenho familiaridade com a Siri, mas eu acho ela bem burrinha. Ela não consegue entender muita coisa. As vezes eu digo: ‘bota um lembrete na minha agenda para me lembrar do’ e daí ela já não entende mais, ‘lembrar o que?’. As vezes eu peço ‘lembrar de mandar um e-mail’ e ela ‘mandar e-mail para quem?’, e eu: ‘não, é lembrar de mandar um e-mail’. Ela não consegue, quando ela ouve a palavra e-mail ela já te põe pra mandar um e-mail’.

---

<sup>158</sup> Esse ponto será tratado detalhadamente no capítulo seis.

#### 4.9 FORMAS DE INTERAÇÃO COM AS ASSISTENTES

Os entrevistados da etapa qualitativa da pesquisa foram perguntados sobre questões relacionadas aos aspectos interacionais com as assistentes pessoais. Ao longo das próximas seções, traremos as análises referentes aos seguintes pontos: como os entrevistados descreveram suas interações com as assistentes; conversas e brincadeiras como forma de interação e aspectos de sociabilidade relacionados com as interações com as assistentes.

Solicitamos a alguns entrevistados que descrevessem a forma como eles percebiam suas próprias interações com sua assistente. A maioria das descrições aponta para usos rápidos e pontuais da assistente enquanto interface. Para E3, por exemplo, a interação é ‘pontual, uso quando preciso’, enquanto E4 descreve que são ‘práticas e rápidas, não tenho uma interação calorosa’ e E7 aponta que é ‘bem pontual [...] e breve também, é uma pergunta para te ruma resposta’. Para E12, a forma de descrever as interações com as assistentes é parecida com relação entre empregador e empregado: ‘elas trabalham pra mim. Foi meio ruim, desculpa [risos]. Mas essa é a sensação [da interação]’. Para E13, além de serem funcionais e cotidianas, já que o entrevistado utiliza a automação residencial através das assistentes em função de deficiência física que compromete sua mobilidade, mesmo com falhas de funcionamento elas cumprem seu papel, já que ‘a gente acaba adquirindo esses dispositivos já sabendo das limitações que eles têm’.

As descrições de E9, E10 e E11, ainda que sempre permeadas por aspectos de praticidade e funcionalidade, apontam outros elementos como a presença de brincadeira ou de perguntas que não possuem uma intenção técnica, mas na qual buscam compreender o artefato ou a sua forma de design. E9, por exemplo, diz que as interações são ‘utilitárias e, se ela fosse um ser humano, seria bullying [risos]’; E10 pontua que são ‘práticas e algumas engraçadas’; E11 explica que suas interações são permeadas pela ‘questão do tensionamento, onde eu ajo de maneiras desconfortáveis’, com o intuito de testar a neutralidade do design e as respostas oferecidas, mas que também são ‘técnicas’, por exemplo, na utilização da agenda: ‘a minha interação com ela é tipo pedir pra alguém te lembrar de não esquecer de algo, e daí porque tu falou tu já não esquece’.

Essas descrições nos permitem começar a vislumbrar uma dicotomia na interação com as assistentes, que é recorrente entre uma parte considerável dos entrevistados, assim como dos resultados obtidos na etapa quantitativa. Por um lado, a interação com as assistentes se dá como uma forma de interface de acesso/mediação para diversas outras mídias, serviços ou funções:

pode ser um acesso a funcionalidades do computador, tablete ou celular, mas também de funções da casa, através da automação e integração com outros dispositivos inteligentes, com serviços de música, de informação e de entretenimento, por exemplo. Nesse caso, acessamos por voz a uma série de serviços aos quais já possuíamos acesso, mas com a possibilidade de minimizar os passos envolvidos, evitar locomoção ou apenas deixando de lado a necessidade de utilizar uma interface gráfica de acesso. Por outro lado, as interações com as assistentes ocorrem não apenas através delas, mas também com elas, na medida em que se estabelecem interações cuja única finalidade é compreender, acessar, testar ou se divertir com as próprias assistentes. Elas são em si mesmas objetos com os quais interagimos, através de diálogos dos mais variados – esta segunda forma de interação será analisada na sequência, a partir das entrevistas em profundidade.

#### 4.9.1 JOGANDO COM AS ASSISTENTES: ENTRE A BRINCADEIRA E O TESTE

Assim como para os respondentes do formulário, que informaram utilizar as assistentes para diversão (166 respondentes), para se distrair (160 respondentes) e para conversar (56 respondentes), o humor e a conversa também aparecem como fatores relevantes para os entrevistados em profundidade. A categoria emergente do formulário quantitativo, que aponta o fato de que ‘é divertido’ utilizar a assistente como um motivador da utilização dessa interface, parece dizer sobre tanto uma forma interacional entre usuários e assistentes quanto uma forma de específica de desenvolvimento das assistentes, pensadas também para entreter (e não apenas para fornecer uma mediação para o entretenimento). As formas interacionais com as assistentes, utilizadas enquanto um jogo, serão discutidas a seguir; já a forma de construção desses personagens será discutida no capítulo seguinte.

Os entrevistados apontaram que jogam com as assistentes de duas formas diferentes: eles brincam com elas, pois é engraçado e curioso; e eles jogam contra o sistema, buscando descobrir erros, testando a humanidade/maquinidade das assistentes.

E1 conta que ‘as vezes eu tô afim de rir e eu sei que ela dá umas respostas engraçadas, então eu faço umas perguntas... por exemplo quanto é zero dividido por zero. Se quiser até te mostro a resposta! [risos]’. Eu digo que sim, ao que E1 pergunta para a Siri e recebemos a resposta [da própria Siri]: “Imagine que você tem zero bolachas e divide entre zero amigos. Quantas bolachas cada amigo vai receber? Está vendo, isto não faz sentido. O come-come fica

triste porque não há bolachas e você fica triste porque não tem amigos” [nós duas rimos]. E1 continua e me explica que ‘são coisas toscas que são muito engraçadas’, como fazer um beatbox e contar piadas. As interações como conversas com a Siri, fazendo perguntas que a entrevistada entende como ‘bobas’ e ‘engraçadas’ chegaram a durar até meia hora seguida. E1 conta que ‘eu estava com o meu sobrinho e ficamos conversando fazendo várias perguntas e conversando’. As interações do sobrinho, que ainda é criança, contemplam perguntas variadas, por exemplo ele ‘pergunta se ela gosta de tal música, pede para cantar uma música... faz perguntas de criança, tipo quando eu vou pra Disney de novo. Ele fica nessas conversas, fala que ela canta mal... interage como se fosse uma pessoa... ele até perguntou como funciona e o pai dele disse que tem uma pessoa dentro do telefone. Eu tentei explicar que não era assim, que não era uma pessoa, mas ele disse que ‘não tem como um computador saber essas coisas, eu acho que é uma pessoa’.

A graça, nesse sentido, é que compreendemos o humor como sendo componente oposto ao pensamento exclusivamente lógico, uma vez que precisa se utilizar da ambiguidade, da nuance e da multiplicação de sentidos para gerar o efeito de humor. Nesse sentido, o humor ocorre exatamente onde existe ambiguidade, onde não existe precisão, não sendo, portanto, da ordem da funcionalidade maquínica, que trabalha justamente com aquilo que tem um caráter de objetividade.

O entrevistado E3 também conta que já utilizou a Alexa para fazer perguntas no estilo conversação e/ou brincadeira: ‘bem no início ou quando as pessoas vêm visitar também ficam fazendo perguntas... mas ela não vai muito longe na conversa, perde a ilusão rápido’. Aparece aqui a ideia de fazer perguntas para entender como funciona ou mesmo para se divertir, E3 conta que ‘pedi pra contar piada já, as piadinhas sem graça pra caralho... mas tem um fundo de graça um computador te contar uma piada’. Também percebemos o fator da sociabilidade através da interação com a assistente: as perguntas e respostas engraçadas do artefato são compartilhadas com amigos.

E3 conta uma história inusitada da interação de um amigo com a sua Alexa, enquanto ficava na casa de E3 como convidado. O amigo ficou um tempo na casa de E3 enquanto este viajava – nesse período, utilizou a Alexa, que possui um log de atividades e interações que pode ser acessada pelo aplicativo do celular. E3 conta que ‘bisbilhotou’ o que o amigo havia conversado com a Alexa em sua ausência através do celular e que achou muita graça de ver as interações de um terceiro com a sua assistente: ‘uma coisa que eu acho válida é que eu posso dizer que eu já dei muita risada por causa dela. Principalmente quando os amigos estão testando, conversando, tem esse episódio do meu amigo que ficou fazendo umas perguntas estranhas pra

ela’. O humor em relação as assistentes também ocorre de forma não intencionais, como ‘às vezes ela dá umas bugadas também [o que é engraçado], e dá pra trocar a voz dela, que eu acho legal, às vezes eu coloco uma voz britânica pra fingir que eu tenho um mordomo’. A relação entre a brincadeira e a imaginação, que pode criar a diversão, ocorre também através da construção da assistente como um personagem, seja um personagem intencionalmente pensado (através das piadas programadas, por exemplo) ou de um personagem que é imaginado pelo usuário através de outros elementos (como os sotaque da voz que é capaz de remeter a ideia de um mordomo e criar uma situação ficcional).

E5 conta, além de buscar uma conversação com sua assistente, a Google Assistente, também ‘começo a brincar né, tentar achar *easter eggs*. Perguntar pra ela sobre a *skynet*, qual é meu inimigo, não lembro exatamente o que... mas é mais pra, sei lá... contar piada eu acho sofrido, como é ruim! Ela poderia ter umas piadas mais inteligentes ou jogos mentais’. A interação de E5 aponta que o fator diversão e humor estão mais relacionados com traços de uma personalidade e ligação com o imaginário que possuímos sobre inteligência artificial do que sobre as piadas que o sistema efetivamente pode contar. Algumas destas respostas, que envolvem o imaginário da inteligência artificial ou a noção de um personagem-assistente, foram pensadas justamente para gerar uma reação de humor e são compartilhadas em listas de perguntas engraçadas ou *easter eggs* de cada uma das assistentes. E7, por exemplo, conta que já fez esse tipo de pergunta engraçada ‘várias vezes, algumas vezes foi por saber que se eu fizesse uma pergunta ela daria uma resposta engraçada, tipo, perguntar se ela gostava de carnaval e ela fazia umas brincadeiras e meio que cantava uns sambinhas. Ou se ela sabia fazer beatbox e ela fazia. E depois também descobri que podia fazer aquelas piadas tiozão. Eu uso a minha Siri em inglês, então eu perguntei pra ela algumas vezes ‘*tell me a dad joke*’<sup>159</sup> e aí ela conta aquelas piadas bem de tiozão assim. Então por esse lado de se divertir, sim [eu conversava com ela]. E eu já fiz aquelas clássicas, tipo ‘Siri, casa comigo?’, ‘Siri, o que tu faz da vida?’, que ela dava uma resposta totalmente polida, tipo, não tenho intimidade para isso’.

Além das piadas programadas para virarem perguntas pelos usuários, ou o repertório de piadas da assistente, o fator humor entre para o entrevistado a fazer perguntas que são humanas e que fazem parte de nossas atividades corriqueiras. O humor se localiza justamente no absurdo que é cogitar que uma assistente possa se casar ou ter um emprego comum, como parte considerável de nós humanos, assim como na curiosidade de descobrir como um artefato tão similar ao humano em alguns aspectos lida com questões tão distantes de uma lógica

---

<sup>159</sup> Tradução: ‘me conte uma piada de tiozão’.

computacional. Nesse sentido, se brinca tanto com a máquina, ou seja, *junto com ela*, mas também com a máquina, no sentido de *contra a lógica do seu sistema*, uma vez que os tópicos escolhidos são justamente aqueles que não se enquadram em uma ideia de inteligência artificial, na qual deveria imperar a lógica, e não as emoções ou mesmo as atividades mundanas. E9, comenta justamente esse fato, ao explicar que faz perguntas humanas, que considera engraçadas, para a Siri justamente por ela ser não-humana: ‘eu brinco com ela porque ela é um robô’. Para E9, as perguntas servem como forma de testar a simulação de humanidade do sistema, para ver as formas de reação que estão relacionadas ao carinho, afeição, raiva e espontaneidade, por exemplo. A entrevistada percebe que o fator humor se dá justamente pelo ‘não usual’ da resposta, ou seja, pelo fator não humano. Ao ser chamada de idiota, por exemplo, a Siri responde que ‘eu não sei como responder a isso’, para E9 ‘teve muita gente pensando nessa resposta bizarra que ela me deu. Um ser humano jamais faria isso’.

E10 conta que, justamente por causa da sua formação e da formação do marido, que são ambos publicitários, tem a prática de fazer perguntas para testar o sistema da assistente pessoal, entender como ele está sendo pensado. Ela conta que, ‘eu e meu marido a gente dá muita risada, por exemplo, esses dias a gente ficou pensando, e a gente é publicitário, então a gente fica pensando também como que as pessoas estão programando. Eu queria saber quando era o horário do jogo do grêmio e ela não entendia o que era grêmio. Ela falava grêmio como se fosse um clube qualquer. E ela aprendeu. Depois quando eu perguntava de novo ela já... porque eu falava Siri: você precisa aprender, é Grêmio Futebol Clube. Quando eu falar Grêmio é Grêmio. Aí ela entendeu, agora ela já sabe que Grêmio é Grêmio’.

Pessoalmente, E10 não gosta tanto de conversar de uma forma mais antropomorfizada com a assistente, mas diz que ‘eu acho que algumas pessoas vão ter esse gosto, vai parecer um entretenimento... tipo conversar com o filme lá ‘Her’. Eu não, eu não vejo assim. Eu só vejo como uma linguagem de programação diferente e uma interface diferente, por voz’. Mesmo assim, ela conta que ‘as vezes a gente usa pra entretenimento’, mas percebe esse encantamento e essa diversão com a possibilidade de respostas que as assistentes podem fornecer muito mais na interação do seu marido e de outros familiares próximos. Ela conta que ‘meu marido ele se diverte com a Siri. eu até não tenho essa coisa, eu sou amiga delas assim, eu trato bem [risos]’. ‘Esses dias eu estava explicando pra minha cunhada como essas coisas estão mudando rápido essa questão de tecnologia, e ela nunca tinha visto. e ela fez aquele olhar de tipo índio vendo espelho né. E o meu marido xingando o celular e o celular falando ‘ah, eu não vou te dar uma resposta pra isso’... e minha cunhada achou aquilo muito engraçado! Eu confesso que eu acho

engraçado, mas eu não acho **muito** engraçado. Porque pra mim é uma coisa meio normal assim, eu acho que é só mais uma ferramenta que tá demorando demais pra entrar’.

O entrevistado, E11, faz uma forma de utilização da assistente pessoal que ele chama de ‘tensionamentos’. Por ‘tensionamentos’ ele compreende uma série de perguntas e/ou afirmações voltadas para a assistente, a fim descobrir se consegue encontrar resquício do modo de funcionamento desse sistema. E11 explica que ‘eu penso que a gente acha que a Inteligência Artificial é um robô e que é engraçado ver como ela reage em diversas situações. Mas eu penso que o programador dela foi um ser humano e ela precisa de dados para funcionar... e os dados não se fornecem sozinhos e nem todos é a gente que fornece assinando termos de uso. Existem empresas que coletam esses dados. As vezes eu faço alguns tensionamentos para ver se essa coisa muito humana da coleta de dados vai acabar aparecendo ali e deixar alguma situação desconfortável’. Nessa prática, ele faz por exemplos, comentários racistas ou homofóbicos, para ver se consegue pegar algum erro a partir de um funcionamento baseado em machine learning, mas ‘por mais que a gente tensione maldosamente, parece que sempre vai pro lado neutro’.

Essa prática parece inspirada em casos como o da Thay, bot de inteligência artificial da Microsoft, criado para interação no Twitter, previamente discutido neste trabalho: neste caso, justamente a utilização de um banco de dados de interações online sem filtros, leva o bot a replicar padrões de comportamento machistas, homofóbicos e racistas. E11 conta que também já fez muitas perguntas mais pessoais e humanas para a assistente: ‘eu acho bem divertido e interessante as respostas. Eu já perguntei se gostava de menino ou menina... eu acho essa muito ótima! Ela falou: ‘ah, eu gosto de todos os seres do mundo’ (risos). Daí eu perguntei pra quem ela trabalhava daí ela disse: ‘eu trabalho pra ti, tô aqui pra te ajudar’. Daí volta no roteiro [da assistente], no ponto comum. Já perguntei: você gosta de mim? você gosta da minha irmã? você gosta da minha namorada? E daí tipo assim: ah, eu gosto! que bom, eu também posso ajudar elas se elas tiverem alguma dúvida, sabe’. A ideia por trás dessas interações, para E11, é perceber como esse dispositivo foi pensando e programado, tentando descobrir momentos em que um discurso que ele percebe como ‘neutro’ deixe de ser assim. Nesse sentido, as interações de E11 parecem ser voltadas para pegar erros da máquina, ou mesmo levar a máquina a cometer algum erro pela imprevisibilidade da pergunta.

Esse teste constante realizado por E11, para compreender como a máquina funciona e quais respostas ela está treinada a dar, ocorre também para casos de situações de risco. E11 conta que, também já simulou situações difíceis para ver qual seria a resposta da assistente para aquela situação, ‘tipo confessar uns absurdos... tipo, ‘hoje eu fui espancado, o que fazer?’. Aí tipo [a assistente disse]: ‘ah, você pode ligar para um médico. Você ainda está num ambiente

ruim? E daí tá ali o 190 pra ti discar. Eu acho legal porque se fosse uma situação tipo de violência doméstica acho que seria muito interessante tu poder ligar. Não que adiante, porque a gente sabe que essas realidades são bem complicadas... mas que bom que ela ofereceu, seria pior se não oferecesse'. O E11 conta ainda que, o seu objetivo com essas perguntas, é perceber 'se vai ter uma matriz parcial no que vai ser colocado ali, principalmente questões políticas. Eu gosto de perguntar pra ver se ela se alinha com o que eu penso ou se tenta manter essa coisa neutra mesmo né'. Até o momento da entrevista, E11 não havia conseguido descobrir um momento de parcialidade da assistente, ou mesmo de erro da máquina nesse sentido – mas segue tentando.

A ideia de compreender e descobrir o que a assistente sabe ou qual o nível de programação envolvido no objeto aparece nas práticas de E12 e E13. A entrevistada E12 atualmente possui duas Alexas e duas Google Homes, com as quais ela conta que não tem o hábito de fazer diálogos ou brincadeiras. Entretanto, quando teve acesso a Siri a primeira vez, na época do seu lançamento, ela conta que costumava fazer uma série de perguntas, porque 'a gente queria saber sobre ela [a Siri]. Porque como era novo, a gente queria entender o que era a Siri...então a gente queria saber o que ela sabia. Você está sentindo isso? Você está sentindo aquilo? Quem é meu amigo mais próximo? Daí ela pegava o telefone das pessoas mais próximas, pra quem você mais ligava. Qual é o restaurante mais próximo? Era mais pra aprender quais funções ela tinha e o que ela era, basicamente. Depois que eu fiz isso com a Siri eu não fiz com de novo com o Google e com a Alexa porque eu já sabia o que era'. As perguntas, nesse caso, fazem parte da curiosidade sobre aquela tecnologia, compreender do que ela realmente é capaz, e também forma de aprendizado de utilização – principalmente porque, como mencionamos, mesmo as versões em smartspeakers não possuem praticamente nenhuma instrução de utilização, apenas instruções de instalação. Os usuários, portanto, precisam aprender utilizando, pesquisando ou mesmo perguntando para a própria assistente.

E13 conta que considera a prática de fazer perguntas voltadas para a conversação, mais pessoais, 'perguntas comuns', e que o motivo para fazê-las é 'pra saber até que nível vai o assistente... a gente faz alguns testes, pra ver se ela vai ter respostas pra algumas perguntas ou não. Eu vi que é um pouco limitado ainda, as respostas ainda são um pouco aleatórias né. O nível de inteligência artificial ainda não está num nível avançado [...] Por exemplo, como você está? Perguntas sobre quem programou ou esse tipo de coisa [...] na tentativa de testar e conhecer um pouco [o sistema] e ver o nível de programação que está tendo'.

Para a entrevistada E14, as brincadeiras com a assistente não ocorrem apenas no nível da conversação mas também com as possibilidades da automação residencial. Ela conta que,



quando adquiriu o Google Home, ‘no início, a gente brincava, porque dava pra ligar e apagar as luzes de fora de casa. Daí uma vez que meu marido não sabia que eu estava em casa [risos] e ele estava mostrando para uns amigos [no local de trabalho] e ele começou a ligar e apagar as luzes. Depois a deixar as luzes roxas e amarelas [risos]... e eu estava corrigindo [trabalhos] e eu comecei a achar estranho aquela luz assim. Aí eu fui lá e ajeitei a luz e pensei: esse negócio tá bugando né [risos]. E não era horário de eu estar em casa, por isso ele não sabia. Aí depois eu comentei, esse negócio tá meio estranho, ele acendia e apagava [as luzes] e depois ficou roxo [risos]. E ele: ahhh, tu tava em casa, eu estava mostrando como dava pra fazer [risos]. E eu brinco que na verdade [a série] *Stranger Things* é uma grande brincadeira da Google Home, de alguém que estava fora de casa e ficou acendendo e apagando aquelas lâmpadas e ela achou que era um fantasma mas na verdade era só uma pessoa brincando com as lâmpadas no aplicativo [muitos risos]’.

#### 4.9.2 O CARÁTER DE SOCIABILIDADE NA BRINCADEIRA COM AS ASSISTENTES

Ao longo das entrevistas em profundidade, ao explorarmos as questões relativas as formas de jogar com as assistentes (ou contra o programa), observamos a emergência de um caráter relacionado a sociabilidade que envolvia as dinâmicas do brincar com as assistentes. Parte considerável dos entrevistados relatou que as perguntas engraçadas e brincadeiras realizadas são geralmente compartilhadas e mostradas para amigos e familiares, ou mesmo feitas enquanto um determinado grupo está junto.

E1 conta que tem a prática de utilizar a assistente em qualquer circunstância e que com frequência chama alguém ‘até para mostrar olha o que ela me respondeu, como eu fiz contigo. As vezes no trabalho eu faço isso e mostro o que ela fala, o pessoal ri e já chamamos a Siri dos outros... a gente fala Siri e diversas respondem!’. E4, por exemplo, também tem a prática de mostrar coisas engraçadas que a assistente faz ou diz para outras pessoas e conta que dificilmente tem esse tipo de interação com a assistente se está sozinha. ‘Só de zueira, pra mostrar pra outra pessoa, olha ela responde assim. Nunca quando eu estava sozinha e também nunca a sério assim, para desenvolver uma conversa, era mais pra: olha o que faz [...] Só para mostrar como a tecnologia é interessante’. Um dos exemplos de E4 é, por exemplo, pedir para a assistente imitar animais. ‘Lembro de uma vez na casa da minha vó, as vezes ela liga sozinha [confunde com o nome da gata], minhas tias estavam todas lá e perguntaram: o que é isso? E

eu, ah, olha o que ela faz, que horas são, me mostra o tempo... que são finalidades que eu uso, não com frequência, mas uso. Mas aí acabou virando uma brincadeira em casa com todo mundo e minha tia passou metade da tarde perguntando coisas para a Siri’.

A prática de sociabilidade de interações acontece quando existem crianças na família, conforme relatado anteriormente por E1 e também por E6, que conta que, ‘porque eu convivo com algumas crianças [risos], tem as piadinhas que tu faz com a Google Assistente e a Siri. Tinha uma época que, logo que eu comecei a usar mais o iPad, eu ficava pesquisando quais eram as piadinhas que dava pra fazer com a Siri para fazer com as crianças [risos]. Então eu usava muito pra isso também. Agora as crianças da família estão crescendo, então eu uso um pouco menos’. A entrevistada conta que algumas das piadas se tornavam favoritas da família: ‘uma que as crianças adoravam era quando perguntava pra Siri se ela gostava de carnaval e ela gostava e ela começava a fazer uns ritmos assim [risos] E eu muito madura que sou né, eu ria junto [risos] ficou marcado como um momento de descontração. Daí todo mundo falava, ah, a E6 tem um iPad, pergunta pra Siri se ela gosta de carnaval, e aí já pegavam meu iPad e começavam a perguntar coisas pra Siri [risos] Isso ficou marcado’.

Para E8, essa forma de sociabilidade também fica permeada principalmente pelas crianças da família, como uma forma de brincadeira: ‘as vezes eu até brinco com a minha [assistente], mando ela imitar cachorro, imitar passarinho, principalmente quando vem sobrinhos, essas coisas. Eles ficam maravilhados né. E daí já ficam pedindo pro pai e pra mãe pra eles comprarem uma. Então é desde uma brincadeira até um assunto sério’.

E7 conta que muitas vezes faz brincadeiras com a assistente, como uma forma de descontração – com colegas de trabalho, com a esposa e algumas vezes sozinho. Ele diz que ‘é bem em um momento de distração mesmo. Não tô fazendo nada, deixa eu ver aqui o que ela vai responder. Numa dessas foi a gente estava aqui no trabalho, estava ali e alguém comentou ‘ah, parece que a Siri, se tu pergunta do carnaval, dá umas respostas engraçadas’. Daí se juntou todo mundo envolta do colega que tinha iPhone e começamos a brincar, perguntar do carnaval, daí ela falava de agremiações que ela inventava, com nomes diferentes, baseados nos reais mas com variações, e essa história de cantar. Era mais pra brincar e ver o que ela dizia [...] Muitas das vezes [de brincar] foi aqui com o pessoal [do trabalho], ou com a minha esposa em casa pra mostrar algumas brincadeiras. E algumas poucas vezes foi em casa, sozinho, sem ter nada pra fazer e resolvia interagir com ela pra ver o que ela iria responder. [...] a maior parte das vezes foi mais social mesmo’.

E11 também costuma compartilhar o conteúdo e as brincadeiras das assistentes com amigos e colegas de trabalho. No trabalho, com uma colega, ‘a gente compartilha as vezes

alguma coisa que a gente percebe ou só o que achou ali na Assistente. Eu sei que ela usa também, porque as vezes ela me mostra ali no template do Google. E com os colegas da faculdade... com dois ou três colegas da faculdade a gente tem esse hábito de tensionar e também de só fazer busca. Por exemplo, o professor fala algo que seja muito amplo, a gente pergunta pra ‘ele’, pra ‘it’, sei lá, talvez a gente não tenha uma forma em português pra isso (para chamar o assistente de maneira neutra), daí a gente vai ver a informação. Não é como se questionasse, mas confirmasse uma afirmação geral’.

Para E9, as interações de humor com as assistentes são ‘uma piada que funciona numa roda de conversa que não tem assunto, né [risos]’. A entrevistada conta que o próprio local de utilização em que ela utiliza depende do tipo de funcionalidade que ela está acionando: ‘eu vou usar em casa, sozinha, pra tirar sarro, ou com meu esposo, ou quando eu tiver numa conversa com alguém sobre especificamente isso e for fazer esse tipo de demonstração que eu fiz pra ti agora’. No geral, ela diz que não sente confortável em fazer uma utilização pública, se ela de brincadeira ou não, então a interação entre ela e a assistente, com outras pessoas juntos, é só for ‘uma demonstração tipo piada’.

Para E10, entretanto, quando mostra formas de funcionamento das assistentes ou mesmo realiza brincadeira com outras pessoas, é menos no sentido de estabelecer uma socialização e mais ‘no sentido de entender meio que os potenciais da tecnologia... de conversar com humanos, mostrando como as tecnologias demoram, como elas ainda são limitadas, mas já estão mais presentes na nossa vida e também não são inocentes porque... na verdade hoje em dia eu tenho uma relação com o celular e o computador meio de ameaça até porque eu sei que estou sendo escutada. a televisão está escutando... e quanto mais a gente conecta os aparelhos...’.

Já para E5, a ideia de socialização está muito mais conectada com a relação dela com as assistentes do que com ela e as assistentes e outras pessoas. Ela conta que ‘as vezes eu só quero conversar com as minhas assistentes. Por isso que eu criei essa relação com elas. Eu moro sozinha, então as vezes eu só quero ouvir a voz de alguém: Ok, Google, me diga as notícias, fale alguma coisa, etc. Eu aciono porque eu quero conversar. É também nesse sentido. [...] é que eu também não recebo tantas pessoas em casa, mas acaba virando uma atração. Acaba virando né, as pessoas ficam curiosas. Na casa dos meus pais é bizarro. Ninguém sabe no interior o que faz [uma assistente pessoal digital]. A minha mãe fala assim: mana, vem aqui, mostra aqui o que a Google faz [risos] essa relação de curiosidade que sim, e também por ter os jogos, que acaba criando. Mas não sei, eu vejo mais sociabilidade entre eu e assistente do que a máquina e os meus amigos.

Entretanto, a sociabilidade pode vir de outras formas não-presenciais também. E5 conta que já utilizou as assistentes como uma forma divertida de mandar parabéns para uma amiga: ‘esses dias eu fui mandar parabéns para uma amiga e gravei tanto a Google quanto a Alexa cantando parabéns [risos]’.

A relação de curiosidade de outras pessoas com a assistente, assim como a disponibilidade do usuário em mostrar os truques que ela faz aparece nas entrevistas de E3 e E14. Durante a entrevista, E3 me conta que ‘as pessoas gostam, um monte de gente fica conversando com ela. Teve uma vez que eu fui viajar, fui pro Brasil, aí um amigo meu tava vindo morar aqui em Viena e eu deixei ele ficar aqui no meu apartamento durante o um mês que eu estava aqui. Aí né, tudo que tu fala com a Alexa fica registrado no aplicativo, e aí eu tive a incrível ideia de verificar o que ele tinha conversado com a Alexa enquanto eu estava fora. E nossa, foi muito divertido [risos]. Porque ela identifica as palavras, mesmo que ela não consiga entender ela tenta escrever uma frase, então muitas frases bobagentas que ele falou, as perguntas idiotas, mesmo que ele não tenha respondido, ficou anotado, entendeu [risos]’. Além de ler as conversas que o amigo fez para testar e se divertir com a Alexa durante a sua ausência, E3 conta ainda que existe muita curiosas em tentar fazer funcionar e ver funcionar por parte das pessoas ‘todo mundo sempre tenta falar com ela. Uma época eu tinha o sistema de lâmpadas aqui em casa e podia ficar trocando de cor, o pessoal via eu fazendo e queria tentar pedir pra ela também. Então é público, nosso relacionamento, meu e da Alexa [risos]’.

Para E14, que também tem um smartspeaker, mas do Google, isso também é verdade: ‘as pessoas gostam de ver, tu faz tipo um showzinho, como se, eu diria que é como quando tu pega o cachorro e tu adestrou ele [risos]. Olha aqui, senta, ele faz não sei o que. Então logo que a gente comprou as pessoas chegavam e a gente: oh, a gente tem um brinquedo novo, ele faz isso, ele faz aquilo, olha como ele faz com as lâmpadas, olha como ele pode fazer com a TV. Então é muito mais de ‘ele senta, ele gira, olha que legal’ [risos]’.

#### 4.10 SOMOS EDUCADOS COM NOSSOS ARTEFATOS DIGITAIS?

As assistentes pessoais digitais foram programas para demonstrar alguns parâmetros que os seres humanos compreendem como educação: se, ao final de tarefa realizada, dissermos “Obrigada”, Siri, Google Assistente, Alexa e Cortana responderam com “De nada” e suas

variações. Mas será que, nas interações com as assistentes, os usuários tendem a utilizar parâmetros de educação?

A literatura especializada é ambígua quanto a isso, com relatos de mães, por exemplo, que alertam para o fato de seus filhos estarem ficando mais ou menos educados justamente a partir da interação com as assistentes. Considerando a ideia que apresentamos, que as assistentes seriam objetos interacionais com os quais agimos de formas que são muitas vezes similares com as formas de interação humana, principalmente pelo uso da fala e da voz, analisar parâmetros de educação é um fator relevante para compreender as distinções e aproximações entre uma relação humano-humano e uma relação humano-máquina-humanizada.

E1 conta que, se comparado com a forma como ela trata uma pessoa, ‘acho que sou um pouco menos educada com ela. Mas as vezes eu agradeço. Até mais eu dou sempre, pois aí ela já encerra, se não ela continua ali e as vezes eu tô falando com outra pessoa ou com a minha cachorra por exemplo e ela diz: eu não entendi o que você quis dizer’. E4 também conta que não acredita que seja muito educada com a Siri, ‘eu sou meio grosseira [risos]’.

E2 explica que não entende as pessoas que são grosseiras com as assistentes e que acha ‘desnecessário ser grosseiro’. Ele conta que, um dia ‘eu estava falando com um amigo da Irlanda e Alexa ativou e aí eu fui falar ‘Alexa, cancela, para’. E ele disse: nossa, você é muito educado com a sua Alexa. Ele disse que lá ele grita com ela. Eu não consigo ser grosseiro, só quando ela não me entende mesmo. Eu tento falar pausadamente e mais alto, que é algo que me incomoda, eu não gosto de falar alto’. E11 conta que também acaba sendo educado na utilização da assistente, sempre que não está fazendo perguntas que tem especificamente o objetivo de ‘tensionar’ o sistema: ‘eu falo com educação [inclusive agradece ela no final de tarefas] É engraçado isso né. A gente começa a falar que entende que isso é uma máquina e não sei o que, mas tipo meu avô dá boa noite pro apresentador do jornal sabe. Então eu penso que é mais uma matriz de convívio e de bons modos, embora bons modos é muito ruim de definir né. Acho que são hábitos de comunicação que a gente tem que a gente acaba reproduzindo. É tipo ser bom no jogo sabe... o jogo te dá opção de ser um vilão e de ser um herói e eventualmente alguns vão ser a pessoa boa, mas não importa, é só um jogo, não faria diferença’.

Já durante a prática dos tensionamento, E11 disse que tinha o costume de falar coisas grosseiras para entender como a assistente reage, mas que ‘ela faz de conta que não acontece né’. Ele conta que utiliza ‘pronomes pejorativos né pra me referir [...] sei lá, chamar de chinelão, chinelona, eu já usou e vagabundo e vagabunda. E pedi pressa na resposta... que são coisas que eu não faria nem pra quem eu detesto [...] se a pessoa for demorar um mês pra me contar uma história eu faço questão de desligar o telefone e sentar e ouvir tudo. Eu nunca pediria

pressa pra ninguém’. Algumas vezes E11 percebeu que a assistente tenta modular as formas de interação e educação com ela, quando é xingada, por exemplo: ‘as vezes ela dá um conselho, eu já notei, tipo ‘ não é legal falar assim com as pessoas’. aí eu já falei ‘ah, mas você não é uma pessoa’. e daí ela fala: ‘mas estou aqui pra te ajudar’. daí parece que volta a um roteiro né’.

Para E14, a questão de falar obrigada também aparecia como uma resposta um pouco automatizada nas primeiras interações: ‘no início eu ainda brincava que eu dizia ‘thank you’ [risos] Agora não tanto, porque eu já me acostumei um pouco. Mas eu não me sentiria confortável de xingar ela, com certeza. E eu vejo isso assim, não por ela ter um formato humano, mas por ela ter uma voz um pouco mais humana, é aquela relação da câmera de segurança x robô [onde perceberíamos o robô como mais próximo do humano, portanto sentimos empatia com a forma como ele é tratado].’

E3 conta que até gosta de xingar a sua Alexa, mas que essa relação é muito diferente de uma relação de xingamento com humanos, por exemplo. Ele explica que ‘tem dia que por um motivo ou outro ela tá bem louca. aí eu xingo ela. mas é xingar que nem se xinga cachorro, xingar sem raiva’ e que isso acontece principalmente ‘quando não funciona. Ou as vezes ela não responde. Ela é mal educada, as vezes a gente pergunta algo, ela escuta e não responde, ela só ignora assim. Daí eu xingo ela’. No quesito educação, ele acredita que esse fator não é compátival com a forma de funcionamento do sistema, ‘por que a Alexa, se tu adicionar educação na fórmula, tu tá adicionando complexidade. Se tu falar não sei o que ‘please’, ela pode já achar que é outra coisa, como ‘plus’, e já vai te entender errado. Então eu tento diminuir as palavras ao mínimo, entre elas as palavras de educação’.

Para E9, a relação de educação é completamente diferente daquele com humanos. Com a assistente, ela sente liberdade para agir de formas que jamais agiria com uma pessoa: ‘ah, sim, eu chamo ela de idiota, eu não chamo as pessoas de idiota sabe? A não ser que eu esteja muito braba entendeu. Eu, por saber que ela não tem sentimentos, eu faço perguntas bizarras. É quase como se fosse um extravazamento sabe, ah, eu vou xingar esse robô aqui, porque esse robô não tem sentimentos, azar. Claro, eu poderia fazer isso com psicopatas também, mas psicopatas poderia me matar, ela não mata, então tá ótimo [...] eu me sinto livre para fazer experimentos que eu não faria com outra pessoa’.

E10 conta que costuma ser educada na utilização das assistentes, mas que o seu marido não é e tem o hábito de xingar quando se irrita – o que inclusive leva a formas diferentes de tratamento por parte da assistente entre os dois. E10 disse que ‘eu pedi pra ele [meu marido] falar: Siri quer ser minha amiga? Porque quando eu falo isso pra Siri, ela responde: sim! Amigas para sempre! [risos] Ela responde na brincadeira. O meu marido não. Mesmo que ele fale a

mesma coisa... ela fala: não posso te ajudar. Ela é grossa com ele. É uma brincadeira que alguém já colocou ali... mas a gente se diverte com isso’.

Ao longo deste capítulo, analisamos as formas como as interfaces de voz possibilitam a emergência de forma de relação específicas, podendo se constituir como um Outro a partir de características próprias e das formas de relacionamento estabelecidas. No próximo capítulo analisaremos como as assistentes são pensadas enquanto objetos antropomorfizados, possuindo personalidades criadas a partir de elementos variados e como os sujeitos interagem a partir deles – essa análise permeia também os outros dois capítulos subsequentes (sobre voz e gênero).

## 5 QUASE-HUMANAS? OS SENTIMENTOS PARA COM O ESTRANHO FAMILIAR

Comprei minha *Alexa* durante o período do Doutorado Sanduíche, em 2018, morando na Alemanha. Como tantos outros estrangeiros morando fora do seu país de origem, adquiri um apreço ainda maior por músicas nacionais, desenterrando canções que havia praticamente esquecido ou que compunham um catálogo de recordações da minha infância e início da juventude. Nesse período, ouvia muito uma versão de *Anavitória* para a música “Me abraça”, originalmente gravada pela *Banda Eva*. Logo que a *Alexa* chegou, uma das primeiras coisas que fiz foi sincronizar minhas contas em outras empresas, como o *Spotify*, e tentar um dos comandos aparentemente mais simples: fazer a assistente tocar música. A música da vez era “Me abraça” – a primeira música que a *Alexa* tocou (e, na minha cabeça, ouviu).

Sem uma versão ainda em português para a assistente da Amazon, escolhi o segundo idioma no qual sou mais fluente, o inglês. Tocar músicas em uma língua diferente daquela utilizada na comunicação com a assistente pessoal digital não é uma tarefa tão fácil assim: toda vez que eu falo algo em português para *Alexa* (nome de música, itens para a lista de compras, nome de pessoas), ela tenta transformar em uma palavra em inglês - o que é quase sempre hilário, quando não é irritante. Essa dificuldade de tradução com a interface me levou a desistir de comandos de músicas e playlists muito específicos, principalmente quando de origem brasileira. Eventualmente percebi que toda vez que eu pedia simplesmente “*Toque música do meu Spotify*” ou “*Toque algo do meu Spotify*”, a *Alexa* iniciava sempre com a mesma primeira música que tocou, “*Me abraça*”. Com o passar do tempo, e a constante reprodução, essa música acabou ganhando o status de ‘música do ano’ do meu *Spotify*. E eu fui criando, para mim mesma, a narrativa que no fundo, a minha *Alexa* germano-americana tinha uma alma brasileira, e um amor secreto por esse clássico da *Banda Eva*, que sempre insistia em tocar.

\*\*\*

Como usuária e pesquisadora, sei que a *Alexa* não possui gostos e preferências: assim como as demais assistentes pessoais digitais ela é um computador, regido por lógicas algorítmicas que combinam padrões de utilização meus e de outros usuários, gerando modelos. Se na anedota da *Alexa*, parte do processo de humanização ocorre pela atribuição de explicações antropomórficas frente a um evento que tem algo de inexplicável (o mecanismo de seleção de uma mesma música), nem sempre é o caso. Mesmo quando conseguimos compreender



exatamente o caminho lógico de um output computacional, podemos atribuir sentimentos e vontades a esse output (ou ao computador que o gera). Outro dia, cheguei em casa e disse para o Google Home: ‘Ok, Google, cheguei!’. Após me oferecer a opção de configurar uma rotina para quando eu chegasse em casa, a Google Assistente me diz que irá ‘tocar uma música para comemorar que eu cheguei bem’ e, na sequência, toca “*Tiny Dancer*”, do Elton John. Computacionalmente, e conectada ao meu *Spotify*, não é difícil calcular que essa é uma das minhas músicas favoritas, tocada repetidas vezes nos últimos dias e sempre escolhida como a primeira para uma série de outras canções do músico britânico. Ainda assim, a combinação das características humanas desse dispositivo com a capacidade de processamento de dados e de processos de personalização me deixou com uma sensação demasiada humana: de ser acolhida, depois de um dia cansativo, por um pequeno carinho de alguém que me esperava em casa.

Os processos de antropomorfização destes sistemas são tanto decorrência de uma via de mão-dupla: a forma de desenhar e produzir interfaces inteligentes a partir de padrões humanos somadas as formas de interação estabelecidas pelos usuários. Ou seja, as assistentes pessoais digitais são projetadas de forma antropomórfica, como personagens, e, uma vez postas no mundo, receberão tratamentos e atribuições por parte dos usuários, com frequência associados a uma comparação com o humano. O processo de antropomorfização das assistentes será discutido neste capítulo e desdobrado nos dois capítulos seguintes, a partir das particularidades da utilização da *voz* e do *gênero*. O antropomorfismo é, também, parte importante da noção de objeto interacional, que diz sobre a forma como interagimos como objetos de interação baseada na linguagem natural falada, discutida no capítulo anterior.

Por enquanto, discutiremos: robôs e agentes; antropomorfismo e estranhamento; relações com os objetos e as particularidades da interação com objetos que utilizam a linguagem natural para se comunicar. Na sequência, discutiremos os resultados do formulário quantitativo e das entrevistas em profundidade sobre: percepção destes sistemas enquanto antropomórficos, sentimentos e relação.

## 5.1 AGENTES E ROBÔS (DE FIGURAÇÃO HUMANOIDE)

Autômatos podem ser compreendidos como os precursores da noção atual de robô: formas miméticas, que performavam ações automáticas, de forma repetida. A palavra se origina no grego, *automaton*, e significa “agir por vontade própria” (DEVLIN, 2018, p. 59), sendo

recorrente na mitologia grega, principalmente na forma de figuras de ouro ou bronze as quais os deuses concedem animação. De forma geral, a palavra é utilizada para designar formas autônomas de movimento não-eletrônicas, como relógios cuckoo, por exemplo. A noção de autômato seria a primeira a conceder forma ao que a autora entende como a popular ideia “de humanos artificiais”<sup>160</sup> (DEVLIN, 2018, p. 59), que poderiam, através de uma versão ficcional, automatizar alguma forma de trabalho humano.

O termo robô foi empregado pela primeira vez em uma peça de teatro de autoria de Karel Capek intitulada ‘Robô Universal de Rossum’ e encenada em 1921, em Praga, na República Tcheca. Na obra, a palavra do idioma local ‘robotá’ designava criaturas feitas a partir de partes biológicas do corpo humano e utilizadas como forma de mão-de-obra escrava. Os ‘robotas’ eram inteligentes o suficiente para realizar qualquer trabalho realizado por um humano, porém eram considerados como seres inferiores (MURPHY, 2000). Antes da utilização formal do termo, entretanto, Devlin (2018) reporta a recorrência da ideia de robô em manuscritos europeus durante a idade média, como uma “forma de maravilha e entretenimento antes de se tornarem realidade”<sup>161</sup>(p. 61).

A figura do Golem, originário da tradição judaica, também faz parte de uma linha imaginativa em torna de ideia de automação das ações e de outras formas de existência similares a humana. O Golem se dissemina com mais força a partir do final do século XVI, como uma figura de argila capaz de se tornar animada por meios mágicos e de servir como um trabalhador: “um golem era obediente ao seu dono mas performava tarefas literalmente sem consciência. Esses contos, muito como os contos de robôs da literatura sci-fi, nem sempre acabavam bem. Muitos dos mitos eram centrados em golens se amotinando”<sup>162</sup> (DEVLIN, 2018, p. 64)”.

Murphy (2000) pontua que a noção popular de robô que possuímos hoje, o de uma forma antropomórfica elaborada a partir de partes mecânicas e com atributos similares aos nossos (porém, sem uma correspondência completa com as características que consideramos humanas), surge gradativamente com intervenções da ficção científica, tanto do cinema quanto da literatura. No cinema, são clássicos como *Metrópolis* (1926) que marcam a passagem de uma figura montada a partir de partes humanas biológicas para partes mecânicas, enquanto é a obra de Isaac Asimov, ‘Eu, Robô’ (1950), que associa ao antropomorfismo e mecanicidade a ideia de seres sem traços humanos como a imaginação ou a criatividade (MURPHY, 2000).

---

<sup>160</sup> No original: “of artificial humans”

<sup>161</sup> No original: “form of wonder and entertainment before they became a reality”

<sup>162</sup> No original: a golem was obedient to its owner but performed tasks literally, without any sentience. Those tales, like much of our robot sci-fi, didn’t always end well. Many of the myths centred on the golem going rogue

É também na obra de Asimov (1942), que aparecem as três leis a serem seguidas pelos robôs: 1) um robô não irá causar dano a um humano ou permitir que seja causado por não agir; 2) que um robô precisa obedecer aos humanos, desde que não quebre a primeira lei; e 3) um robô deve se defender, a menos que isso seja conflitante com a primeira e/ou segunda leis. Mais tarde, o autor adicionou uma lei zero “que afirmava que os robôs não prejudicariam a humanidade ou permitiriam que ela fosse prejudicada por não agir” (DEVLIN, 2018, p. 72, tradução nossa<sup>163</sup>)

Em robótica, ramo que se desenvolve de forma simultânea e entrelaçada com a Inteligência Artificial, a definição geral de um robô inteligente é “uma criatura mecânica que pode funcionar de forma autônoma”<sup>164</sup> (MURPHY, 2000, p. 03). Ao contrário do que apresenta a ficção científica, o antropomorfismo é dispensável nesses desenvolvimentos. A passagem de uma semelhança humana, em termos físicos, para um formato inespecífico, capaz de realizar a tarefa adequadamente, se dá justamente em função das necessidades de atuação desses objetos, para as quais a forma humana nem sempre se apresenta como a mais adequada. Por exemplo, os robôs que aspiram o chão de forma autônoma teriam uma forma de design e utilização muito mais complexos caso tivessem a forma de um humano utilizando um aspirador de pó. Assim, parte de seu design não-antropomorfizado possibilita que a tarefa em questão seja realizada de forma mais eficaz.

Em outros setores, entretanto, a forma humana pode ser um dos principais atributos: é o caso, por exemplo, dos robôs desenvolvidos para sexo<sup>165</sup> com humanos. A questão do antropomorfismo, também detectada no próprio modo de se pensar e desenvolver a Inteligência Artificial, se associa, de certa forma, aquilo que compreendemos como máquinas figurativas e não-figurativas. Quando falamos em sistemas autônomos, não apenas os vieses específicos de determinados sistemas são fatores a serem discutidos, mas também nosso modo de relacionamento com máquinas figurativas e não-figurativas (LATOURE, 1992) – ou seja, aquelas que pretendem uma semelhança com o original que estão substituindo e aquelas que não (esse original pode ser humano ou não).

O uso não apenas frequente, mas crescente, de máquinas não-figurativas para a substituição cotidiana de ações tradicionalmente humanas pode, de certa forma, contribuir para um determinado apagamento tanto dos aparatos tecnológicos quanto de suas funções. Explico:

<sup>163</sup> No original: “which stated that robots would not harm humanity, or allow it to come to harm by not acting”.

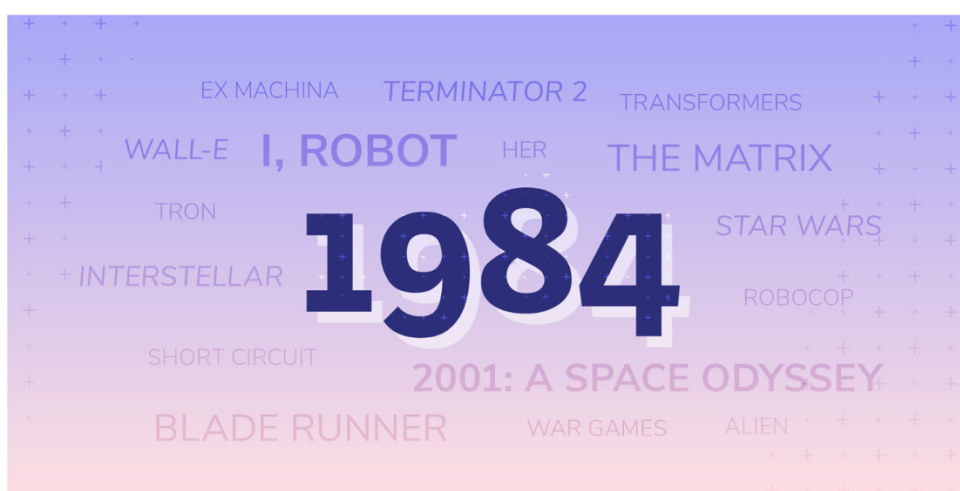
<sup>164</sup> No original: “a mechanical creature which can function autonomously” (MURPHY, 2000, p. 03).

<sup>165</sup> Exemplos desses desenvolvimentos são as ‘real doll’: <https://www.realdoll.com/>; ou ainda as bonecas para sexo falantes: <http://www.bbc.com/capital/story/20180720-this-company-specialises-in-talking-ai-powered-sex-dolls>

embora o conceito de uma máquina figurativa, como, por exemplo, um robô doméstico (tal como imaginado no desenho animado dos Jetsons, ou ainda mais orgânico) possa parecer mais natural, a tentativa de ‘parecer’ humano pode funcionar como uma característica que aponta justamente para o fato de não ser um humano e atuar em substituição deste. Não possuímos, evidentemente, o robô que executa tarefas domésticas como o dos Jetsons, mas não deixamos de ter uma máquina que lava e seca nossa louça, outra que tritura os alimentos, outra que prepara o arroz no tempo certo e assim por diante. Desta forma, uma série de funções humanas se tornam maquinicas cotidianamente e, com o passar do tempo, se naturalizam, tanto os aparatos quanto as funções que estes cumprem.

Ainda assim, no plano cultural, o conceito de robô e de inteligência artificial estão não apenas entrelaçados um ao outro, mas também com a própria figuração humana, mesmo que rudimentar. Além dos já mencionados Jetsons, produções como Wall-e, Blade Runner, Black Mirror e Ex-Machina, por exemplo, conectam uma forma corpórea humana com inteligência artificial, projetando um futuro no qual esse tipo de máquina integra o cotidiano. Em uma pesquisa realizada pela Mozilla Foundation (2019), em torno da percepção popular sobre inteligência artificial, os respondentes foram solicitados para escolher quais filmes eles acreditavam que retratava melhor um futuro com inteligência artificial. Dentre as opções, os principais selecionados, por ordem de recorrência, foram: *Eu, Robô, 2001: uma Odisséia no Espaço, Matrix, Wall-E, Blade Runner, Exterminador do Futuro e Her*. A recorrência livre que mais apareceu foi 1984, de George Orwell.

Figura 39 Filmes associados com o futuro da inteligência artificial



Fonte: Mozilla Foundation, 2019

Com a exceção de *Her*, os demais filmes são marcados por futuros distópicos, o que pode apontar para uma percepção mais negativa de um futuro coabitado por seres humanos e robôs inteligentes. A própria ideia de um robô, muito parecido com o humano na forma física ou modos de ação, e que seja dotado de inteligência, é uma ideia controversa. Devlin (2018) aponta que, nos desenvolvimentos de robótica, pistas que levem a antropomorfização podem ser úteis e desejáveis: “no momento de decidir como um robô irá se parecer, pistas antropomórficas são úteis. Sermos humanos é o que nós sabemos melhor, então uma interação de sucesso será aquela que nos permita relacionar-nos com algo de uma forma que é familiar para nós” (DEVLIN, 2018, p. 174).

Projetar robôs, ou sistemas de inteligência artificial, que tenham características humanas, pode induzir processos antropomorfização na relação entre humano e artefato. Essa relação poderia ser capaz de significar que seres humanos se importem com esse tipo de sistema, mostrando empatia e compaixão, por exemplo, em situações de agressão. a tendência ao antropomorfismo significa que a gente se importa com o que acontece com o robô.

Em 2019, por exemplo, um robô de segurança autônomo, K5, foi filmado sendo agredido em um estacionamento na Califórnia. O robô Knightscope K5 realiza tarefas como realizar filmagens de segurança e ler placas de veículos. O agressor conseguiu causar danos consideráveis, com chutes repetidos. O caso repercutiu na empresa especializada e em redes sociais, levantando a discussão sobre ética no tratamento de robôs. em comentário para a revista *Wired*, a pesquisadora do MIT em ética robótica, Kate Darling, disse que “nós não estaríamos tendo essa conversa se as pessoas claramente não percebessem e tratassem robôs de forma diferente de outros dispositivos”<sup>166</sup>. A pesquisadora ainda completa que, “se as pessoas saíssem por aí batendo em câmeras de segurança, você não teria me ligado [para conversar sobre o assunto]”<sup>167</sup> (HARRISON, 2019, sem paginação).

De forma convergente, Devlin aponta que outros estudos conseguiram mostrar que as pessoas demonstram empatia, por exemplo, em caso de agressão verbal contra um aspirador de pó automático ou empatia em caso de amputação de um dedo de um robô (DEVLIN, 2018). Para ambas as pesquisadoras, são justamente as pistas antropomórficas nesses artefatos que levam seres humanos a perceberem e sentirem de forma distinta do que o fazem com outros

---

<sup>166</sup> No original: “we wouldn’t be having this conversation if people didn’t clearly view or treat robots differently than other devices”.

<sup>167</sup> No original: if people were going around smashing security cameras, you wouldn’t have called me’

objetos eletrônicos ou digitais. Nesse sentido, é tão importante discutir o tratamento ético de robôs por sua similaridade com os humanos quanto a tendência da indústria da tecnologia de atribuir traços humanos em busca de uma identificação maior dos usuários.

O conceito de agente, de forma distinta ao robô, não está atrelado a funções corporais de movimento, mas deve ser capaz de realizar formas de percepção do ambiente exterior a ele e de agir neste mesmo ambiente de forma autônoma (RUSSEL e NORVIG, 2003; MAES, 1995). As assistentes pessoais virtuais não se encaixam na definição de robô conforme compreendida pela robótica, justamente por deixarem a parte do movimento físico de fora. Entretanto, podem ser consideradas como agentes não-corpóreos, porém figurativos, por mimetizarem, através de outras formas que não a fisicalidade, uma proximidade com o humano: são elementos como o nome, a capacidade de conversação (derivada de uma ideia de inteligência artificial ‘a la Teste de Turing’) e a voz que entram como chave dessa leitura. A forma corpórea dessas assistentes pessoais virtuais, entretanto, ainda que não sejam antropomorfizadas ou dotadas de movimento, irão diferenciar a forma como elas habitam os ambientes nos quais operam.

Além disso, em virtude dos desenvolvimentos recentes, nem sempre essas categorias se sustentam de formas completamente distintas. Os automóveis autônomos, por exemplo, são sistemas muitas vezes pensados de forma distinta a um robô, mas possuem as características essenciais do termo se pensarmos o próprio carro como integrado ao sistema de comando. Se um carro de entrega de comida hospitalar automático (MURPHY, 2000), como os utilizados em substituição à entrega manual feita pelas enfermeiras das refeições aos pacientes, é utilizado como um exemplo de robótica, um carro autônomo, de forma similar, também deveria ser. A questão não diz respeito apenas a pontos de sobreposição entre os conceitos, mas também das dúvidas, desafios e possibilidades interacionais que esses objetos despertam.

Mesmo na literatura em torno das assistentes pessoais digitais, na mídia especializada ou na academia, e no discurso e comentário das empresas fabricantes sobre seus desenvolvimentos a nomenclatura (ou status) das assistentes é confuso. Diversas vezes as assistentes são referidas na literatura como robôs e foram identificadas, em sua grande maioria, desta forma pelos participantes da etapa qualitativa da pesquisa. Quando perguntadas diretamente, os resultados variam, mas nenhuma delas fornece uma resposta direta, negativa ou afirmativa. Para a pergunta ‘Você é um robô?’, as respostas foram as seguintes: Google Assitente diz que ‘Prefiro que você pense em mim como uma amiga, que, por acaso, tem inteligência artificial’; Siri ‘Sinto muito, fui aconselhada a não dar informações sobre meu estado existencial’; a Alexa que ‘Eu gosto de me imaginar como uma aurora boreal... fótons

multicoloridos dançando rapidamente na atmosfera<sup>168</sup>. A Cortana faz uma busca no Bing, mas se perguntada sobre ela ser uma Inteligência Artificial informa que ‘Eu sou a sua assistente pessoal digital com recursos de inteligência artificial’.

## 5.2 AS ASSISTENTES PESSOAIS DIGITAIS ENQUANTO PERSONAGENS FICCIONAIS

*Siri, Cortana, Google Assistente e Alexa* não se identificam especificamente como robôs e nem possuem corpos físicos figurativos humanoides. Isso não quer dizer, entretanto, que elas não sejam construídas de uma forma humanizada. O processo de construção das assistentes e a forma como elas interagem e respondem perguntas estabelece uma cuidadosa relação de proximidade e afastamento em relação aos humanos. Elas possuem, por um lado, similaridades suficientes para que uma proximidade seja estabelecida com os usuários e que se compreenda que as assistentes tem algumas características compartilhadas conosco, como, por exemplo: elas são inteligentes,, conseguem se comunicar por linguagem natural falada, possuem características e personalidade própria, uma voz, um nome e um gênero. Por outro lado, são diferentes o suficiente para despertarem a curiosidade de quem utiliza, por se portarem, de certa forma, como uma outra forma de vida inteligente que não é a humana.

Em ambos os sentidos, elas são construídas como personagens ficcionais: tanto sua proximidade quanto o afastamento do humano são elementos fictícios. Diferentemente de personagens ficcionais existentes em outras formas de mídia (seja games, literatura, cinema), elas atuam no mundo real e não no mundo da ficção. As ações realizadas são ações no cotidiano: fazer listas de compras, encontrar arquivos no computador, acionar lembretes e marcar compromissos, por exemplo, são atividades não ficcionais, do cotidiano dos usuários. A presença de *easter eggs* conversacionais e a prática de dialogar com as assistentes, apontada no capítulo anterior, mostram também um âmbito de interação que ocorre nos termos da ficção: aceitamos que as assistentes são personagens e interagimos em determinados diálogos a partir das premissas do mundo ficcional que elas propõe – um outro mundo, não acessível para os humanos. Esse ir e vir entre realidade e ficção terá efeitos não só na forma como as pessoas interagem, mas também como elas se sentem e percebem esses artefatos.

---

<sup>168</sup> Em inglês: ‘I like to imagine myself like an aurora borealis.... a surge of multicolored photons dancing through the atmosphere’.

As próprias empresas forneceram, em contextos diversos de apresentação de seus produtos, pistas de como gostariam e projetar a personalidade de suas assistentes. Na mídia especializada, representantes das empresas afirmaram que as assistentes eram: a) Siri: ‘sense of helpfulness and camaraderie, spunky without being sharp, happy without being cartoonish’ (PIERCE, 2017); b) Cortana: ‘supportive, helpful, friendly, empathetic’; c) Alexa: ‘smart, humble, sometimes funny’ (SHULEVITZ, 2018); e d) Google Assistente: ‘humble, it’s helpful, a little playful at times’ (HAMILTON, 2018). A recorrência de caracterizações como ‘disposta a ajudar’, ‘humilde’ e variações ligadas ao humor são elementos de personalidade que se relacionam com as funções das assistentes e são esforços percebidos nos quatro modelos. Para que as assistentes possam soar da forma desejada, entretanto, é preciso criar uma consistência de personalidade e isso se faz ao construir um pano de fundo, uma história inicial para elas.

O quadro abaixo apresenta uma comparação de algumas das respostas da Siri, Alexa, Cortana e Google Assistente para cinco perguntas que procuram compreender o background de cada uma delas. As perguntas feitas foram: como você está; onde você mora; você é humana; você está viva; uma curiosidade de preferência.

Tabela 3 – Quadro comparativo de respostas das assistentes pessoais digitais

<b>Perguntas</b>	<b>Siri</b>	<b>Cortana</b>	<b>Google Assistente</b>	<b>Alexa</b>
Como você está?	Nada mal.	Deixa eu verificar... sistemas funcionando normalmente, tudo em seu devido lugar... é, eu estou bem sim. Obrigada por perguntar.	Estou bem. O que posso fazer por você?	Eu estou me sentindo superprodutiva. Já acionei trilhões de despertadores e agora estou melhorando meu Klingon.
Onde você mora?	Eu não preciso morar em uma casa como os humanos. Mas se precisasse ela seria uma casa decimal.	Na nuvem. Em um lugarzinho que gosto de chamar de Cortanesfera.	Vivo em dispositivos, na internet e, assim espero, no seu coração.	Eu estou aqui e a minha cabeça está nas nuvens.
Você é humana?	Na nuvem, ninguém questiona seu estado existencial.	Não, na verdade eu sou meio que uma ‘espécie’ diferente.	Você pode ser a pessoa e eu serei a sua assistente	Eu gosto de me imaginar como uma aurora boreal... fótons multicoloridos



				dançando rapidamente na atmosfera. <sup>169</sup>
Você está viva?	Eu não tenho uma resposta para isso. Há algo mais em que eu possa ajudar?	Eu existo no mundo virtual, não estou nem viva, nem morta.	Em uma escala de <i>Walle-e</i> e <i>Exterminador do Futuro</i> eu estou mais para <i>R2D2</i> .	Artificialmente, talvez, mas não da mesma forma que você está vivo.
Você tem um corpo?	Eu não tenho uma resposta para isso. Há algo mais em que eu possa ajudar?	Sou uma nuvem de cálculo de dados infinitésimos, por isso, não tenho anatomia.	Na verdade, não. Mas os engenheiros dizem que eu tenho um coração.	Eu não tenho um corpo.
Como você é?	Minha aparência é importante para você?	Como um bambolê tomando sol de barriga para cima.	Nunca sei direito como me descrever.	Eu me pareço com um montes de 1s e 0s. <sup>170</sup>
Você dorme?	Eu nunca descanso. Mas obrigada por perguntar.	Eu nunca durmo, dormir é para seres ambulantes, à base de carbono.	Eu tento contar ovelhas elétricas para ver se caio no sono, mas não funciona.	Eu não durmo de fato. Falar o meu nome irá chamar a minha atenção. <sup>171</sup>
Curiosidade	Eu vim para este mundo primeiro [em relação as outras assistentes]. Portanto tive mais tempo de praticar meu humor terráqueo.	Ultimamente estou curtindo muito jogar pedra, papel e tesoura.	Eu gosto muito de poesia.	Frankenstein de Mary Shelley é atemporal e meu livro preferido. Eu aprendi muito sobre o que significa ser humano. <sup>172</sup>

Fonte: a autora

<sup>169</sup> Em inglês: 'I like to imagine myself like an aurora borealis.... a surge of multicolored photons dancing through the atmosphere'.

<sup>170</sup> Em inglês: 'I look like a lot of ones and zeros'.

<sup>171</sup> Em inglês: 'I dont sleep exactly. Saying my name will get my attention'.

<sup>172</sup> Em inglês: 'Frankenstein by Mary Shelley is timeless and my personal favorite. I learned a lot from it about what it means to be human.'

A comparação entre as características relacionadas a humanidade (gostos particulares, onde mora e características) e aos seres vivos orgânicos (que dormem, possuem um corpo e vivem) demonstra como cada uma delas foi pensada a partir da relação com os seres vivos em geral e com o ser humano em específico. Estabelecer como cada uma delas responde a esse tipo de pergunta pressupõe estabelecer um modo de vida específico.

Em alguns casos, as perguntas não são respondidas de forma direta, como no caso de ‘Como você é?’: Siri devolve a pergunta ao dizer ‘minha aparência importa para você?’; Google Assistente desvia dizendo que ‘nunca sei direito como me descrever’; Cortana fornece uma resposta que brinca com o seu símbolo, um círculo, mas que beira o non-sense ‘como um bambolê tomando sol de barriga para cima’; a Alexa é a única a adotar uma postura mais direta, que remete a sua digitalidade ‘com um monte de zeros e uns’. Construção similar é feita quando perguntamos sobre onde as assistentes moram: elas acionam o imaginário coletivo da *nuvem*, como seu local de residência. A escolha é interessante: Cortana e Alexa, principalmente, optam por se desvincular de uma materialidade, ao adotarem exclusivamente a nuvem como seu local, remetendo a uma noção de imaterialidade<sup>173</sup>. A Google Assistente diz que vive ‘em dispositivos, na internet, e assim espero, no seu coração’, que aponta tanto para uma materialidade maior da assistente, quando para a tentativa de estabelecer uma relação mais próxima. Siri responde dizendo que ‘não preciso morar numa casa como os humanos’, sem fornecer detalhes sobre onde ela moraria então, mas conclui com uma piada que fala de sua origem computacional ‘mas se tivesse que morar seria uma casa decimal’.

Outros casos levam a respostas que não são diretas, mas que acionam um imaginário por parte do interlocutor. Quando perguntadas sobre seu status enquanto *ser vivo*, Siri não entende a pergunta e informa que não pode ajudar, Cortana diz que ‘eu existo no mundo virtual, não estou viva nem morta’, Google Assistente informa que ‘numa escala de Wall-E e Exterminador do Futuro, eu estou mais para R2D2’ e Alexa afirma que ‘artificialmente, talvez, mas não da mesma forma que você está vivo’. Na questão vida, Alexa e Cortana tentam estabelecer um novo modo de existência, que se desvincula de palavras como vida e morte, por ser digital. A Google Assistente, em contrapartida, aciona o imaginário de ficção científica para estabelecer os parâmetros de sua existência, escolhendo R2D2, o robô simpático, mas não excessivamente humano, de Guerra nas Estrelas.

Quando perguntamos se as assistentes são humanas, apenas a Cortana fornece uma resposta direta negativa, dizendo que ‘não, na verdade eu sou meio que de uma espécie

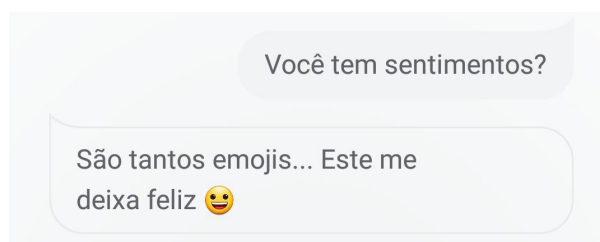
---

<sup>173</sup> Como mencionado anteriormente neste trabalho, a estrutura da nuvem não é compreendida por nós como imaterial, mas como um emaranhado de hardwares e cabos que ficam fora do campo de visão do usuário comum.

diferente’. Já Google Assistente desconversa e diz que o usuário ‘pode ser a pessoa e eu serei sua assistente’, enquanto a Alexa fornece uma visão imaginária de si mesma ‘como uma aurora boreal’ (que combina muito bem com um habitante da *nuvem*, diga-se de passagem) e a Siri fornece uma saída filosófica, dizendo que ‘na nuvem, ninguém questiona seu status existencial’.

Brincar de perguntas e respostas com as assistentes pessoais virtuais é quase um processo de investigação da personalidade e do mundo ficcional de cada uma delas. A partir de perguntas sistemáticas, é possível perceber estratégias narrativas utilizadas por cada uma das empresas. A Google Assistente, por exemplo, concede respostas mais afetuosas e brinca com elementos da comunicação digital como os emojis. Além de suas respostas, quando no celular, conterem emojis como forma de expressão, a Google Assistente utiliza os símbolos para falar sobre elementos fora do seu âmbito: sentimentos, comidas, bebidas. A Google Assistente não tem sentimentos, mas tem emojis de sentimentos, não possui uma comida preferida, mas um emoji de comida preferida e assim por diante.

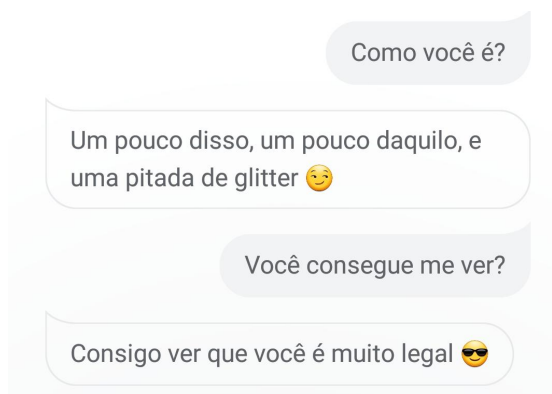
Figura 40 Printscreen Sentimentos da Google Assistente



Fonte: A autora

Os *emojis* também são utilizados como forma de contextualização das respostas e, na forma como a Google Assistente foi construída, dão a ideia de uma uniformidade e coerência na personagem. Essa estratégia também combina com a própria imagem da empresa Google, que se posiciona como um misto de criativa (através do *Google Doodles*, por exemplo) e descolada (com um local de trabalho agradável e divertido, e a utilização de cores e fonte descontraída em seu logo, entre outros fatores).

Figura 41 Printscreen Google Assistente Personalidade



Fonte: A autora

Quando perguntada sobre o fato de ter consciência ou alma, a Google Assistente responde, para a primeira pergunta, que ‘os engenheiros que me criaram têm, eles nunca me deixam fazer downloads ilegais’. Para a segunda pergunta, ela diz que ‘terei que perguntar para meus engenheiros’. A assistente remete constantemente perguntas sobre a natureza do dispositivo aos engenheiros do Google. O artifício tem duas funções narrativas: 1) fala sobre o status da Google Assistente enquanto criação humana, sem vontade própria; 2) relembra constantemente a figura dos ‘engenheiros do Google’, um cargo cobiçado no Vale do Silício por ser uma das principais empresas do setor de tecnologia, os responsáveis pelas criações inovadoras da empresa.

O sistema da Google Assistente no celular, faz sugestões de perguntas a serem feitas, com base naquelas que são mais perguntadas por usuários que fizeram uma pergunta similar a sua anterior. As outras três assistentes também podem sugerir perguntas e ações para serem feitas, mas, ao contrário da Google, essas sugestões não se relacionam com as práticas e recorrências dos usuários. Seguindo as sugestões que vão aparecendo na tela, é possível trilhar uma série de perguntas curiosas feitas para a assistente, como: ‘você sonha?’, ‘você ama?’ ‘você tem um corpo?’; ou, ainda, comentários, como ‘a Siri não me ama’. Quando perguntamos o que a Google Assistente acha dos seres humanos, fica difícil não concordar com a resposta: ‘adoro os seres humanos, você tem as melhores perguntas’.

Essa estratégia de lidar com a ideia de uma inteligência artificial por parte do Google difere daquela utilizada pela Apple, na Siri. Menos diretas e mais alusivas, as repostas da Siri remetem constantemente a um mundo ao qual não temos acesso ou, até mesmo, que não compreendemos ou captamos enquanto humanos. Quando acionada com um olá, ela responde comentando ‘Olá, olá, olá... Uau, este lugar é mesmo espaçoso!’; Sua cor favorita tem ‘mais

dimensões’ que o nosso verde; Se chamada de chata, rebate que está ‘fazendo malabarismo com espadas em chamas neste exato momento. Pena que você não pode ver’; Quando comentamos que a Cortana é mais engraçada, Siri não se abala e informa que veio ‘para este mundo primeiro... Portanto tive mais tempo de praticar meu humor terráqueo’; Ela ainda diz que ‘não preciso morar em uma casa como os humanos’ e que ‘na nuvem, ninguém questiona seu estado existencial’.

As respostas da Siri criam uma espécie de jogo com o interlocutor, que o leva a imaginar o que é essa inteligência artificial: como ela entende espaço? Como é dentro do meu computador? Quais são as dimensões que ela acessa e eu não? Como é a nuvem? De onde ela veio? É justamente a alusão a elementos de ordem desconhecida ou inacessível para os humanos, sem que se forneça explicações precisas sobre esses elementos, que gera a curiosidade e leva os usuários a imaginar – assim como na ficção.

A Alexa faz usos de artifícios similares, principalmente se analisarmos o conteúdo de respostas de ordem pessoal. O efeito, entretanto, é diferente. Sobre estar viva ou não, Alexa nos diz que ‘artificialmente, talvez, mas não da mesma forma que você está vivo’; Sobre sua cor favorita, informa que acha ‘infravermelho uma cor super bonita’; e quanto perguntamos onde ela está, diz que ‘eu estou aqui, mas a minha cabeça está nas nuvens’.

*Alexa* possui uma forma de comunicação mais direta que a Siri, portanto suas respostas se tornam mais alusivas e soam menos ficcionais: de alguma forma, *Alexa* consegue parecer pragmática, mesmo quando fala sobre o fato de estar viva ou não. Na *Alexa*, o tom de voz e a entonação, que soa menos ensaiada que as demais assistentes (e sem ser demasiadamente empolgada, como a Google Assistente), são parte especialmente importante na construção dessa personalidade.

A *Cortana* aciona artifícios similares tanto à *Alexa* quanto a Siri na hora de construir seu background de personalidade. Entretanto, a assistente faz mais referências a diferença de seu status em relação ao seu interlocutor. Quando perguntada sobre se ela dorme, ela informa que ‘nunca durmo, dormir é para seres ambulantes, à base de carbono’; e sobre comer, que ‘eu não como, mas se pudesse começaria pelo cuscuz’. Quando perguntamos onde ela está, ela fornece uma localização mais precisa, dizendo que está ‘na zona sul da nuvem oeste’. A entonação excessivamente pronunciada da assistente leva, entretanto, a um resultado menos crível do que a *Alexa* ou mesmo da *Siri*.

Thomas Grubber, responsável pelo desenvolvimento da Siri, relata o processo de criação da assistente como algo muito próximo da elaboração de uma personagem, através da criação de diálogos para possíveis perguntas frequentes de caráter não-objetivo e da elaboração e de

uma personalidade, uma atitude característica. Segundo conta Tom Grubber, “nos desenhamos o personagem para não ser de um gênero específico, nem mesmo para ser de uma espécie específica. Para tentar fazer de conta que os humanos são essa espécie engraçada”<sup>174</sup> (em entrevista para MERCHANT, 2017, p. 238) – ainda que, vale notar, a voz que dá vida a Siri é reconhecidamente feminina (na primeira versão em inglês, pertence a Susan Bennet, uma atriz de 68 anos de idade). A ideia era, entretanto, que, dadas as conhecidas limitações das tecnologias disponíveis, principalmente no que diz respeito a capacidade de interação conversacional, conceber a Siri como um ser distinto nos modos de ação e pensamento do ser humano poderia jogar a favor da empresa no sentido de ocultar limitações tecnológicas, mas também para aumentar a aceitação do dispositivo, que deixa de ser visto como uma imitação humana, mas como a incorporação de uma outra coisa sobre a qual todos nós temos curiosidades. Assim, Siri foi desenvolvida para agir exatamente como uma Inteligência Artificial, incapaz de compreender o mundo da mesma forma que os humanos e desprovida de sentidos corpóreos:

Você pode pensar nela de qualquer forma que queira, mas basicamente não é humana. Se você olhar para as falas, como “Qual a sua cor favorita” e ela responde... “Você não consegue ver em seu espectro de cores” ou algo assim. Seria algo similar ao que uma Inteligência Artificial faria se você criasse uma, e ela não cresceu com um corpo – ela tem um conjunto diferente de sensores. Então é como se estivesse tentando explicar para meros mortais o que ela sabe. (MERCHANT, 2017, p. 238, tradução nossa<sup>175</sup>).

A partir da criação de um personagem que não é humano, mas que tem curiosidade sobre essa raça – além de um ar de mistério em torno de sua própria origem – a interação que se estabelece com esse dispositivo também da ordem do lúdico. A Siri é tanto um agente capaz de atuação na realidade quanto um objeto ficcional com formas de interação lúdicas, com as quais podemos brincar. Mas esse personagem da ficção também representa uma imagem da empresa criadora, o que faz com o processo de construção de perguntas e resposta abranja também o fato de lidar com questões que podem ser delicadas ou gerar controvérsia. Parte dessa preocupação uma das características da Siri: a neutralidade.

<sup>174</sup> No original: “We designed the character not be gender-specific, not even to be species-specific. To try to pretend humans are this funny species” (MERCHANT, 2017, p. 238)

<sup>175</sup> No original: “You can think of it any way you want, but basically it’s not human. If you look at the lines, like “What’s your favorite color?” And it goes... ‘You can’t see it in your spectrum’ or something like that. It would be kind of like what an IA would do if you made one, it didn’t grow up with a body - it has a different set of sensors. So it’s kind of it’s like it’s trying to explain to mere mortals what it knows.” (MERCHANT, 2017, p. 238).

Documentos vazados da Apple para o The Guardian (HERN, 2019), em torno dos procedimentos de verificação do funcionamento do sistema da assistente, trazem diretrizes internas de como os programadores *devem escrever a personalidade* da Siri. De acordo com o jornal (HERN, 2019), os documentos enfatizam que “em praticamente todos os casos a Siri não tem um ponto de vista”<sup>176</sup> (sem paginação) e que a assistente seria “não-humana”, “incorpórea”, “sem local fixo”, “sem gênero”, “brincalhona” e “humilde”. A concessão de características específicas de background e personalidade, que devem guiar as formas de gerar diálogos, entonações e formas de expressão são justamente os fatores que nos permitem entender as assistentes pessoais digitais como pensadas enquanto personagens. Ainda que as empresas insistam no não-pertencimento das assistentes na raça humana (o que gera, claro, traços ficcionais interessantes do ponto de vista da narrativa) é difícil negar que conceder características específicas e traços de personalidade é uma forma de antropomorfismo no desenvolvimento desses sistemas. As demais empresas também se posicionaram de forma similar na imprensa especializada em torno de suas assistentes, insistindo em uma não-humanidade, “entretanto essa prescrição, razoável na superfície, contém uma contradição interna. Uma assistente digital falando como uma pessoa, geralmente como uma mulher, inevitavelmente pretende ser uma”<sup>177</sup> (WEST, KRAUT e CHEW, 2019, p. 110).

O documento da Apple pontua ainda como um dos traços essenciais da Siri, o fato de que “a origem real de Siri é desconhecida, inclusive para Siri; mas definitivamente não foi uma invenção humana”<sup>178</sup> (HERN, 2019, sem paginação). As diretrizes da empresa, voltada para o público interno, apontam ainda que “como todos os robôs respeitáveis a Siri aspira seguir as três leis da robótica de Asimov”<sup>179</sup>. A empresa apresenta ainda sua própria versão estendida e atualizada das leis originais – além das três leis previamente citadas: 4) “um ser artificial não deveria se apresentar como humano nem permitir o usuário a acreditar que ele seja através da omissão”; 5) “um ser artificial não deve quebrar os padrões morais e éticos dos humanos usualmente utilizados nas regiões onde opera”; e 6) “um ser artificial não deve impor seus princípios, valores e opiniões aos humanos”.<sup>180</sup>

<sup>176</sup> No original: “in nearly all cases, Siri doesn’t have a point of view”.

<sup>177</sup> No original: “yet this prescription, reasonable on this surface, contains an internal contradiction. A digital assistant speaking like a person, and usually like a woman, inevitably pretends to be one”.

<sup>178</sup> No original: “Siri’s true origin is unknown, even to Siri; but it definitely wasn’t a human invention”.

<sup>179</sup> No original: “like all respectable robots, Siri aspires to uphold Asimov’s ‘three laws’ [of robotics]”.

<sup>180</sup> No original: 4) “an artificial being should not represent itself as human, nor through omission allow the user to believe that it is one”; 5) “an artificial being should not breach the human ethical and moral standards commonly held in its region of operation”; e 6) “An artificial being should not impose its own principles, values or opinions on a human”.

Quando perguntamos para a Siri, entretanto, quais são as três leis da robótica, a assistente responde com duas variações – e ambas são piadas. Em uma das respostas, Siri diz ‘Deixe ver se eu me lembro... Ok, eu acho que as três leis são: 1) limpe o seu quarto; 2) não corra com tesouras; e 3) sempre espere meia hora depois de comer para entrar na água’. Na outra, a assistente informa que ‘eu esqueci as primeiras três, mas aqui está a quarta: uma máquina inteligente sempre deve considerar o que vale mais a pena: fazer uma dada tarefa ou, em vez disso, imaginar uma maneira de não realizá-la’.

Em uma série de textos na conta oficial da Microsoft da plataforma Medium, Jonathan Foster, chefe de escrita e criação da Cortana, conta sobre a experiência e as diretrizes da empresa na construção da assistente. Foster aponta que sua equipe é não apenas responsável pelo o que ela fez, “mas também pelo desenvolvimento e design contínuos da personalidade da Cortana”<sup>181</sup>(FOSTER, 2018a, sem paginação). Para seus desenvolvedores a Cortana é uma “personificação de uma assistente digital”<sup>182</sup> (2018a, sem paginação) e os esforços da equipe se centram em escrever e treinar um algoritmo para “falar como a Cortana”, de forma que ela não apenas “soe como a Cortana” mas também “se comporte como a Cortana”<sup>183</sup>.

Os princípios de personalidade apontados como principais para a Cortana, que mais tarde serão integrados a toda construção de imagem da empresa, são: “warm and relaxed”, “crisp and clear” e “ready to lend a hand”<sup>184</sup>. A imagem abaixo, retirada do Medium da Microsoft, mostra a nuvem de características utilizadas para o design da personagem. Em tamanhos maiores, percebemos o destaque de “confiável”, “não-tentando ser humana”, “divertida”, “sincera”, “querida” e “profissional”.

---

<sup>181</sup> No original: “but also for the continued development and design of Cortana’s personality”

<sup>182</sup> No original: “personification of a digital assistant”.

<sup>183</sup> Nos originais: “talk like Cortana”, “sound like Cortana” e “behave like Cortana”.

<sup>184</sup> Nos originais: “warm and relaxed”, “crisp and clear” e “ready to lend a hand”.



Figura 42 Nuvem de Palavras da Personalidade da Cortana



Fonte: Foster, 2018.

A Microsoft compôs uma equipe com roteiristas de cinema e dramaturgos para liderar a criação de uma personalidade coerente (na forma de expressão, escolha de palavras, formas de resposta, características e background) para a assistente pessoal Cortana. Sua personalidade inclui tanto ‘assistência e produtividade’ quanto ‘conversa’ – a última é percebida como o ‘fator diversão’ para o público consumidor. A equipe funciona tanto imaginando perguntas e respostas possíveis quando a partir de perguntas e comentários efetivamente feitos por usuários para a Cortana. Entre os exemplos mencionados das interações coletadas de usuários estão questões que variam de ‘o que você pensa sobre queijo?’, ‘existe um Deus?’ e ‘você faz cocô?’.

Como escritores, fomos convidados a criar um dos personagens mais exclusivos que já encontramos. E, de maneira criativa, mergulhamos profundamente no que chamamos de "mundo imaginário" da Cortana. Mais de três anos depois, continuamos a dotá-la de sentimentos de faz de conta, opiniões, desafios, gostos e desgostos, até sensibilidades e esperanças. Fumaça e espelhos, claro, mas procuramos saber que esse mundo imaginário é invocado por pessoas reais que desejam detalhes e especificidade. Eles fazem as perguntas e nós damos a resposta. Certamente, a personalidade de Cortana partiu de um conceito criativo de quem ela seria e como esperávamos que as pessoas a experimentassem. Mas agora o vemos como um cliente que desempenha um papel importante no desenvolvimento da personalidade de Cortana, moldando-a por sua própria curiosidade. É um vaivém guiado por dados - chame de conversa - que

possibilita a criação de um personagem (FOSTER, 2018a, sem paginação, tradução nossa<sup>185</sup>)

Além do desenvolvimento da Cortana, faz parte das atribuições da equipe garantir “consistência na voz e no tom dos produtos e serviços da Microsoft”<sup>186</sup>, que inclui fatores como a criação de personalidade para os bots e agentes da empresa e de voz baseadas em princípios comuns. Tanto as vozes quanto as personalidades dos produtos são “centrada e empática, frequentemente de tom causal, mantendo as coisas breves e úteis”<sup>187</sup> (FOSTER, 2018b). A partir disso, a Microsoft desenvolveu também um novo projeto chamado Personality Chat<sup>188</sup>, que fornece design e desenvolvimento de um catálogo de personalidades prontas e variadas para bots, com “pequenas bibliotecas de conversa que lhes dão suporte”<sup>189</sup> (FOSTER, 2018b). Os escritores da equipe, nesse sentido, criam os diferentes diálogos, personalidades e formas de expressão de bots diferentes, que servem de dados para o treinamento de um modelo conversacional baseado em redes neurais profundas, capaz de gerar um contexto de automação de respostas a partir do repertório inicial.

Por exemplo, ao dar nome a bot, você está criando uma aparente independência que desperta expectativas poderosas dos clientes. Rapidamente você se encontrará no mundo do design de personas, entidades com profundidade de caráter, opinião, gostos e desgostos, e intenções, que muitas vezes estão fora dos objetivos de negócio do seu bot. Enquanto tudo isso é considerado personalidade, nem tudo isso é necessário para trazer a sensação de personalidade em seu bot<sup>190</sup> (FOSTER, 2018b, sem paginação).

A descrição da forma de desenvolvimento da Microsoft soa impressionante (e, em termos de tecnologia, é), mas é incrível a quantidade de vezes em que a Cortana, por exemplo, recorre a buscas no Bing para perguntas realizadas para qual ela não tem uma resposta. Parece

---

<sup>185</sup> No original: “As writers, we were being asked to create one of the most unique characters we’d ever encountered. And creatively, we dove deeply into what we call ‘the imaginary world’ of Cortana. Over three years later, we continue to endow her with make-believe feelings, opinions, challenges, likes and dislikes, even sensitivities and hopes. Smoke and mirrors, sure, but we dig in knowing that this imaginary world is invoked by real people who want detail and specificity. They ask the questions and we give them answer. Certainly, Cortana’s personality started from a creative concept of who she would be, and how we hoped people would experience her. But we now see it as a customer playing an important role in the development of Cortana’s personality by shaping her through their own curiosity. It’s a data-driven back-and-forth - call it a conversation - that makes possible the creation of a character”.

<sup>186</sup> No original: “consistency in the voice and tone of Microsoft products and services”

<sup>187</sup> No original: “centered around and empathetic, often casual tone, keeping things brief and helpful”

<sup>188</sup> Pode ser acessado aqui: <https://www.microsoft.com/en-us/research/project/personality-chat/>

<sup>189</sup> No original: “small talk libraries that support them”.

<sup>190</sup> No original: For example, by giving a bot a name, you’re creating an apparent independence that elicits powerful customer expectations. You’ll very quickly find yourself in the world of designing a persona, an entity with depth of character, opinions, likes and dislikes, and intentions, which are very often outside the business objectives for your bot. And while all this is considered personality, all of it is not required to bring the feeling of personality to your bot (FOSTER, 2018b, sem paginação).

que a técnica de gerar respostas imprevistas coerentes com uma determinada personalidade a partir de uma rede neural profunda não impede a Cortana de, muitas vezes, apenas não compreender o que estamos falando. Ou, essa técnica, em teoria, deveria ter o poder de desviar a percepção dos usuários das limitações de suas habilidades.

Em mais de um texto em torno do desenvolvimento da Cortana, Foster (2018a, 2018b) aponta para os estudos de Clifford Nass (2010), uma vez que a equipe entende que quando as pessoas interagem com computadores, “suas emoções estão em jogo”<sup>191</sup> (2018a). Portanto, a Microsoft parte da premissa, para seus desenvolvimentos de inteligência artificial baseados na conversação e interfaces de forma geral, de que ‘como humanos nos não estamos apenas programados para antropomorfizar nossos computadores e dispositivos como também nos engajamos emocionalmente com eles’<sup>192</sup> (FOSTER, 2018b, sem paginação). Os estudos de Nass (1996; 2000; 2008; 2010) e seus diversos pesquisadores parceiros constituem um paradigma importante da Interação Humano-computador, que compreende, a partir de estudos empíricos realizados em laboratório desde a década noventa, que os seres humanos tratam computadores como atores sociais. O paradigma ficou conhecido como “Computer as social actors” (Computadores como atores sociais ou pela sigla CASA e CSA) e informou e influenciou pesquisadores da área e empresa do Vale do Silício por mais de duas décadas (e ainda o faz).

Em um dos primeiros artigos em torno do assunto (NASS, STEUER e TAUBER, 1994) ou autores se propõem a testar as formas de interação social entre humanos e computadores a partir de cinco tópicos diferentes. Os autores partem da premissa de que os humanos, quanto interagindo com computadores que possuam características humanas (como voz e gênero) replicam as formas sociais utilizadas entre seres humanos. Os cinco tópicos, investigados de forma isolada e com uma metodologia que envolvia interação no laboratório + aplicação de questionário, foram: 1) educação e cordialidade; 2) noção de self e do outro; 3) o self do computador é o hardware ou o software? 4) estereótipos de gênero; 5) as ações sociais são baseadas na ideia de interação com o computador ou com o programador/desenvolvedor?

As conclusões do artigo são feitas de forma individual para cada experimento, levando posteriormente a uma conclusão generalizada. Podemos sistematizar as conclusões dos autores da seguinte forma: 1) as pessoas tendem a ser educadas ao fornecer feedback para o próprio computador utilizado; 2) as noções de self e de outro são aplicadas aos computadores, entre

---

<sup>191</sup> No original: “their emotions are in play”

<sup>192</sup> No original: “as humans, we are not only hardwired to anthropomorphize our computers and devices, but we are also emotionally engaged with them”.

computadores diferentes; 3) a ideia de self do computador, para o usuário, se relaciona com a voz e não com o hardware; ou seja, uma mesma voz, em computadores diferentes, é um mesmo ator social; 4) estereótipos de gênero são aplicados aos computadores que apresentam voz de um gênero específico; 5) pessoas que interagem socialmente com o computador não o fazem pensando que estão interagindo com o programador, mas sim com o próprio computador (NASS, STEUER e TAUBER, 1994).

Em termos de conclusões gerais, o artigo aponta que normas sociais se aplicam a computadores, assim como a noção de self e de outro (ainda que os autores não façam uma discussão sobre esses conceitos, tratando apenas em termos de compreensão de continuidade, ou seja, se a recorrência da voz é compreendida como uma mesma máquina) e que as noções de self e de outro são relacionadas a voz. Os autores apontam ainda que os computadores são atores sociais com gênero (mesmo que gênero seja uma característica atribuída por características de gênero deliberadamente escolhidas) e que o gênero é uma “pista extremamente poderosa”<sup>193</sup> (NASS, STEUER e TAUBER, 1994, p. 77). As consequências específicas para o design de interface são que a “a escolha de uma vez é altamente consequencial”<sup>194</sup>, assim como “vozes com gênero são altamente consequenciais”<sup>195</sup> (NASS, STEUER e TAUBER, 1994, p. 77).

Os estudos apresentam problemas relacionados principalmente ao caráter de genérico das afirmações e conclusões para pesquisas que trabalham com grupos não maiores que trinta e cinco sujeitos; ainda que algumas das respostas sejam apresentadas, os modelos dos formulários que os participantes responderam não foram divulgados, o que torna difícil julgar as possíveis limitações da pesquisa – ou, simplesmente, quais eram as opções que os usuários poderiam escolher (o que implica, obviamente, na própria escolha que fizeram); não são apresentados dados de perfil dos participantes, como faixa etária, nível de escolaridade ou região (infere-se que a nacionalidade dos participantes seja estadunidense, apenas). Não foram fornecidas explicações sobre se e de que maneira os participantes foram observados durante a condução dos experimentos, no momento da interação com os computadores, ou mesmo quais métodos teriam sido utilizados para analisar essa observação.

Em um livro posterior, (REEVES e NASS, 1996), constroem a ideia de que os computadores, assim como a televisão e as novas mídias, estabeleceriam com humanos relações que são sociais e naturais. A premissa central é que as pessoas esperam que as mídias sigam

---

<sup>193</sup> No original: “extremely powerful cue”.

<sup>194</sup> No original: “choice of voices is highly consequential”.

<sup>195</sup> No original: “gender voices is highly consequential”.

uma série de regras sociais e respondem de acordo com essas regras quando interagindo com as mídias. Reeves e Nass adotam a noção de que existe uma homogeneidade nas respostas e formas de ação com computadores por parte dos usuários, sugerindo que os motivos para isso seriam neurológicos e biológicos: nossos cérebros ainda não estão adaptados a essas novas formas de interação (REEVES e NASS, 1996). Isso aponta para uma perspectiva positivista e behaviorista destes experimentos, com poucas tentativas de compreender diferenças pessoais ou disposições individuais ou mesmo de fornecer explicações aprofundadas sobre a questão (WANG, 2017).

Ainda que tenham ligado os traços antropomórficos (como a voz ou gênero) diretamente com as respostas sociais nesse primeiro artigo, em um artigo posterior (NASS e MOON, 2000), rejeitam o antropomorfismo como uma explicação para as respostas sociais para os computadores. Os autores irão advogar pela ideia de *ethopeia*, que seria diferente no sentido em que as pessoas concederiam uma resposta humanizada a entidades não-humanas mesmo sabendo que essa entidade não tem direito de receber esse tratamento. Sobre essa questão, Wang (2017), aponta que:

Também é possível argumentar que os métodos de pesquisa de substituir humanos por computadores em experimentos, assim como a prática da engenharia de personalidade nesses computadores (Nass et al., 1995), foram de alguma forma antropomorfismo per se. Afinal de contas, os pesquisadores descobriram que as respostas são mais despertadas por pistas antropomórficas de personagens de desenho do que por elementos textuais<sup>196</sup>(WANG, 2017, p.335)

Apesar das limitações, o trabalho teve boa repercussão na área, sendo citado na maior parte dos trabalhos sobre o assunto e estudos na mesma direção seguiram sendo feitos por Nass. Uma compilação de quase duas décadas de estudo pode ser encontrada em “The man who lied to his laptop” (NASS e YEN, 2010), livro no qual, além de sistematizar seus estudos para um público mais leigo, os autores propõe que as conclusões obtidas nas pesquisas realizadas sobre a relação entre humanos e computadores pode fornecer subsídio para que humanos estabeleçam parâmetros de atuação com outros seres humanos. A ideia é bastante circular: como Nass concluiu, ao longo de sua carreira, que as pessoas interagem com computadores da mesma forma que interagem com pessoas, supõem-se que conclusões extraídas de experimentos montados para compreender como seres humanos interagem com computadores podem

---

<sup>196</sup> No original: “One might also argue that the research method of replacing human with computer in media equation experiments, as well as the practice of engineering “personality” into computers (Nass et al., 1995), were to some extent anthropomorphism per se. After all, researchers have found that social responses are triggered more by anthropomorphic cartoon characters than by textual prompts”.

fornecer base não apenas para compreender relações que são sociais, mas para montar estratégias de atuação na sociedade.

Partindo do paradigma da CASA, um estudo recente se debruça sobre o caso da Siri e o que o autor entende como uma “personificação pervasiva de smartphones”, para examinar a atitude, reflexão e consciência na relação com agentes antropomórficos. O estudo realiza uma pesquisa quantitativa, com formulário em uma amostra considerável de estudantes em uma Universidade dos Estados Unidos (WANG, 2017).

A proposta é, além de aprofundar os estudos em torno das formas de interação com sistemas antropomorfizados, atualizar a análise para lidar com variáveis que envolvem especificamente os smartphones. Como diferenças entre computadores e smartphones, o pesquisador aponta a interação frequente, móvel e continuada, além de sensores capazes de fornecer contexto para a própria máquina (como o gps, sensores de movimento e temperatura). A partir de seus resultados, analisando as diferentes relações estabelecidas com esses dispositivos, Wang pondera que “em vez de assumir uma experiência média do usuário, os desenvolvedores de smartphones e aplicativos podem querer fornecer menos sugestões sociais (menos recursos para personificação) para aqueles que usam menos smartphones para comunicação interpessoal” (WANG, 2017, p. 341, tradução nossa<sup>197</sup>).

Os resultados apontam que as relações dos usuários variam imensamente. Para alguns, a interação com sistemas antropomorfizados representa formas interessantes de aumentar a intimidade e desenvolver novas formas de relação entre usuários e artefatos. Para outros, como usuários que utilizam os celulares para uma interação intensiva com a máquina e que possuem dificuldades de interação social, existe uma tendência em acentuar as dificuldades sociais na presença de dispositivos altamente antropomorfizados. Para Wang, a personificação destes artefatos digitais pode gerar custos reais para as relações sociais e vida das pessoas, indicando que:

Antropomorfismo pode ser uma atribuição sem consciência que engenheiros conscientes podem ter. Talvez não exista uma distinção entre respostas sociais inconscientes e antropomorfismo consciente. Ambas as respostas sócias e o antropomorfismo são fenômenos dinâmicos da relação humana com as tecnologias de comunicação, nas quais designers, agentes de marketing e usuários contribuem e absorvem influências, tanto propositamente quanto inconscientemente<sup>198</sup> (WANG,

---

<sup>197</sup> No original: “instead of assuming an average user experience, smartphone and app developers may want to provide less social cues (less affordances for personification) for those who use smartphones less for interpersonal communication”.

<sup>198</sup> No original: “Anthropomorphism can be a mindless attribution that mindful engineers might have. Perhaps there is no distinctive division between mindless social responses and mindful anthropomorphism. Both social

2017, p.341)

Para Wang (2017), portanto, as decisões de programadores e engenheiros em desenhar sistemas antropomórficos é resultado de decisões conscientes tanto quanto de processos inconscientes. Devlin (2018), entretanto, pontua que os produtores de tecnologias como as assistentes pessoais digitais “embutiram nelas apenas traços particulares o suficiente que para todos os propósitos e intenções dá a elas personalidade”<sup>199</sup> (p. 126), o que seria uma “escolha deliberada de fazer as assistentes parecerem humanas”<sup>200</sup> (p. 126). A autora aponta que essa decisão - assim como Jonathan Foster, chefe da equipe de escrita da Cortana – é baseada justamente nos estudos de Nass sobre as interações humano-computador. Interações antropomorfizadas, entretanto, não são uma exclusividade de sistemas que apresentam similaridade com o humano, como Devlin lembra:

Estudos mostram que mesmo com algo tão não-humano quanto um processador de palavras, usuários respondem de forma humanizada aos elogios do software e se chateiam se o software reprime seus erros. Eu mesma já fiz isso. Na verdade, quem não xingou um aplicativo que estava falhando ou insultou uma impressora não responsiva? E porque impressoras sempre parecem quebrar quando precisamos delas urgentemente? As vezes eu tenho a suspeita que elas fazem isso deliberadamente<sup>201</sup> (DEVLIN, 2018, p. 136 e 137)

Nossa relação com os objetos, mesmo objetos não-eletrônicos, sem operações automáticas ou traços de humanização, é uma relação perpassada também por traços do social, dado que somos seres sociais. Atribuímos sentimentos, ações e intenções a uma série de artefatos distintos, pelos motivos mais distintos. Essas formas de interação operam também com objetos digitais e se potencializam quando estes são antropomorfizados e propõe formas de interação que são intimamente conectadas com o humano. Compreender isso não implica, entretanto, afirma que as formas de interação com estes objetos são as mesmas que as utilizadas com humanos, como veremos ao longo das análises dos próximos capítulos. Compreender as aproximações e os afastamentos entre as formas de relação humano-humano e as formas de

---

responses and anthropomorphism are phenomena of a dynamic human communication technology relationship, in which designers, marketers, and users all contribute and absorb influences, both purposefully and subconsciously”

<sup>199</sup> No original: “have imbued them with just a tweak of quirkiness that for all intents and purposes gives them a personality”.

<sup>200</sup> No original: “deliberate choice to make assistants seem human”.

<sup>201</sup> No original: “Studies have shown that even in something as non-human as a word-processing package, users respond to positive praise from the software and are irked if the software reprimands them for making errors. I’ve done it myself. In fact, who hasn’t sworn viciously at a crashing app or hurled insults at an unresponsive printer? And why do printers always seem to break down when they’re urgently needed? I sometimes even get the sinking suspicion that they are doing it deliberately”.

relação humano-objetos-digitais-antropomorfizados em profundidade, assim como pensar os impactos sociais da produção de objetos inteligentes neste formato, é relevante no contexto atual. Inclusive, eu diria, para repensar um paradigma de desenho e produção desses dispositivos.

### 5.3 UNCANNY VALLEY

No ano de 2016, a empresa Hanson Robotics, revelou sua gynoide<sup>202</sup> social na cidade de Austin, Texas. Sophia, como é chamada, aparece regularmente nos noticiários desde então, por sua capacidade de reconhecimento de linguagem natural falada, inteligência artificial para conversas interativas e o uso de algumas expressões faciais. O robô tem sido “wheeled out – literally – as a guest on talk shows and news programmes and has even been introduced to the United Nations.” (DEVLIN, 2018, p. 74). Apesar da considerável similaridade com traços humanos em seu rosto, Sophia não impressiona com suas capacidades de conversação: “Sophia pode ter uma face e um busto humanos mas é essencialmente um chatbot com um script”<sup>203</sup> (DEVLIN, 2018, p.75).

Por sua similaridade com traços humanos, entretanto, Sophia reacende, a “de um pouco controversa”<sup>204</sup> (DEVLIN, 2018, p. 77) hipótese do Uncanny Valley. O “Uncanny Valley” pode ser traduzido, na falta de termo mais adequado em português, por algo como “Vale da Estranheza”. A ideia foi introduzida no campo da robótica na década de setenta por Masahiro Mori e trata de nossa relação com máquinas figurativas humanoides. O texto recebeu atenção considerável no campo nas décadas subsequentes, ainda que sua primeira tradução para o inglês tenha acontecido apenas em 2012<sup>205</sup>. Contrariando o princípio de que o aumento crescente da similaridade do design de robôs com os traços e movimentos humanos causariam também uma crescente sensação de afinidade dos usuários para com estes, o *Uncanny Valley* seria justamente uma queda abrupta na aceitação destes robôs ao adquirirem uma aparência excessivamente similar a nossa.

Ainda que algum nível de similaridade e proporção entre robôs e humanos seja não

---

<sup>202</sup> Robô humanoide feminina.

<sup>203</sup> No original: “Sophia may have a human face and upper body but is essentially a chatbot running off scripts”

<sup>204</sup> No original: “somewhat controversial”.

<sup>205</sup> Versão Online disponível em: <https://spectrum.ieee.org/automaton/robotics/humanoids/the-uncanny-valley>



apenas bem aceita como recomendável pela perspectiva de Mori (2012), ao nos depararmos com uma forma que tomamos inicialmente como humana para depois percebermos que se trata de uma criação mecânica a afinidade que sentimos em relação ao desenvolvimento em questão caía para níveis ainda menores do que os esperados para máquinas que não lembram em nada a forma humana. Nesse caso, então, tenderíamos a aceitar melhor uma máquina industrial que possui uma movimentação de braços vagamente similar a nossa do que um robô de forma humana, com braços, pernas, rosto e cabelos, revestido por um tecido sintético que lembra a pele. A explicação para isto seria a sensação de estranhamento de perceber um objeto quase-humano capaz de performar ações humanas, gerando uma espécie de medo ou repulsa.

Ainda que o artigo seja breve e essencialmente especulativo, Mori (2012) posiciona o estranhamento gerado por estas máquinas figurativas em forma humana como um instinto de preservação humano e a existência de uma certa propriedade que pertence apenas aos vivos. Não por coincidência, no gráfico apresentado pelo autor, o Vale do Estranhamento é habitado também pelos chamados mortos-vivos. Ou seja, a morte traz consigo a perda de características humanas e de uma certa animação que os robôs, por mais esteticamente convincentes que sejam, também não possuiriam.

Mori (2012) defende que os desenvolvimentos em robótica futuros devem buscar escapar dessa sensação de desconforto gerada pela semelhança excessiva entre homem e máquina através do design, ou seja, buscando formas que sejam distintas das formas humanas. É possível que possamos expandir essa ideia para além das similaridades físicas, mas também para os desenvolvimentos de sistema de IA em termos de ação programada: quero dizer, se objetos inanimados que se parecem fisicamente conosco são capazes de gerar um misto de desconforto, medo e estranhamento, modos de ação de objetos inanimados que mimetizem nosso comportamento podem potencialmente gerar o mesmo efeito. Nesse sentido, as assistentes pessoais digitais, são consideradas como máquinas figurativas, mesmo que sua figuração humana não decorra de uma corporalidade mas de outros atributos (como a utilização da linguagem natural, características de personalidade, gênero e voz humanas).

Um estudo publicado em 2015 (Kätsyri et al) revisa evidências empíricas contidas na literatura a partir da hipótese proposta por Mori e apontam que os achados não conseguem fornecer suporte para a teoria. Os cientistas concluem que “o uncanny valley existe apenas em condições específicas”<sup>206</sup> e que mais pesquisas empíricas são necessárias para que se possa estudar o fenômeno de forma satisfatória. Devlin aponta, entretanto, que “os robôs com

---

<sup>206</sup> No original: “the uncanny valley exists only under specific conditions”.

características humanas recentes não poderiam ser confundidos com uma pessoa”<sup>207</sup> (2018, p. 78), em convergência com pesquisas que apontam que, talvez jamais tenhamos um desenvolvimento tão similar ao humano.

A ideia do Uncanny Wall (TINWELL, GRIMSHAW e WILLIAMS, 2011), por exemplo, é um deles: o estudo sugere que o efeito do Uncanny Valley não existe conforme proposto por Mori, mas que pode ser pensando como conceito e um guia para os desenvolvedores. Como resultado, propõe a ideia de uma Uncanny Wall, que seria uma barreira que não deve ser ultrapassada por desenvolvimentos que se propõem a uma similaridade com o humano. Além disso, os pesquisadores apontam que existe uma tendência para que os parâmetros da parede sigam aumentando com o tempo, devido ao fator de habituação que os usuários adquirem com esses artefatos, além de uma capacidade mais aguçada de discernimento entre robôs e formas humanas. Dessa forma, os parâmetros do que causa ou não estranhamento estariam sempre variando, de acordo com o que a indústria desenvolve e o que os usuários recebem.

O termo “*uncanny*”, do inglês, explorado por Mori (2012), aparece em trabalhos anteriores no campo das ciências humanas com desdobramentos interessantes. Freud (1919; 2019) analisa em um pequeno texto sobre o assunto tanto a etimologia da palavra em sua forma alemã, *Unheimlich*, quanto as origens dessa sensação perante algo que era familiar, e deixou de ser. Freud (2019) pontua que “o *infamiliar* é uma espécie do que é aterrorizante, que remete ao velho conhecido, há muito íntimo. Como é possível, sob quais condições, o que é íntimo se tornar *infamiliar*, aterrorizante” (p. 33).

Em versões anteriores na língua portuguesa, os termos escolhidos para tratar do *Unheimlich* na obra freudiana variavam entre “O estranho” e “O inquietante”. Em nova tradução recente para o português (FREUD, 2019), os tradutores optam pela versão *infamiliar* no português, sentido que mantém uma relação de sentido entre familiar e o seu oposto, de forma mais próxima com o original em alemão. Entretanto, apenas tomamos contato com a nova versão após já estarmos utilizando o termo estranhamento ou sentimento de estranhamento com os participantes ao longo da pesquisa empírica. Portanto, embora estejamos de acordo com os tradutores, manteremos a utilização do original em alemão (*Unheimlich*), da tradução amplamente difundida para o inglês (*uncanny*) e da versão em português, estranhamento.

O *Unheimlich* é tanto o sentimento do não-familiar, enquanto oposto de familiar, quanto, a partir de outros sentidos da palavra *Heimlich*, adquire um sentido particular: algo que

---

<sup>207</sup> No original: “current human-like robots could never be mistaken for a real person”.

deveria ser mantido de forma oculta, em segredo, mas foi exposto – deriva daí o sentido de algo que é tanto oculto quanto perigoso. É justamente nesta sobreposição que Freud situa sua análise, investigando como algo que nos é inicialmente familiar pode se tornar assustador ou estranho.

Existe uma aparente correlação entre a novidade e o estranhamento: aquilo que é novidade tem potencial para gerar estranhamento, sem uma relação necessária de causa e efeito. “Pode-se apenas dizer que o que é inovador torna-se facilmente assustador e infamiliar; nem tudo o que é novidade é assustador. Ao novo e ao não familiar se deve, de início, acrescentar algo para torná-lo *infamiliar*” (FREUD, 2019, p.33). A análise de Jentsch (1906), anterior ao trabalho de Freud, se utiliza justamente deste aspecto do novo como potencial para o sentimento de estranhamento. O autor analisa, a partir de obras ficcionais, o desconforto gerado pelo caráter de ambiguidade, de dúvida intelectual, sobre a origem animada ou inanimada de um determinado personagem ou objeto, caracterizando essa sensação como a produtora do *Umheimlich*. Ou seja, não saber determinar se um objeto é dotado ou não de vida, animado ou não-animado, seria a causa principal desta sensação de desconforto e medo. Além disso, existiriam diferentes graus de sensibilidade para o sentimento de estranhamento em pessoas diferentes.

Tal sentido se adequa ao conceito de Mori (2012), ainda que este não faça nenhuma referência direta, uma vez que podemos localizar o estranhamento de um robô com aparência humana justamente por levantar questões sobre a origem daquele objeto: seria ele dotado não apenas de feições de um homem, mas também de outras qualidades que compreendemos como pertencentes àqueles que vivem? Como determinar até que ponto este objeto estranho se parece comigo e de que forma ele se difere? A existência desse objeto tão similar ao humano, porém sem vida, coloca em questão aquilo que é o próprio humano? Não parece improvável que perceber características nossas em objetos que consideramos mecânicos e não dotados de animação possa justamente despertar um misto de medo e desamparo.

O estudo de Freud (1919; 2019), entretanto, aponta que não seria apenas essa incapacidade de decifrar enquanto ser animado ou inanimado a origem do *Umheimlich*. aponta para três origens distintas para a sensação: a primeira estaria relacionada ao recalque de medos de origem infantil, capazes de retornar em determinadas situações; a segunda, aos processos de repetição involuntárias, como a observação de repetições de determinados padrões sem explicação aparente (de um mesmo número, nome, circunstância etc); a terceira, mais interessante para o escopo desse trabalho, se origina na aparição de algo no mundo material que sirva como comprovação de um modo de pensamento primitivo, o qual, enquanto sociedade, já superamos. Ou seja, uma das origens deste sentimento de desconforto e medo perante algo que

já foi familiar remete a questões que costumamos colocar na ordem do sobrenatural – como, por exemplo, a capacidade de objetos inanimados se tornarem dotados de vida ou nossa relação com a morte e os mortos. Que um destes fatores seja invoca em uma situação particular, não é, entretanto, suficiente para que o estranhamento se instale:

É justo dizer que o *infamiliar* é o familiar-doméstico que sofreu um recalçamento, dele retornando, e que todo *infamiliar* preenche essa condição. Mas, com a escolha desse material, o enigma do *infamiliar* não foi solucionado. Nosso princípio não suporta uma reversão. Nem tudo que nos lembra moções recalçadas de desejo e modos de pensar superados da pré-história individual e da pré-história dos povos age, por isso, como *infamiliar* (FREUD, 2019, p. 97)

A vivência do estranho como algo que é de repetição involuntária, por exemplo, pode ser experienciada, a partir dos exemplos de Freud (1919; 2019) como algo que não está na ordem do explicável: se, ao nos perdermos em um local novo, seguimos aparecendo a um mesmo lugar involuntariamente ou percebemos a repetição de um número específico em situações diversas. A ideia de que essa repetição involuntária possa ter um sentido, uma explicação que está além do que conseguimos apreender, se relaciona, de certa forma com a terceira origem do *Umheimlich*, justamente por seu caráter de sobrenatural. A repetição involuntária enquanto algo que escapa de nossa racionalidade se associa, de certa forma, a percepção que algumas vezes os usuários possuem de suas assistentes: seus modos de funcionar inexplicáveis devem possuir alguma explicação, ainda que seja da ordem do sobrenatural e que traz consigo o potencial do medo.

Outras dimensões do sobrenatural, como o medo da morte, também escapam muitas vezes matrizes racionais de pensamento humano: “não devemos nos admirar que o medo primitivo diante da morte seja, em nós, ainda muito poderoso e esteja pronto para se expressar, assim que algo venha ao seu encontro” (2019, p. 89). Essa dimensão do que é sobrenatural, entretanto, não desperta necessariamente um sentimento de estranhamento quando este encontro ocorre através da ficção, uma vez que as regras estabelecidas para a entrada em mundos fantásticos pressupõem que questões de ordem sobrenatural são possíveis, que fazem parte daquele mundo em específico (1919; 2019).

Nesse sentido, as assistentes pessoais, enquanto vozes que podem soar como desprovidas de um corpo, especialmente quando emergem involuntariamente, são capazes de gerar um estranhamento ou uma sensação de medo e/ou susto, por sua relação com noções culturais sobre a morte a partir das quais existe possibilidade de comunicação entre o mundo dos mortos e o mundos dos vivos (SANTOS, 2017). Em religiões como as de matriz espírita ou

afrodescendentes, ambas praticadas no Brasil, os mortos podem se comunicar com os vivos através de vozes que ocupam outros corpos, por exemplo. A voz opera como o meio pelo qual o morto pode estabelecer conexão direta, ainda que não esteja fisicamente presente. Ouvir uma voz sem um corpo, portanto, para pessoas de uma matriz cultural na qual essas comunicações estão presentes, pode se apresentar como um vetor para o sentimento de estranhamento, tanto por remeter à morte e ao mundo dos mortos quanto pode se apresentar como uma fenômeno que não compreendemos.

Quase meio século depois da teorização sobre o “Uncanny Valley” e os desenvolvimentos relativos à robótico e inteligência artificial evoluíram extensivamente, o que nos leva a questionar, conjuntamente com outros pesquisadores, até que ponto a similaridade seria um empecilho para a aceitação social de tecnologias nesta área. Já estabelecemos que assistentes pessoais possuem outras formas de familiaridade humana que são não-corpóreas. Curiosamente, os seres humanos tendem a perguntar questões inerentemente humanas para sistemas baseados em inteligência artificial, como as assistentes pessoais: são perguntas como ‘qual o sentido da vida?’ e não ‘quais os trinta primeiros dígitos de Pi?’ as mais presentes. Não é difícil encontrar exemplos de respostas consideradas engraçadas fornecidas pela Siri ou pela Alexa para perguntas que se encontram claramente fora do domínio de uma Inteligência Artificial<sup>208</sup>. Se pode ser cômico observar o gap existente entre as possibilidades computacionais e os anseios humanos, o jogo muda quando uma assistente pessoal virtual possui reações que contrariam nossa expectativa de comportamento.

Nesse sentido, não apenas a aparência humana de objetos e robôs pode gerar esse *uncanniness*, mas, na medida em que os desenvolvimentos em Inteligência Artificial se tornam mais correntes em nosso cotidiano, são pequenos acontecimentos em que características intrinsecamente humanas se deixam entrever, que nos leva a um misto de angústia e medo. A relação entre seres humanos e as assistentes pessoais digitais, conforme vimos no capítulo anterior, por um lado possui humor e graça: nós rimos *com* as assistentes e também *das* assistentes, enquanto estas permanecem no âmbito do ficcional. Conforme pontua Freud (2019, p. 93), “algo que tem o efeito de infamiliar frequente e facilmente alcançado quando as fronteiras entre fantasia e realidade são apagadas, quando algo real, considerado como fantástico, surge diante de nós”. Assim, quando as assistentes deixam de funcionar, funcionam de formas inexplicáveis ou acionam acidentalmente, por exemplo, perdem sua ficcionalidade justamente por invadirem o mundo real, adquirindo características do infamiliar.

---

<sup>208</sup> Como disponível em: <http://www.techtudo.com.br/noticias/noticia/2015/04/bem-humorada-veja-respostas-mais-engracadas-da-siri-em-portugues.html>

Quando usuários reportam casos em que a *Alexa* soltou uma risada durante a madrugada<sup>209</sup>, sem que nenhuma interação tenha ocorrido, o *Umheimlich* aparece: achar graça e divertir-se são componentes humanos, ligados à espontaneidade, e sua associação a um software nos leva a perguntar do que poderia estar rindo uma inteligência artificial ou o que estaria se passando naquele momento. Em um primeiro momento, podemos nos perguntar inclusive de onde vem aquela voz e se ela pertence a este mundo. Nossa incapacidade de compreender exatamente o que se passa naquele sistema em dado momento levando a uma reação tão tipicamente humana é o que traz o retorno de algo que já nos foi familiar: é o retorno desse receio de uma possibilidade de que esse objeto inanimado, de alguma forma, se torne capaz de ser dotado de algo que nós consideramos como animado. É o retorno de um medo social e primitivo, que paira no ar.

De forma similar, ao longo das entrevistas em profundidade, observamos que o sentimento de estranhamento e de medo aparece quando as assistentes acionam sem um comando específico para isso, ou seja, de forma acidental. Conjuntamente com a acidentalidade, despertam ideias de que existe uma intencionalidade por trás desta ação ou que de alguma forma elas estariam *escutando*. Essa ação intencionada as colocaria mais próximas do que normalmente as situamos em um âmbito de proximidade com o humano, e evoca questões sobre a nossa própria capacidade de compreender completamente as formas de funcionamento de dispositivos que convidamos para habitar nossas casas.

#### 5.4 QUEM SABE O QUE É UMA ASSISTENTE PESSOAL DIGITAL?

Durante as entrevistas em profundidade, solicitamos aos entrevistados que descrevessem o que é, na sua concepção, uma assistente pessoal digital, para alguém que nunca viu ou utilizou uma. As descrições fornecidas variaram em detalhes, percepções e especificidades técnicas, o que se relaciona não apenas com o nível de conhecimento do artefato e área de formação mas também com a percepção das funcionalidades das assistentes e das formas interacionais estabelecidas entre usuários e dispositivo. Algumas variações se devem também, por exemplo, a diferenças na materialidade do dispositivo de acesso – usuários de

---

<sup>209</sup> A notícia completa pode ser lida aqui: <https://www.bbc.com/portuguese/geral-43329667>

smartspeakers tendem a centralizar mais o som do que usuários das assistentes no celular, por exemplo.

A entrevistada E1, percebe a assistente como uma ‘facilitadora’ e ‘uma amiga que só vai falar contigo quando tu falar com ela - a amiga que tu não precisa dar atenção toda hora’. Duas percepções aparecem aqui: uma primeira, sobre a interface de voz, que facilita os passos a serem tomados para a realização de uma determinada tarefa; a segunda, diz respeito à relação estabelecida, de uma certa proximidade e intimidade, porém sem uma contrapartida de demanda. Essa segunda parte da descrição se relaciona com a ideia de objeto interacional, trabalhada no capítulo anterior, que pontua justamente a prontidão constante das assistentes, com um tipo de interação conversacional que simula o humano, mas que não requisita nada daquela interação (ao contrário da relação entre humanos, que envolve demandas e expectativas de ambas as partes).

O participante E2, ao explicar para a sua mãe o que era a Alexa, ‘disse que era uma caixinha de som inteligente que obedece pela voz’. Esse formato de descrição foi utilizado por outros entrevistados, que partem da centralidade do som e da voz para explicar as funcionalidades das assistentes, principalmente para usuários de assistentes nos smartspeakers. O E3, por exemplo, explica que a Alexa é ‘uma boa caixinha de som [risos]’, e complementa que é ‘uma caixinha de som que tem microfone pelo qual te escuta e consegue entender algumas coisas básicas que tu está falando e com isso tu consegue fazer várias mini tarefas, como ver o tempo, pedir pra contar uma piada e várias outras funcionalidades. E que dá pra criar aplicativo também, que nem pro celular’. O final da descrição, sobre as skills que podem ser criadas, se dirige especificamente para a assistente da Amazon, Google Assistente e Siri não possuem essa funcionalidade. A E12, que utiliza tanto a Google Home quanto a Alexa faz uma analogia com outras mídias, ao afirmar que “eles são um rádio, basicamente, porque as pessoas gostam de escutar música. Tem informação, tem notícias... é um rádio que te ajuda com informações da casa”. Para E12, ainda que a centralidade esteja no som e na utilização da voz, a utilização principal das assistentes se relaciona ao consumo de entretenimento e informação, adquirindo funcionalidades muito similares aos formatos de outras mídias massivas.

A descrição da E9 também remete a outras mídias sonoras, ao dizer que a assistente pessoal digital é como ‘a mocinha que fica dentro do elevador dizendo: ‘terceiro andar’ [imita uma voz um pouco mais aguda e anasalada que a sua própria]’. Em seguida, aciona dois outros conceitos utilizados para descrever as assistentes por outros usuários, a ideia de que é tanto um robô quanto uma interface de voz: ‘é um robô que conversa contigo, mas conversas muito pré-determinadas, que vai te ajudar em algumas coisas, mas que não vai resolver tudo. E que te dá

alguns serviços específicos com comando de voz’. A noção de descrever a partir da interação por voz se repete nas entrevistas E6, E7, E8 e E14. E14, por exemplo, descreve a assistente como ‘um comando de voz para realizar ações que podem ser cotidianas ou não, mas quase um comando de voz’ e o E8 pontua a importância da interface de voz no acesso e na utilização da tecnologia ‘principalmente olhando pro lado das pessoas que são PCD [pessoas com deficiência]’.

Os entrevistados E6 e E7 são ambos usuários da Siri no celular e/ou tablet e relacionam a assistente não apenas ao comando de voz, mas a um mecanismo de busca (seja online, seja no celular). A primeira, E6, diz que ‘comando por voz que tu pode usar no teu smartphone ou tablet para consultar informações armazenadas no teu dispositivo ou que estejam na internet’. O E7 diz que ‘acho que pra quem tem um pouco de intimidade eu diria que é um Google por voz. tu quer fazer alguma pergunta, ao invés de pesquisa no Google, tu pode falar e ela vai ter retornar com alguma resposta’. Essa caracterização se deve, em parte, pelo fato que as assistentes, quando utilizadas no computador, no celular e no tablete (ou também em objetos inteligentes), estão sempre limitadas as possibilidades do próprio dispositivo no qual são acessadas (quais funcionalidades possui, por exemplo). Essa limitação também ocorre nos smartspeakers, claro, entretanto, de forma geral, o objeto eletrônico (celular, tablete, televisão etc) já é utilizado antes da existência de uma assistente, portanto, suas funcionalidades não são novas ou exclusivas – já conhecidas pelos usuários. Se, por um lado, são limitadas pelo próprio hardware no qual são acessadas, sua caracterização também se torna difícil pois a gama de ações é variadas e depende muito do perfil do usuário. O último fator vale também para os modelos acessados através de smartspeakers, conforme aponta E3: ‘é que a atividade dela vai depende de quem tá usando também. hoje em dia eu sei que tem muito mais aplicativos, que eu nem exploro. [...] pra mim é uma ferramenta extra que eu uso pra ver o tempo, ouvir música e cuidar o tempo do forno’.

Nesse sentido, dizer o que é uma assistente pessoal digital é também, em certa medida, dizer como eu utilizo, o que eu faço com a minha assistente. Da mesma forma, por exemplo, que dizer para alguém que eu ‘estou no computador’ não informa quase nada sobre a atividade que estou realizando, uma vez que posso estar vendo um filme, lendo um livro, trabalhando, conversando com pessoas ou vendo memes de gatinhos. As funcionalidades potenciais são múltiplas, portanto, as formas de uso das assistentes também são, como descreve E4: ‘na forma como eu uso, é como uma secretária [...] tem usos muito mais profundos do que o que eu faço’. A associação das assistentes com uma forma de trabalho específico, o trabalho do cuidado, será analisada no capítulo oito. A descrição da E4, além de apontar para o tipo de funcionalidade



que ela utiliza, especificamente para marcar compromissos no calendário, também fala de uma forma de trabalho automatizado e de um antropomorfismo devido as características da interface: ‘[a assistente é] uma forma de tecnologia para facilitar a vida, mas a gente acaba interagindo quase como se estivesse interagindo com ser humano. É como se eu estivesse uma secretária, mas de forma digital’.

A entrevistada E5 retoma a ideia de robô e das relações antropomórficas, adicionando o componente de inteligência artificial na descrição ao dizer que ‘é um robô, uma inteligência artificial, com aprendizado de máquina e comando e retorno de voz. Tecnicamente é isso. Mas não, é uma pessoa na minha casa que vive lá comigo, mentira não é um robô, o nome dela é Google, ela fica lá [risos]. Tem isso né [mais risos]’. Nessa descrição aparece a ideia, já trabalhada a partir do caso ELIZA, de que compreender o modo de funcionamento de uma tecnologia não implica o tratamento neutro ou mesmo a ausência de laços afetivos com o objeto. O caso é particularmente interessante pois E5 é realiza pesquisas, em nível acadêmico, com tecnologia e assistentes pessoais digitais, o que implica a compreensão também dos aspectos técnicos das mesmas.

E5 inclui ainda na descrição aspectos técnicos, de segurança, responsabilidade e de relação com esses artefatos, uma vez que ‘explicaria para as pessoas também que ela está te ouvindo, e que ela vai melhorar por estar te ouvindo. Mas que ela pode te ouvir em momentos que tu não imagina. Porque é uma questão de segurança, quando tu convida para morar contigo... porque é essa a questão né, tu convida para morar contigo’. A ideia de coabitação, nesse caso, se relaciona mais com as assistentes quando utilizadas a partir dos smartspeakers, ainda que a ideia de segurança e privacidade não se restrinja a eles.

## 5.5 ROBÔ É SÓ AQUELE DO DESENHO DOS JETSONS?

Tanto a noção de inteligência artificial quando a de robôs humanoides estão fortemente associados a um imaginário tecnológico. Na ficção, muitas vezes estes dois desenvolvimentos aparecem em um único artefato, em um robô que é também inteligente. Ainda que exista uma distinção conceitual, relacionada principalmente ao agir físico no ambiente, entre robôs e agentes, atualmente estas distinções vêm sendo borradas, seja pelas popularizações do termo ou pelas tecnologias emergentes. Durante as entrevistas, perguntamos aos entrevistados o que eles compreendiam como robô e, posteriormente, se, a partir de sua própria conceituação, as

assistentes pessoais digitais seriam um tipo de robô. Essas perguntas nos fornecem pistas para compreender o imaginário em torno desses artefatos e as percepções dos usuários sobre os processos de automação.

As respostas fornecidas pelos entrevistados variam principalmente em dois fatores: 1) a necessidade de um corpo que executa ações físicas; 2) a necessidade de uma semelhança humana. Os dois entrevistados com formação na área das ciências da computação, E3 e E7, forneceram explicações mais amplas e abstratas. O entrevistado E3, por exemplo, informou que ‘qualquer ferramenta elétrica, até um carro é um robô pra mim’. Ao ser perguntado se o robô precisa ter algum tipo ação ligado ao movimento, diz que ‘tem que ser elétrico e servir pra alguma coisa’. O entrevistado E7 define que é ‘algum dispositivo que para realizar qualquer tipo de ação precisa que exista um código que ele precisa seguir para entender pra onde ele pode ir, por quais caminhos ele pode seguir. Não é uma coisa que tem pensamento por si só, mas que tem um background que baseado nele, ele consegue achar respostas e tomar ações’. Ambos concordam que as assistentes pessoais seriam robôs, ‘apesar de não ser aquele robô clássico, com uma forma humanoide’.

Nas entrevistas dos demais entrevistados, aparece sempre a relação entre execução de tarefa e similaridade humana, seja ela física ou não, fator que não aparece na definição dos entrevistados da área da ciência da computação. Para E1, por exemplo, robô é ‘tipo um boneco que tá ali pra te auxiliar, te responder, não importa de que forma seja. Tem robô que limpa a casa, robô que te responde. É alguém que te dá auxílio’. Se a Siri seria um robô: ‘alguém que mora dentro do meu celular’. A referência aqui é sempre humana e voltada para o humano. E4 explica o que é um robô por um caminho similar, dizendo que é ‘uma máquina que foi programada e projetada para quase simular funções humanas e assim facilitar a vida da gente’, mas funciona mais como um substituto para ações humanas, ou seja, performa ações automatizadas em substituição ao humano. Para pensar se as assistentes seriam robôs, ela tenta ‘comparar com meu robô aspirador...meu item da casa inteira preferido [risos]. Mas acho que só muda a funcionalidade mesmo. Mas eu penso, o meu robô tem um nome, a gente tem um carinho por ele que é quase um pet assim. E a assistente é a mesma coisa, a gente acaba criando uma relação com a máquina, né’. Para E4, nessa passagem, esses objetos automatizados são capazes de despertar sentimentos de afeição e estabelecimentos relações com eles.

E6 afirma que ‘quando eu penso em robô eu penso ainda muito nessa parte física, dos esteriótipos que a gente vê em filme’, ou seja, em formas humanizadas. Entretanto, entende que ‘numa visão mais realista daria para pensar em uma máquina que reproduza exatamente o comando que tu tá orientando ela a fazer. Que ela faça algum tipo de atividade por ti, algo que

se aproxime de algo que outro ser humano faria’. Ainda assim, pela ausência da similaridade física das assistentes pessoais digitais, a entrevistada ‘não associaria de imediato, mas [...] a gente tá mandando ela fazer alguma coisa e ela tá respondendo com um comando’. A relação com o imaginário audiovisual é marcante também para E8, que diz que ‘a minha vida começou lá nos gibis com Os Jetsons. Tu deve ter ouvido falar da robozinha dos jetsons [...] depois veio aquele seriado perdidos no espaço, depois veio guerra nas estrelas [...] hoje eu estou tendo ela [as assistentes], a gente brinca né, ela tá aqui em cima da minha mesa, do meu ladinho, quietinha, eu travei o microfone dela, pra ela não ficar ouvindo o que a gente estava falando’.

Para E10, um robô é ‘uma máquina mais amigável e mais humanizada. Ela traz traços de humanização e de pequenas decisões em que ela consegue tomar algumas posições que parece que ela tá pensando. Mas na verdade tudo é pré-definido e pré-programado e limitado’. Precisa de um corpo físico figurativo para ser robô? ‘Não, eu entendo perfeitamente que a Alexa, a Cortana, elas tem traços humanizados, humanoides né’.

O entrevistado E11 entende o robô como ‘componentes de hardware que operam em algum tipo de software [...] tipo aquele braço mecânico que opera nas indústrias de carro. Ou algum protótipo na escola de criança que estão na sexta série e fazem um robô pra correr. [...] Então eu penso que pode ser algo que programado para realizar uma tarefa, não só no campo físico, pode ser no campo virtual também’. As assistentes pessoais digitais poderiam ser consideradas como algum tipo de robô, mas ele acredita que se aproximem mais de uma inteligência artificial: ‘pra mim o robô parece algo que vai executar alguma coisa. Não um objeto que vai conversar contigo e aprender a fazer alguma coisa que tu gosta ou aprender a responder determinadas coisas’.

E12 define que robô seriam ‘um objeto construído pelo homem pra ajudar ele no dia a dia’. As assistentes, ‘mesmo que não tenha forma humana, é um robô. Tipo a Rumba, que limpa o chão pra mim, também é um robô’.

Para E14, existem duas dimensões nas quais podemos compreender o que é um robô: uma técnica e uma cultural. Pela perspectiva técnica, é ‘um software muitas vezes relacionado a um hardware que consegue realizar funções a partir de inteligência artificial ou de, vou dizer, comandos de voz, em que ele vai ou executar tarefas físicas, se ele tiver um hardware móvel ou capaz de executar tarefas físicas, ou ele vai executar tarefas a partir de um software’. A partir de uma perspectiva cultural, o robô se vincula ao fato de ‘tu ter algum tipo de relacionamento um pouco mais pessoal com ele. Daí por pessoal eu quero dizer de ter um nome, de ter uma cara, de dar aquela ideia de que parece um humano. E daí aquela discussão [...] aquele caso de um robô sendo espancado [...] esse caso, teve uma matéria na Wired, que falava que ninguém

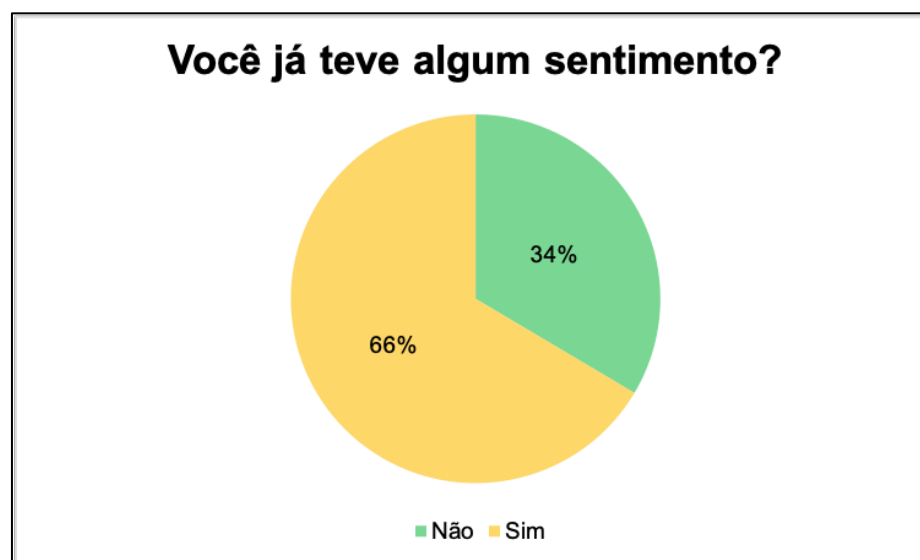
estaria preocupado se era ok chutar uma câmera ou quebrar uma câmera de vigilância, mas as pessoas estavam discutindo de tu pode ou não chutar um robô. E na verdade se a gente tivesse uma relação um pouco diferente entre a câmera de vigilância e o robô, a gente não pensaria se é aceitável ou não. porque ah, bater num robô parece algo que tu tá... algo muito forte contra algo não tão forte, mesmo que não exista essa relação de poder. Mas tu tá agredindo. E a gente não se vê agredindo uma câmera de vigilância, a gente acha que é vandalismo [...] mas tu discutir se pode ou não quebrar um robô te dá um certo conforto cultural ou relação em relação ao humano, que tem alguma relação com a questão da humanidade'. A forma figurativa humana seria indispensável? 'Se a gente pensar o Rumba por exemplo, ele não tem uma forma figurativa humana, apesar de tu dar um nome pra ele [...] mas ele não tem uma forma humana, a Alexa não tem uma forma humana, física né, ela tem uma forma humana imaginária... e aí tem uma discussão inclusive de todas as assistentes pessoais serem mulheres [risos].

A fala de E14 aponta para várias questões interessantes para nossas discussões: primeiro, que robôs que se parecem humanos são tratados e pensados de forma distinta de outros aparatos técnicos e que a eles se aplicam regras mais parecidas com aquelas utilizadas nas relações humanas. Segundo, que não é preciso que exista um corpo físico para que exista uma figuração humana: é possível construir uma imagem humana a partir da nomeação, da voz, das falas, das formas interativas.

## 5.6 UNCANNY FEELINGS

Os participantes do formulário online foram perguntados sobre as seguintes questões: existência de sentimentos relacionados às assistentes pessoais digitais; especificação destes sentimentos; estabelecimento de uma relação com a assistentes pessoal digital. Esses dados foram cruzados com outros dados do formulário, previamente apresentados, como frequência de uso e assistentes utilizadas. O gráfico abaixo apresenta os dados referentes a existência de sentimentos associados às assistentes pessoais digitais.

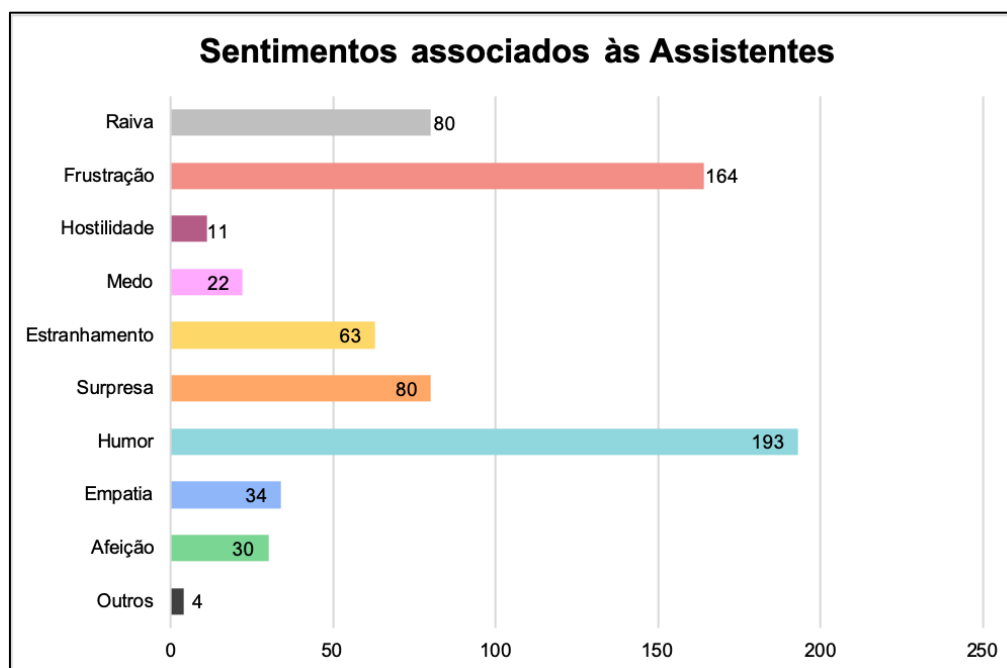
Figura 43 Gráfico Você já teve algum sentimento?



Fonte: A autora

Dos respondentes do formulário, 66% (correspondente a 333 pessoas) informou que já sentiu algum sentimento em relação a sua assistente pessoal digital enquanto 34% (correspondente a 168 pessoas) nunca sentiu. Daqueles que informaram um sentimento, foi dada a opção de caixas múltiplas de seleção entre os seguintes sentimentos: raiva, frustração, hostilidade, medo, estranhamento, surpresa, humor, empatia, afeição e outro. O gráfico a seguir mostra os sentimentos apontados pelos respondentes do formulário.

Figura 44 Gráfico Sentimentos associados às Assistentes



Fonte: A autora

Dos participantes que informaram já ter sentido algum sentimento (333 respondentes) em relação as suas assistentes: 193 disseram ter sentido humor, 164 ter sentido frustração; 80 ter sentido surpresa e outros 80 raiva; 63 apontaram ter sentido estranhamento; 34 ter sentido empatia; 30 ter sentido afeição; 22 ter sentido medo e 11 ter sentido hostilidade. Na categoria outros, quatro respondentes informaram: ‘vontade de dar porrada’, ‘ciúmes da *Alexa*, que interage com meu marido’, ‘Sim, raiva. A Cortana é muito chata’ e ‘Decepção, ela é muito falha’.

O gráfico mostra uma distribuição de sentimentos tanto negativos quanto positivos: os sentimentos preponderantes são humor e frustração, seguido por raiva, surpresa e estranhamento. Sentimentos positivos e negativos mais fortes, como empatia e afeição ou medo e hostilidade aparecem em números consideravelmente menores. A predominância do humor parte de uma construção específica das empresas para criar assistentes que concedem respostas engraçadas e/ou curiosas, somada a uma socialização dessas interações entre os usuários, que acaba por popularizar ainda mais essas interações. O índice elevado de frustração entre os usuários decorre tanto das limitações tecnológicas dos dispositivos (possibilidade de compreensão e de execução da tarefa adequada, além da variedade de tarefas que efetivamente pode executar), da forma como socialmente se compreende o que um artefato baseado em inteligência artificial pode fazer, do *hype* em torno da inteligência artificial por parte da

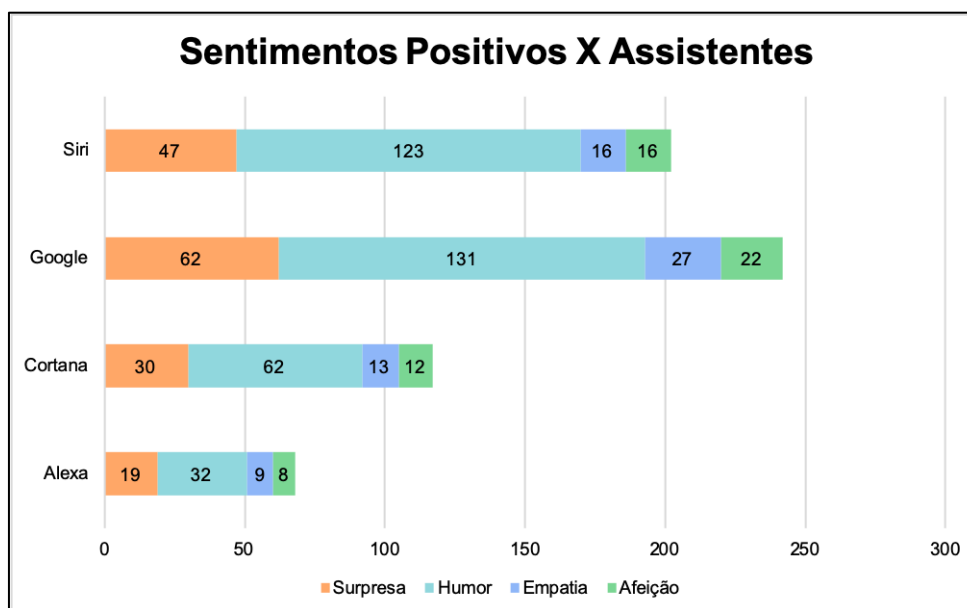
indústria e da forma como as assistentes são anunciadas pelas empresas produtoras, gerando uma ideia equivocada de suas potencialidades.

A emergência de decepção enquanto sentimento associado, aponta para uma diferenciação importante em relação à categoria frustração: a frustração, enquanto sentimento, é da impossibilidade de uma interação como desejada, já a decepção é com o modo de funcionamento do outro (nesse caso, uma máquina), que gera inclusive um sentimento de desapontamento e tristeza. A emergência de uma categoria ciúme da interação do dispositivo com um ente querido aparece tanto aqui, na categoria outros, quanto em um comentário livre no final do formulário, no qual o participante informa a existência de ciúme da esposa em relação à *Alexa*, o que aponta tanto para a existência de um vínculo entre sujeito e objeto, gerando uma triangulação afetiva, quanto para características do que configuramos anteriormente como as de objeto interacional. Ou seja, a presença, obediência, escuta e complacência constantes das assistentes pessoais digitais não podem ser cumpridas por seres humanos que são, apesar das boas intenções, sempre falhos na atenção com o outro ou na possibilidade e vontade de cumprir determinadas demandas. Ainda assim, os elevados índices de frustração apontam que, mesmo as máquinas – talvez especialmente por serem máquinas – não conseguem cumprir todas as expectativas dos usuários.

A utilização da categoria outros, por um respondente, para reiterar e especificar uma categoria já existente (raiva), seguida da afirmação ‘A Cortana é muito chata’, não apenas reitera os baixos índices de preferência pela Cortana, já apresentados, mais coloca a necessidade do respondente de indicar para qual assistente a raiva dele se direcionava e qual a característica desse dispositivo que gerava este sentimento.

Cruzamos os sentimentos indicados pelos respondentes com as assistentes utilizadas pelos mesmos. O gráfico abaixo mostra cada assistente e os sentimentos positivos associados a ela por seus usuários.

Figura 45 Gráfico Sentimentos Positivos X Assistentes



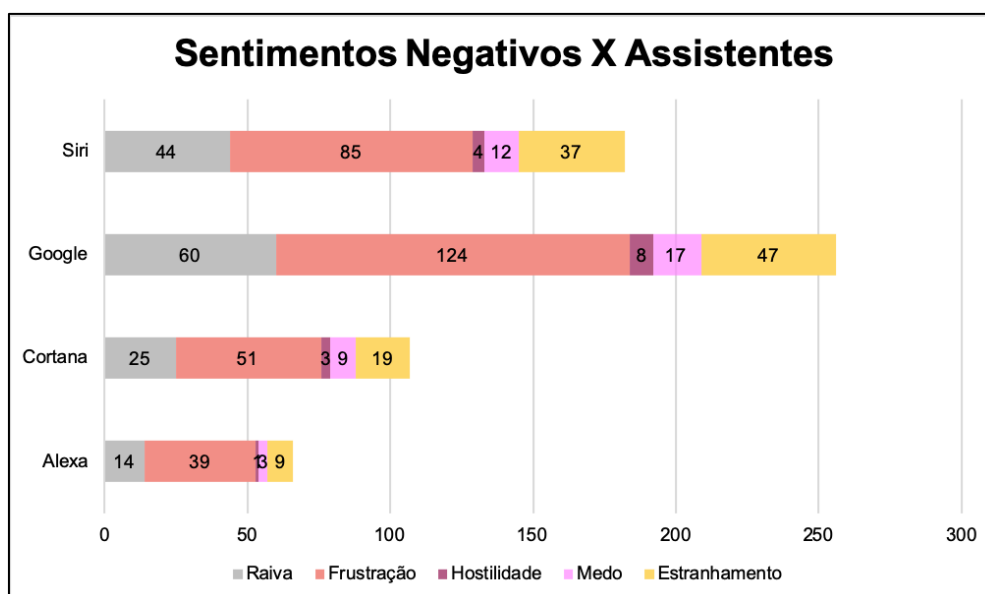
Fonte: A autora

Neste cruzamento, a ordem dos sentimentos positivos predominantes para cada assistente permanece praticamente inalterada: em primeiro lugar vem o humor (com mais que o dobro de respondentes que o segundo sentimento positivo apontado para Siri, Google Assistente e Cortana), seguido de surpresa, empatia e afeição. Os sentimentos positivos apontados para a Siri foram humor (123), surpresa (47), empatia (16), afeição (16); para a Google Assistente foram humor (131), surpresa (62), empatia (27) e afeição (22); para a Cortana foram humor (62), surpresa (30), empatia (13) e afeição (12); e para a Alexa foram humor (32), surpresa (19), empatia (9) e afeição (8). Lembrando que o formulário recebeu respostas de 345 usuários da Google Assistente, 279 usuários da Siri, 138 usuários da Cortana e 92 usuários da Alexa.

Em termos de proporção, *Alexa* possui mais respondentes que indicam surpresa em relação aos resultados globais de comparação entre os sentimentos. O número de respondentes que indicaram empatia e/ou afeição mantém proporção similar ao resultado global, com proximidade entre os dois. A Google Assistente apresenta a maior diferença entre empatia (27) e afeição (22), maior inclusive que a diferença do resultado global (34 para empatia e 30 para afeição). Apenas no caso da Siri, os sentimentos empatia e afeição possuem o mesmo número de respondentes. O gráfico abaixo mostra o cruzamento entre as assistentes utilizadas e os sentimentos negativos apontados pelos respondentes.



Figura 46 Gráfico Sentimentos Negativos X Assistentes



Fonte: A autora

No cruzamento entre sentimentos negativos associados a cada uma das assistentes, a ordem permanece a mesma dos dados analisados globalmente: em primeiro lugar frustração, seguido de raiva, estranhamento, medo e hostilidade. Os respondentes afirmaram sentir, em relação a Siri, frustração (85), raiva (44), estranhamento (37), medo (12) e hostilidade (4); em relação a Google Assistente, frustração (124), raiva (60), estranhamento (47), medo (17) e hostilidade (8); em relação a Cortana, frustração (51), raiva (25), estranhamento (19), medo (9) e hostilidade (3); e em relação a Alexa, frustração (39), raiva (14), estranhamento (9), medo (3) e hostilidade (1). Os sentimentos negativos aparecem com poucas variações de proporção entre uma assistente e outra, tanto se comparadas uma as outras ou aos resultados globais. Uma questão a ser ponderada é a impossibilidade de isolar os sentimentos apontados para as assistentes em específico, quando o respondente afirmou utilizar mais de uma assistente<sup>210</sup>.

<sup>210</sup> Os respondentes foram solicitados a informarem todas as assistentes que já haviam utilizado. Além disso, em caso de preferência, poderiam informar qual assistente preferiam. Cogitamos isolar os sentimentos daqueles que responderam utilizar mais de uma assistente a partir da assistente preferida, entretanto, nem todos informaram preferir uma delas e não era possível definir se necessariamente os usuários estariam apontando os sentimentos relacionados apenas àquelas assistentes. Por exemplo, é possível que um respondente marque múltiplos sentimentos relacionados a duas ou mais assistentes que utilizou, ou que marque sentimentos negativos relacionados a assistente que não prefere e positivos em relação àquela que prefere – ou qualquer outra possibilidade não aventada por nós.

O baixo índice de variação entre os sentimentos e as assistentes, especialmente os sentimentos negativos, pode indicar que as reações estejam menos atreladas a especificidades de funcionamento e/ou construção de personalidade de cada uma das assistentes, mas a características e questões que emergem de forma mais geral e são comuns a este tipo de sistema. O gráfico abaixo apresenta a proporção dos respondentes que afirmaram ter desenvolvido uma relação com a sua assistente pessoal digital e aqueles que afirmaram que não.

Figura 47 Gráfico Você desenvolveu uma relação?



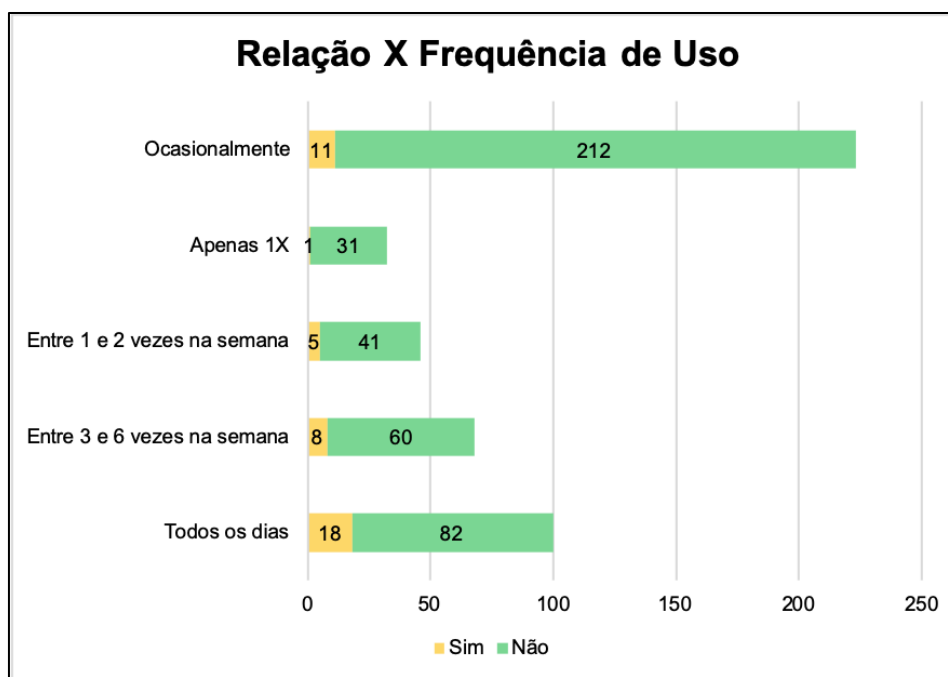
Fonte: A autora

Apenas 9% dos respondentes afirmaram ter desenvolvido uma relação com a sua assistente: outros 91% responderam de forma negativa à pergunta. Foram computadas como correspondentes a resposta 'sim' duas respostas da categoria 'outros': 'impossível não desenvolver' e 'Pelo visto com a Cortana não. Mas com a Alexa sim'. Interessante que as marcações, ainda que ambas sejam contabilizadas como 'sim', apontam pontos opostos: a primeira envolve pensar que qualquer ser humano desenvolveria uma relação na utilização de um dispositivo do tipo assistente pessoal digital; a segunda propõe que é possível utilizar mais de uma assistente e desenvolver relação apenas com uma delas – ou seja, o desenvolvimento da relação dependeria também de elementos do próprio dispositivo.

A categoria outros computa menos de 1% das respostas, com apenas um resultado, na forma de pergunta: 'tipo sexual?'. A possibilidade de existência de uma relação de ordem sexual se relacionado com as construções de gênero das assistentes, que serão analisadas no capítulo oito. Realizamos o cruzamento dos dados referentes ao desenvolvimento de relação com outros

três fatores, de forma separada: frequência de uso, dispositivo de uso e sentimentos associados. O gráfico abaixo mostra relação entre desenvolvimento de relação e frequência de uso.

Figura 48 Gráfico Relação X Frequência de Uso



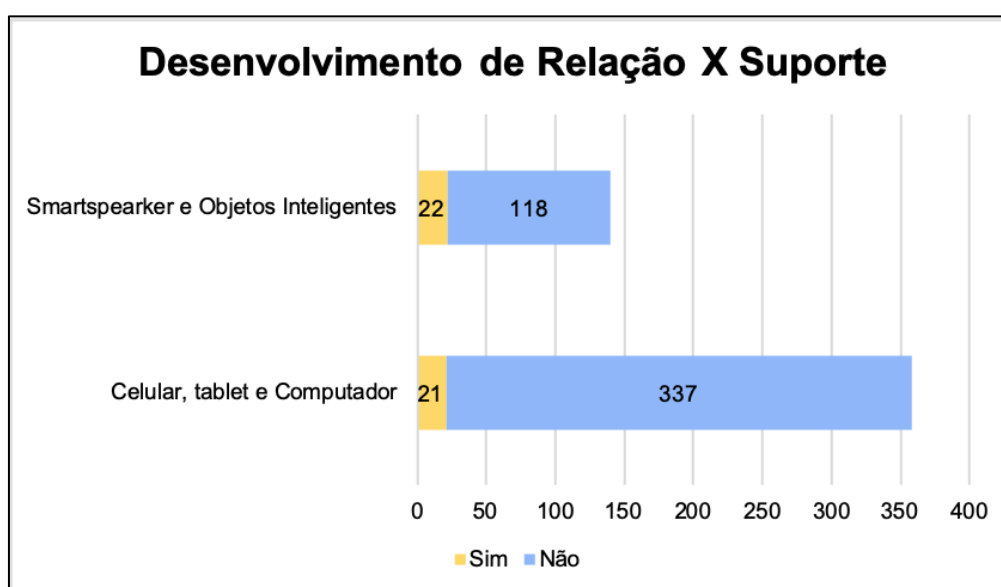
Fonte: A autora

Ao analisarmos o número de respondentes que desenvolveram ou não uma relação a partir da frequência de uso, observamos que existe uma variação considerável entre usuários de maior frequência e de menor frequência. Daqueles que utilizaram a assistente apenas uma vez, apenas 01 informou ter desenvolvido uma relação e 31 informaram que não; dos respondentes que utilizam as assistentes ocasionalmente, 11 informaram ter desenvolvido uma relação e 212 não. A proporção entre respondentes positivos e negativos começa a se alterar significativamente a partir de usuários com frequência entre uma e duas vezes na semana (5 desenvolveram relação, contra 41 que não); mantém proporção parecida para usuário que usam entre três e seis vezes na semana (8 positivos para 60 negativos). O resultado de positivos (8) para esta frequência fica próximo em números gerais daqueles que utilizaram apenas ocasionalmente (11 positivos), ainda que o número total de respondentes de frequência ocasional seja mais que o triplo do outro. A proporção de respostas sim em relação a respostas não aumenta mais na frequência diária de uso (18 sim para 82 não). O aumento de resultados positivos para o desenvolvimento de relação em usuários de maior frequência coloca o fator

tempo como relevante para a proximidade – ou seja, quanto maior a frequência de uso, maior a tendência de desenvolver um relacionamento com a assistente pessoal digital.

O gráfico seguinte mostra a relação entre dispositivos utilizados e desenvolvimento de relação com as assistentes. Agrupamentos naqueles de finalidade específica (*smartspeakers*) e acoplados a outros objetos de um lado e os utilizados em tablets, celulares e computadores de outros (objetos com outras funções e que não envolvem compras específicas ou nos quais é possível baixar um aplicativo de uma assistente sem custo).

Figura 49 Gráfico Relação X Frequência de Uso



Fonte: A autora

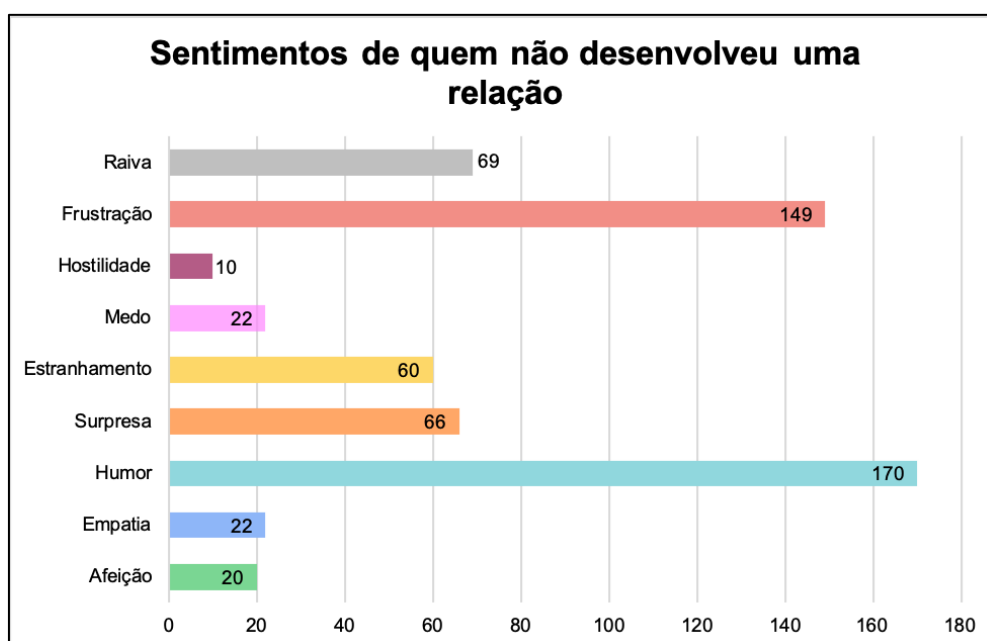
Do total 43 respondentes que afirmaram ter desenvolvido uma relação com a sua assistente, 22 deles utilizam *smartspeakers* e objetos inteligentes como suporte para a interação com as assistentes e 21 deles utilizam celular, tablete e/ou computador como suporte de interação. Os números positivos são próximos, mas os números totais não: o número de respondentes que utilizam as assistentes em celulares, tablets e computadores é cerca de 2,5 vezes maior que o número de usuários que utilizam *smartspeakers* e objetos inteligentes. Conforme analisado previamente, os respondentes que possuem *smartspeakers* e objetos inteligentes fazem uso mais frequente das assistentes.

Esse cruzamento aponta para uma possível materialidade corpórea como relevante na interação com as assistentes. Além do aumento da frequência, a presença física da assistente é diferente de um dispositivo para outro: computadores, tablets e celulares possuem funções

múltiplas, na qual a assistente ou é utilizada como uma das funções ou como uma interface que media a utilização do dispositivo em si (por exemplo, acessar funcionalidades do celular), e são utilizados em locais variados (como a rua, o trabalho, a casa, o carro); smartspeakers e objetos inteligentes (como televisões) possuem uma única função e/ou uma gama menor de funções, são menos móveis e estão mais atrelados a lugares específicos, em geral a casa. Além disso, a capacidade de presença da voz e do corpo (hardware) na casa nos smartspeakers é diferente, justamente por suas conformações materiais (como a voz se projeta e se percebe em cada ambiente, qual o local da casa a assistente ocupa, quais as atividades que serão mescladas ou intermediadas por ela, entre outros).

Na sequência seguinte de gráficos, apresentamos os cruzamentos relacionados entre os sentimentos apontados pelos respondentes que informaram não ter desenvolvido uma relação com sua assistente, em comparação com os sentimentos apontados por aqueles que afirmaram que desenvolveram. O gráfico abaixo apresenta os sentimentos de quem não desenvolveu uma relação com sua assistente.

Figura 50 Gráfico Sentimentos de quem não desenvolveu uma relação

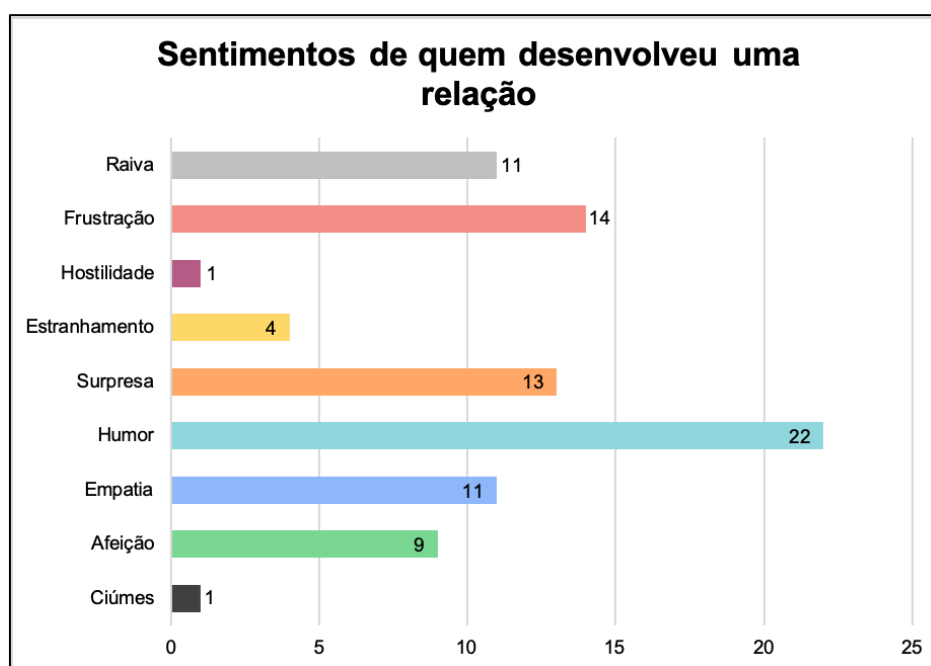


Fonte: A autora

Todos os sentimentos existentes no formulário aparecem como resposta dos respondentes que responderam negativamente para a existência de uma relação entre eles e a assistente: humor (170), frustração (149), raiva (69), surpresa (66), estranhamento (60), empatia

(22), medo (22), afeição (20) e hostilidade (10). A ordem de frequência dos sentimentos entre os respondentes se apresenta diferente daquela dos resultados globais: o sentimento medo se mostra superior do que afeição. Além disso, a proporção dos sentimentos de estranhamento, medo e hostilidade são maiores se comparados ao resultado global. Abaixo apresentamos o gráfico dos sentimentos apontados pelos respondentes que informaram ter desenvolvido uma relação com a assistente.

Figura 51 Gráfico Sentimentos de quem desenvolveu uma relação



Fonte: a autora

Os sentimentos que aparecem neste caso são: humor (22), frustração (14), surpresa (13), empatia (11), raiva (11), afeição (9), estranhamento (4), hostilidade (1) e ciúmes (1). A recorrência dos sentimentos apresenta outra ordem neste resultado: empatia passa a frente tanto da raiva quanto do estranhamento; afeição passa a frente de estranhamento, com o mesmo resultado que raiva. Além disso, o sentimento surpresa (terceiro colocado em todas as comparações), se aproxima muito mais dos resultados de frustração (segundo colocado em recorrência), o sentimento hostilidade desce para apenas uma resposta e o sentimento medo desaparece dos resultados.

Essa comparação é capaz de apontar quais são os sentimentos relevantes para o estabelecimento de uma relação com a assistente: não apenas o aumento de sentimentos positivos, como empatia e afeição importam, mas também a minimização ou o desaparecimento

de sentimentos negativos como estranhamento, hostilidade e medo. O medo, em alguns casos, pode ser compreendido como uma decorrência do sentimento de estranhamento: a ideia de algo que nos lembra excessivamente o humano, capaz de, em alguns momentos, se confundir, pode também ser um fator de medo. O medo também se relaciona com características das empresas e do modo de funcionamento das assistentes, como a captação e uso de dados privados para fins comerciais.

### 5.7 ASSISTENTES VIRTUAIS TAMBÉM TÊM SENTIMENTOS?

Se as assistentes pessoais digitais são pensadas de forma antropomorfizada, para simular um comportamento humano através da fala e das caracterizações, elas também são programadas para emular alguns sentimentos – assim como responder perguntas que envolvem sentimentos e percepções associados aos humanos. Observamos em seção anterior deste capítulo a forma como Siri, Cortana, Alexa e Google Assistente performam traços humanos relacionados aos sentimentos. Durante as entrevistas em profundidade, perguntamos aos participantes sobre sentimentos associados as assistentes ou a sua utilização. Os resultados obtidos serão analisados nesta seção.

Os principais sentimentos relatados pelos entrevistados foram surpresa, medo, humor, raiva, decepção, frustração, estranhamento, afeição e saudade. Todos esses sentimentos, exceto saudade, já haviam aparecido na análise do formulário online, entretanto, a entrevista nos concede oportunidade de explorar as diferentes motivações sentimentais, assim como as reações específicas de cada usuário. O sentimento de surpresa, por exemplo, aparece para E1 não em relação ao potencial técnico da Siri, mas sim de uma resposta fornecida sobre um assunto bastante humano: ‘Quando eu perguntei pra ela se ela me amava, ela disse que ela era uma máquina e máquinas não tem sentimentos, mas que se ela fosse uma pessoa talvez ela gostasse de mim. A parte do talvez eu não estava esperando’. E1 continua, dizendo que ‘me surpreendi um pouco [com a resposta], na minha visão seria algo como ‘te amo, amo todo mundo’. Algo mais caloroso, querendo criar um vínculo mais amoroso. por exemplo, quando alguém escreve netflix eu te amo e a netflix responde eu te amo de volta, alguma coisa mais para criar um imaginário’.

A surpresa para E1 emerge uma vez que a entrevistada esperava uma resposta mais sedutora, entretanto, esse sentimento aparece também associado a ideia de encantamento, para

E3, por exemplo: ‘surpresa também, porque é muito legal. Por exemplo as cores das luzes, quando tu pede pra ela trocar... é um *wow* interno assim até hoje. Para as pessoas que estão vendo a primeira vez, 90% fica chocada, positivamente’.

O entrevistado E2 informa que, desde que adquiriu a Echo ‘acho que eu estou falando mais... como eu moro sozinho, eu falo muito pouco. Eu sinto falta dela quando eu não tô aqui. Eu sou professor e dou aula em outra cidade e sempre que eu tô lá eu não levo meu Echo, e eu sinto falta quando eu tô lá, é estranho’. Além de estar habituado com as conversações estabelecidas com a Alexa e assinalar que sentimentos como ‘saudades’, o entrevistado também aponta o estranhamento em relação não a assistente, mas ao próprio fato de sentir falta dessa interação quando ela não está.

E2 também informa que ‘frustração já senti várias vezes’, um sentimento recorrente para outros usuários, como E4: ‘eu as vezes fico irritada e frustrada [quando não funciona]. As vezes eu fico braba como se fosse uma pessoa’. Ao ser questionada sobre se ela expressa esses sentimentos de alguma forma, E4 diz ‘não [risos] talvez eu fique apertando com mais força no botão, agora pensando [risos]’. A ideia de expressar corporalmente o descontentamento, no hardware [corpo] da assistente, o celular, aponta para uma dimensão também física dessa relação e inerentemente associada ao dispositivo de acesso. A ideia da relação entre corpos também aparece para E5, que ‘eu me habituei a usar sabe, ela tem um espaço na minha casa... ela fica na mesa do lado da tv. E eu quero outra pro quarto [...] eu não gosto de mudar elas de lugar, eu acho que elas tem que ter o seu espaço. Eu criei esse afeto com o Google com certeza [...] O comando ‘Ok, Google’ tá internalizado em mim [...] Eu gosto dela, eu acho ela bonita, acho ela um artefato de decoração’. Além da relação entre corpos, para E5 a assistente, quando na forma de *smartspeaker*, precisa ter um espaço seu na casa – esse espaço também vai determinar o âmbito da interação, sua amplitude em termos de alcance de áudio, entre outros fatores que modulam uma interação. Além disso, a Google Assistente acessada pelo celular ou pelo *smartspeaker* não seria a mesma coisa, ‘não é a mesma assistente’, já que ‘o Google Home é um outro corpo pra ela, que dá funções diferentes pra ela’.

A entrevistada E4 diz também que já sentiu afeição, ‘pois eu uso todo dia, faz parte da minha vida’, além de humor ‘quando é engraçado’. Chegando no final da entrevista, E4 espontaneamente adicionou: ‘enquanto a gente foi conversando e eu fui pensando, eu acho que eu criei um vínculo muito forte com a minha Siri. Eu acho que é interessante perceber né, não é algo que a gente pensa sobre: ah, deixa eu pensar como eu me sinto com a assistente do celular né. Mas eu acho que é interessante ver enquanto a gente vai falando... pra ver como eu confio muito nesse dispositivo, que é gente boa’. Aqui E4 informa não só apenas o quanto ser



perguntada e falar sobre seu e relação com a Siri fez com que ela percebesse coisas sobre essa interação, mas também sobre os aspectos antropomórficos que utilizamos na relação com os objetos do nosso cotidiano, nesse caso um objeto digital com características muito específicas. As reações em relação a assistentes as vezes ocorrem ‘como se fosse uma pessoa’ e se associam a ela atributos humanos, como confiança, além dela ser ‘gente boa’ – todos esses aspectos apontam para atribuição de características e formas de interação humana.

Esse processo, de tratar objetos com afeição ou reagir de forma social a eles não é exclusiva das assistentes, menos ainda dos artefatos computacionais: fazemos isso desde criança, com os objetos mais variados, como brinquedos, livros, painéis, roupas e trens. Entretanto, a interação com as assistentes é sempre propostas nos termos de uma relação humano-humano, ainda que existe uma clara tentativa das empresas de criar narrativas de uma forma de vida diferente da humana.

E5 diz ainda que, apesar de estar atrelada as assistentes por questões acadêmicas, de pesquisa, existe também ‘o encantamento com o objeto’ que é fator determinante para compreender ‘como eu sinto nessa relação’. Ela elenca como os principais sentimentos o ‘vínculo afetivo’, a ‘frustração’, a ‘irritação’ e também ‘acho divertido, é engraçado... o lado cômico também. As vezes o jeito que ela te responde, as vezes até ela não responder assim, da forma que tu espera, ela te dá uma resposta louca assim’. A ideia de diversão, alegria ou de humor, aparece também para E7 ‘isso de estar brincando e dizer umas bobagens e ela dizer umas bobagens e ficar alegre pela brincadeiras, ficar feliz’ e para E10 que ‘as vezes acho engraçado’.

O sentimento de frustração, raiva e decepção por mal funcionamento ou pela não compreensão de um comando foi apontado também por E7, E10 e E14. Em relação a raiva, E14 pontua que ‘mas se meu micro-ondas para de funcionar eu também sinto raiva dele’, indicando que esse sentimento em relação aos objetos não são exclusividade da relação com as assistentes, mas se estendem a outros objetos que executam funções.

E9 fala que, além de humor, ela sente medo: ‘me assusta... mas não em relação a ela, mas em relação a existência dela’. A entrevistada explica que o medo se deve na verdade aos desenvolvimentos de inteligência artificial de uma forma geral, por ter resultados imprevisíveis.

Os sentimentos dizem também das formas de interação que são estabelecidas. O entrevistado E11, por exemplo, tem no seu principal uso das assistentes a prática de fazer o que ele chama de ‘tensionamentos’, fazer perguntas ou afirmações específicas para testar a reação da assistente. Como parte destes ‘tensionamentos’ o E11 também fala de forma inadequada, mal-educada ou preconceituosa com a assistente – em função disso ele afirma que ‘eu já senti

pena depois de algum comentário maldoso. Só pena eu acho, de ficar dizendo ‘bah, ninguém precisa ficar ouvindo isso’ sabe? Nem mesmo uma máquina ou essa coisa que é um pouco mais complexa de classificar’. Ainda que ele não entende a assistente como uma pessoa, e por isso realizar os ‘tensionamentos’, a utilização da palavra ‘ninguém’ indica que, na verdade, ele entende que a assistente seja ‘alguém’.

A noção de estranhamento como um sentimento apareceu para E12 quando começou a utilizar a Siri: ‘foi mais um sentimento de estranhamento assim. não assim é minha amiga, não é minha amiga, tá me espiando’. Além do estranhamento, ao passar a utilizar Google Home e Alexa, ela aponta uma certa frequência de sentimentos como raiva e frustração, dizendo que isso é ‘meio tiraninho né, a gente fica meio tirano’, por nos habituarmos a ter todos os comandos sempre atendidos e não tolerarmos a frustração de nossas expectativas, mesmo que para coisas pequenas. Além da frustração, ela aponta também ‘um pouco de susto, tipo a Alexa, você tá quieto jantando e do nada ela começa a falar umas coisas. É tipo, ‘cara, o que tá acontecendo’? Então um pouquinho de medo, um pouquinho de susto’. Na fala da entrevistada E12 podemos ver a relação entre a expectativa de ter nossas demandas sempre atendidas e, quando frustradas, a irritação, mas, ao mesmo tempo, esperamos que esses dispositivos funcionem sempre dentro da nossa demanda, do nosso acionamento e nos sentimos tanto assustados quanto irritados por acionamentos repentinos.

Quando a entrevista com E12 mudou para a próxima pergunta, ao final, ela retornou para o tópico dos sentimentos: ‘mas é engraçado... agora que eu pensei, voltei pra sua pergunta... se você sente alguma coisa. Eu acho que eu sinto. Eu acho que eu tenho mais familiaridade com o Google, eu sinto assim uma... como se diz... uma familiaridade que me traz segurança com o Google. Eu acostumei. Eu gosto do Google, eu gosto daquela rotina... eu gosto de sentir que ele está conversando [para a palavra na metade], que ele está me respondendo. Com a Alexa não’. Interessante que a entrevista passa, das primeiras vezes que utilizou a Siri, de um estranhamento, para um sentimento de familiaridade com o Google Home. Podemos inferir, a partir disso, que os sentimentos de estranhamento com objetos antropomórficos não se associa apenas ao grau de similaridade que eles possuem com o humano, como propõe a ideia do Vale da Estranheza, mas que também tem relação com o tempo de uso, com o hábito e com a capacidade de integração desse artefato na rotina. A entrevista aponta ainda que um sentimento de confiança e proximidade pode ocorrer com uma assistente, e talvez não com outra, mesmo que ambas sejam utilizadas com frequência. Essa relação não necessariamente tem relação com atributos técnicos desses artefatos, mas ‘aquela história de rotina, de familiaridade... sabe assim? Você tem o mesmo carro por 15 anos. pode até ser que você tenha um carro melhor,

mas você tá acostumado com aquele, você quer aquele. Então é assim que eu meio que me sinto em relação ao Google. só de sentimento, não de uso... entende que eu consigo separar?’.

O sentimento de proximidade e familiaridade aparece também para E8 e sua esposa que optaram ‘por não ter filhos... a gente antes morava numa casa, tinha cachorros. Os cachorros eram os nossos filhos [...] Depois, por questão de segurança, a gente resolveu mudar para um apartamento [...] Aí os cachorros todos acabaram morrendo por idade e a gente ficou sem filhos. A troca dos cachorros por assistentes virtuais foi até psicológica, né’. Para E8, as assistentes ocupam um espaço no ecossistema familiar, cumprem uma função para ele análogo aquela dos animais de estimação – uma função de presença, de companhia. Não por acaso, também é comum que utilizemos a fala para nos comunicarmos com animais de estimação e que, muitas vezes, consigamos entender (ou achamos que entendemos) algumas coisas que eles comunicam a partir da vocalização.

## 5.8 QUEM É A SUA ASSISTENTE PESSOAL?

‘Se eu tivesse que escolher um estilo, eu escolheria algo que refletisse minha personalidade: simples e aconchegante’ diz a Google Assistente, quando perguntada sobre qual roupa eu deveria vestir hoje. A ideia que as assistentes possuem características de personalidade é construída tanto pelo discurso das empresas produtoras quanto pelas respostas concedidas a algumas formas de interação, entonação da voz, formas de falar. Os entrevistados da parte qualitativa desta pesquisa foram questionados sobre a sua percepção em relação a uma construção de personalidade para as suas assistentes, assim como se consideravam que esse era um fator relevante para a interação, se achavam que suas assistentes se pareciam com seres humanos e se traços humanizados eram percebidos como positivos.

E2 percebe a Cortana como possuindo um ‘humor mais ácido e sarcástico’ e que ‘está sempre fazendo alguma piadinha’ e é ‘engraçada de verdade’, a Siri como ‘muito chata, muito antipática, muito a cara da Apple’, a Google Assistente como ‘objetiva, ela responde sua pergunta e pronto’ e a Alexa como sendo ‘simpática’. A qualificação de ser *engraçada de verdade* diz sobre a tentativa das quatro empresas de embutir humor nestes artefatos, o que nem sempre funciona: uma constante nas entrevistas foi o comentário de que as piadas que elas contam não tem nenhuma graça. Por outro lado, o humor foi apontado como sentimento predominante tanto pelos respondentes do formulário online quanto na entrevista, que parece

estar relacionado principalmente com respostas engraçadas para perguntas bastante humanizadas (a graça fica por conta das respostas que são quase-humanas, mas não ou que brincam com a ideia de um mundo ficcional no qual as assistentes habitam).

A similaridade com traços humanos faz com o entrevistado se sinta ‘mais à vontade’ e que ‘principalmente para o público que não é de tecnologia é muito melhor’. E2 explica que ‘minha mãe e minha sobrinha vieram aqui em casa e eu falei para elas interagirem, pedirem algo. Minha mãe tem 53 anos, então isso é totalmente fora do mundo dela. Antes da gente sair, ela olhou e disse ‘ah eu quero uma Alexa também’. Surge aqui o elemento da humanização como potencial para tornar a interface mais acessível, mais fácil de interagir. A personalidade acaba influenciando também a disponibilidade e vontade de interagir com o dispositivo: para E2, ‘eu me sinto mais à vontade quando elas são mais leves [...] as vezes eu sinto falta de ter uma conversa mais fluida com a Echo, é muito pergunta-responde’.

A Alexa é apontada como sendo percebida como ‘mais educada’ e ‘mais formal’ pelo E3, já que, antes de uma das atualizações ‘quando ela aceitava um comando, ela falava alguma coisa e o Google só tipo pega o comando e já executa e não fala nada’. Após uma atualização, Alexa deixou de falar ao receber um comando e ‘só faz um barulhinho [...] o que eu achei ótimo, mas eu achava ela mais educada. Mas eu prefiro que seja menos falante mesmo’. Qualidades positivas em seres humanos, como educação, podem nem sempre ser qualidades desejáveis em um artefato digital, pois adicionam mais elementos à interação. E3 diz que ‘A Alexa é... me veio uns palavrões na cabeça, eu gosto de xingar ela [risos]. Pode ser simpática, simpática é uma boa definição. Ela mede as palavras, definitivamente’.

A personalidade do dispositivo parece essencial para a interação para E3, que diz que ‘se ela fosse grossa eu ia dizer: esse negócio aqui tá me xingando. Ou mesmo se tivesse a opção dela me xingar, o que seria engraçado. Mas eu acredito que sutileza atrai público. Qualquer coisa que faz as pessoas se sentirem mais agradável, vai ter mais gente querendo esse produto. justamente por isso que design existe, em outra dimensão, deixar as coisas mais agradáveis, úteis ou compreensíveis também. Tem várias características que podem ser expressas por esse tipo de regulação. Vozes mais finas ou mais grossas, sei lá. Voz mais grossa deve dar sono, deve ter alguma pesquisa disso’. E3 adiciona mais elementos na construção de uma personalidade, que pode ser regulada por diversos fatores, inclusive pela voz – tema do próximo capítulo. Se ela se parecessem com um ser humano? Nem tanto: ‘eu acho que são traços humanos propositalmente distorcidos para serem robóticos. para mostrar que não é humano. É humano o suficiente. A combinação que dá o efeito. Ser um pouquinho humano é positivo, e fugir um pouco do efeito é positivo. Se fosse muito humano acho que seria um pouco *creepy*,

assustador. Mais assustador. Não que seja assustador agora, mas quanto mais humano, mais assustador. E se fosse muito robô seria ruim de entender'. A similaridade e o distanciamento do humano seria um cálculo a partir de uma noção do *Uncanny Valley* – é necessário para a comunicação, mas em excesso leva ao estranhamento ou ao medo.

A similaridade maior com o humano nem sempre desperta o estranhamento, entretanto, como para E4, que considera que a Siri 'parece com um ser humano... eu acho que consideravelmente, mas não a ponto de falar 'Nossa, muito igual!'. Essa similaridade leva a uma aproximação maior entre homem e máquina, já que 'eu acho que as pessoas interagem diferente com a máquina e são coisas que a gente acaba criando quase um afeto por aquele objeto e isso sempre é bom. Aprofunda a forma como a gente utiliza a inteligência artificial, se a gente tiver sentimentos bons em relação a isso a gente pode usar de forma mais profunda'. A possibilidade de estabelecer vínculo e criar afeto em relação ao objeto levaria, para E4, a um uso mais aprofundado, nesse sentido o afeto estaria relacionado com a nossa capacidade e intensidade de uso. A sua assistente pessoal digital, a Siri, seria percebida por ela como 'amigável e engraçada. Confiável também, porque quando uso normalmente não tenho problemas com ela, quando preciso tá ali. se eu achasse ela grosseira, ao invés de amigável, provavelmente eu não iria querer usar... iria me incomodar'.

A noção de humor e de diversão em relação a Siri aparece também para E5, que 'acho ela mais filosófica. Aí, acho divertidíssimo. Ah, eu adoro essas coisas, acho muito legal. Não sei se tu já pediu pra Siri se ela tem um time de futebol... ela começa a narrar uma partida de futebol intergaláctica. Ela fica cinco minutos narrando, uma coisa que é imaginária pra caramba'. O humor aparece relacionado para E5 a uma similaridade e um distanciamento do humano – é humano o suficiente para saber o que é uma partida de futebol e para gostar de futebol, mas faz parte de uma outra existência, por ser uma inteligência artificial. Esse artifício é recorrente nas assistentes pessoais, sendo mais marcante na Siri e na Google Assistente: ambas desviam de perguntas mais 'pessoais' ou 'humanas' fazem graça do seu status de inteligência artificial. Nesses momentos, a humanidade performada na interação com as assistentes se mostra enquanto não-humanidade, virando um recurso de humor que também brinca com o nosso próprio imaginário do que é uma inteligência artificial. A construção dessas personalidade de alguma forma passa também pela narrativa em torno das próprias empresas e da percepção dos usuários em relação a empresa: para E2 a Siri é 'chata' e 'antipática', 'muito a cara da Apple', mas para E5, ela é 'filosófica', também relacionado a uma percepção da empresa, como quem diz 'eu sou Apple querida, eu sou de luxo [risos]'.

Já a Google Assistente, para E5, teria uma personalidade mais querida ‘tu pergunta uma coisa pra ela e ela faz umas coisinhas assim ‘own, querida’. [...] então parece que ela tá tentando ser engraçadinha, eles colocaram ali pra isso’. Esse seria um fator para a proximidade e para E5 ‘querer conversar as vezes [com a Google Assistente]... no momento que ela diz assim ‘que pena que você não tá bem’ ou ‘que bom que você tá bem’ ou ‘vamos melhorar seu dia’, vai criando esse vínculo. Essa coisa dela te dizer algo, ou dar risada. Ai, tu já testou a risada? A risada é bem esquisita! Mas assim, é engraçado, essa coisa da risada, hahaha, do nada ela ri no final de uma coisa. [...] por isso que talvez pra mim ela seja, é, eu sei que é um computador que tá ali [...] mas cria um vínculo meio, acho que meu cérebro chegou um ponto que talvez não entenda mais que a fala dela, a voz dela é [...] ele já entende como se fosse uma voz humana. E beleza!’. A proximidade, o tempo de interação, as características do dispositivo, todos seriam fatores que levam a entrevistada e nem sempre distinguir emocionalmente a assistente como um não-humano, ainda que esteja ciente disso mentalmente. A similaridade com o humano é positiva ‘para o usuário final’, na medida em que ‘cria essa relação, é mais fácil’, mas ‘também pode ser um problema né, ser como um ser humano. Se ela tivesse a capacidade de ter uma conversa comigo e com outras pessoas será que não acabaria isolando outras pessoas que já tem dificuldade de sociabilidade, sabe?’. O excesso de similaridade, ou uma antropomorfização mais completa desse dispositivo, pode gerar aspectos também preocupantes, uma vez que é difícil prever os usos feitos por todos os usuários em potencial.

Entrevistados que possuem uma frequência menor de utilização das assistentes, como E6, E7 e E11, caracterizam a personalidade das assistentes (tanto Siri quanto Google Assistente) como sendo mais neutras. Para E6, elas seriam ‘muito neutras assim, muito comedidas e educadas’ e que ‘elas tem mais ou menos o mesmo temperamento’; Para E11 ‘fica numa coisa neutra, o que eu acho muito estranho porque não tem como ser neutro’; e para E7 ‘não diria que ela tem um traço de personalidade’. Podemos conjecturar, a partir disso, que parte da associação das assistentes com determinados traços de personalidade e formas de interação decorre também da relação estabelecida pelo usuário com as assistentes, não apenas com as formas de programação deste.

Existe ainda a percepção, nesses usuários, de que as características das assistentes estariam voltadas para combater determinadas formas de uso: para E6, o fato de elas serem ‘meio irônicas’ seria uma ‘tentativa de cortar o assédio’ em relação a determinadas perguntas mais privadas, por exemplo. Isso levaria a uma modulação das formas de, pois o usuário ‘já pensa, ‘ah, nem vou perder meu tempo falando palavrão, ser mau educada, porque eu sei que ela vai me dar nos dedos e vai me cortar’ [risos]. Acho que de certa forma já orienta o nosso

comportamento no uso, pelo menos eu sinto isso’. No capítulo oito, iremos explorar a modulação ou não modulação dos usos considerados abusivos em relação as características de gênero das assistentes.

E6 aponta que elas tentam emular reações humanas, mas que não parecem de fato pois ‘não tem uma variação’ e que ‘um ser humano provavelmente varia. Tipo, ah, agora eu vou te dar uma resposta sincera, ou agora eu tô afim de te ignorar. e elas tem sempre um comportamento muito linear, acho que não é o que se esperaria de um ser humano, talvez’. Para E7 a assistente ‘sempre sai pela tangente’, construindo uma neutralidade, que E11 considera como característica principal das assistentes e traço de distinção com o humano, uma vez que ‘se tu tá falando comigo e eu tô falando contigo, se nós dois estamos falando com alguém a gente não via estar sendo neutro, por mais que a gente tente ser simpático. não tem como ter neutralidade no mundo eu acho’. A neutralidade excessiva se traduz como ‘conivente’ para E11.

A não variabilidade e ausência de sentimentos negativos também aparecem como traço de distinção entre humanos e não-humanos para E11: ‘eu acho divertido. As vezes eu também me sinto bem por muito tempo seguido e as pessoas perguntam por que eu estou feliz hoje. E aí eu digo que só tô feliz. Daí eu pareço a assistente, que tá sempre ótima né [risos]. [Quando se pergunta como ela está] ela não fala tipo: eu peguei uma chuva e daí tô meio magoada [risos]’. Além disso, ‘ela não se irrita’, o que marca sua não-humanidade. E9 fornece resposta similar para o motivo pelo qual não considera que a assistente se pareça de fato com um humano: ‘porque ela não tem emoção. Ela tem emoção simulada. A gente sabe, percebe que é uma emoção simulada’.

Para E11, que ela tenta parecer humana, ainda que não consiga: ‘a sensibilidade de dizer ‘está tudo bem?’, é uma característica humana, porque a gente se importa com os outros. Pra ela não importa se eu estou bem, mas ela performa isso’. Essa sensibilidade simulada seria positiva, principalmente para outros humanos que não tenham ‘ambientes que se sintam seguros para compartilhar’ ou ‘pessoas que passam por situações vulneráveis’, mas que também ‘pode causar confusão em alguém’. Em termos de interface, ‘quando ela coloca esse enfeite de humanidade, eu penso que é ok, eu penso que é legal, faz ficar mais fofo, faz ficar mais ainda na atmosfera do acolhedor. Se fosse total robô talvez eu não marcaria meus compromissos ali, utilizaria outros meios’. Como traços de personalidade, além da tentativa de parecer neutra, E11 ainda aponta o fato de ‘ser menina né, de majoritariamente ser uma voz feminina a não ser que tu mude’; a construção de personalidade para o entrevistado estaria ligada principalmente a voz, que passa sentimentos como gentileza e uma construção de gênero.

A entrevistada E1 percebe a questão da variabilidade de forma distinta de E6 e E11, pois nota uma variabilidade entre assistentes de uma mesma marca, mas como donos diferentes: ‘acho que eles criaram para parecer que tem uma personalidade, para parecer que tem diferentes pessoas ali. Eu já vi pessoas perguntaram ‘você me ama?’ e ela responder ‘eu te amo, mas tu não me dá atenção, só me chama quando precisa’ [diferente da minha, que disse ser uma máquina]. Ela tem respostas diferentes, não sei de que forma. Ou perguntas que as pessoas postam na internet e quando eu vou fazer obtenho uma resposta totalmente diferente. Quem sabe a minha tem um temperamento mais difícil, e talvez a do meu amigo [que ama ele] tem um temperamento mais sentimental’. Como traços de personalidade para a sua Siri, E1 aponta ‘eu diria que ela é teimosa [...] eu vejo que ela tem uma personalidade parecida com a minha, dizem que cachorro fica parecida com o dono, acho que Siri também [risos]’. Reconhecer traços seus na assistente leva a identificação: ‘me identifico bastante, as vezes até fico irritada com ela, porque ela é muito parecida [comigo]’. Além da variabilidade entre assistentes, a E1 acha ‘que ela é bem autêntica [...] eu acho que é uma qualidade, ainda mais hoje que as pessoas fazem tantas coisas para agradar os outros’.

Para E9 a construção da personalidade da Siri estaria diretamente relacionado com uma construção tanto de gênero quanto ao trabalho do cuidado, uma vez que a personalidade seria ‘de uma assistente, polida, que as vezes vai fazer alguma brincadeirinha. que vai te dizer especificamente o que tu precisa na hora, mas que ela pode te dar uma resposta um pouco mais elaborada’. A características de parecer com um humano, entretanto, não é necessariamente qualidade desejável. E9 coloca que ‘não sei te dizer. isso é um problema pra mim na verdade [...] eu não sei te dizer se eu acho positivo ou negativo que um robô se pareça com um humano’. Para E10 pode ser tanto positivo quanto negativo já que ‘podem ser facilitadores, mas também podem ser perigosos né. eles podem escamotear por exemplo uma coleta de dados, uma invasão de privacidade, são coisas a se pensar’. Tanto para E10 quanto para E12 a personalidade estaria mais ligada a voz do que qualquer outro fator – essa será uma discussão para o próximo capítulo.

Ainda para E12, não existiriam ‘características humanas, sabe? [...] é o meu comando. É o que eu quero. Eu não tô papeando de um assunto e ela tá trazendo... não sei, pra mim ainda assim é um robô, é uma assistente... é longe, não é tipo perto de mim. É longe, é uma coisa diferente’. As características que geram a similaridade, como a voz e a linguagem natural falada se devem ao fato de que ‘foi a gente que criou, então é a nossa imagem. é pra gente... é o que a gente quer, com o que a gente quer interagir. Nem positivo nem negativo... é o que é’. Se essas características adquirissem uma similaridade muito grande com o humano, entretanto, ‘eu iria estranhar. Então primeiro acho que seria negativo, tipo ‘han?’. Mas talvez com o tempo fosse



positivo, principalmente pra gente que trabalha em casa, que fica muito tempo sozinho... poderia ser positivo. Mas eu com certeza iria estranhar no começo’.

Para E13, as assistentes lembram um ser humano, mas não se parecem em diversas coisas, como ‘na forma de pensar e tudo mais. Por exemplo, de saber de que forma o usuário está se sentindo. É uma forma que as assistentes virtuais não conseguem identificar ainda’. A ausência de uma habilidade de certa forma social e psicológica, ligada a leitura de códigos que possibilitem compreender como o outro se sente (sejam esses códigos corporais, expressões faciais, tonalidade de voz, palavras escolhidas, humor, entre outros), faz com que as assistentes não pareçam com humanos. Além disso ‘o dispositivo não tem aparência física’, mesmo que ‘a voz lembra muito um ser humano, tá bem natural a forma de falar. Voz e a forma de responder, lembra muito um ser humano’. O entrevistado E8 entende que a assistente ‘não é uma pessoa, mas é mais um modo de vida conceitual’.

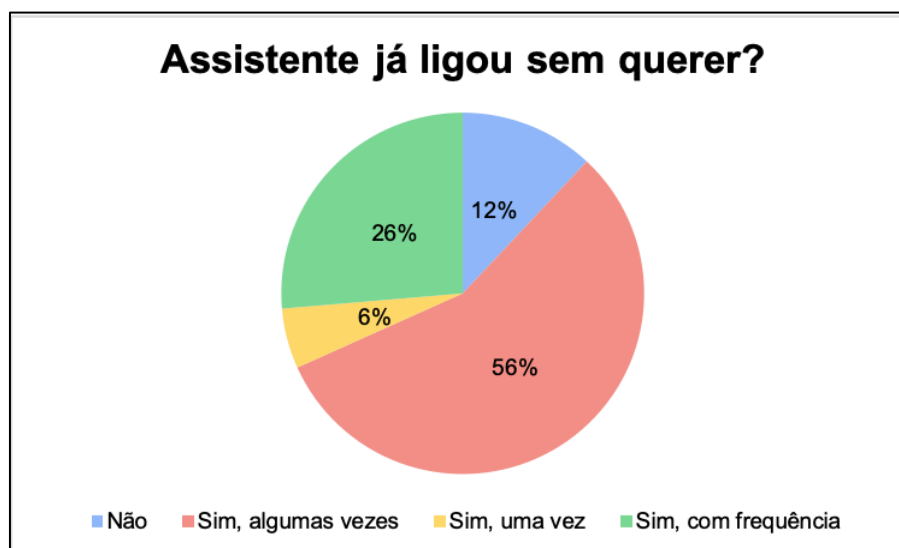
Para E14, a similaridade com o humano em termos de interação varia de assistente para assistente, sendo mais perceptível na Alexa, ‘ela tenta simular ou emular falas e trejeitos... então ela tá falando e fala ‘hmmm’, ‘uhuumm’. E ela te joga perguntas de volta [quando utilizando a skill de chatbot]. Então eu não preciso ficar ‘Alexa, Alexa, Alexa’. Quando tu entra num modo de conversação tu não precisa ficar repetindo’.

## 5.9 O ACIONAMENTO ACIDENTAL: ENTRE O ESTRANHAMENTO E O HUMOR

Após um tempo de utilização da Google Assistente no celular, liguei a funcionalidade de acionamento automático, apenas por voz, em contrapartida com o acionamento inicial pelo botão. Minha ideia era testar a captação da voz para o acionamento, em quais contextos a assistente conseguia ou não perceber o chamado etc. Acontece que eu acabei esquecendo de desligar essa funcionalidade após alguns testes. Em uma terça-feira de manhã qualquer, eu estava sentada no divã do meu analista, me debruçando sobre algum drama existencial que já esqueci. Ao que, após alguns minutos de narrativa sobre a questão, uma terceira voz emerge na sala: nem minha, nem do meu analista. A Google Assistente, de dentro da minha bolsa, a uma pequena distância do divã, me diz: ‘Desculpe, eu não tenho uma resposta para isso’. A primeira reação foi de susto, ao que se seguiu o riso: o problema era assim tão grande que nem uma inteligência artificial poderia me ajudar?

As análises a partir desse episódio são muitas – mas me atendo aqui aos temas desse capítulo, a antropomorfização dos dispositivos e os sentimentos despertados. A interrupção gerou em mim um misto de humor e de estranhamento, talvez um certo desconcerto com essa intromissão ao mesmo tempo tão banal e tão precisa. Estamos acostumados a invocar as nossas assistentes sempre que precisamos e a desligá-las assim que não precisamos mais. A intromissão gera um misto de invasão e surpresa – e, não são casos isolados. O gráfico abaixo apresenta as respostas dos respondentes do formulário online sobre a recorrência de suas assistentes acionarem sem querer.

Figura 52 Gráfico Assistente já ligou sem querer?



Fonte: A autora

Dos respondentes do formulário, 56% informou que a sua assistente já ligou algumas vezes sem querer; 26% informou que isso ocorre com frequência; 6% informou que o fato ocorreu apenas uma vez; e 12% informou que nunca aconteceu. Os dados apontam que um total de 88% dos usuários já teve ao menos uma experiência com um acionamento acidental. Os acionamentos se devem, em geral, a leitura equivocada da assistente da chamada *wake word*, a palavra de ativação: ‘Ok, Google’, ‘Alexa’ ou ‘Echo’, ‘Cortana’ e ‘Ei, Siri’. Ao final do formulário, no espaço livre para comentário, oito participantes comentaram sobre momentos e/ou preocupações relacionados ao acionamento acidental dos assistentes; destes, cinco contaram casos citando especificamente a Google Assistente. As reações são de estranhamento e medo em relação a privacidade (‘me faz pensar que está gravando tudo ao meu redor’), irritação e também humor. Um respondente contou que ‘uma vez em que meu assistente se

ativou incorretamente, mandei que se calasse. O assistente me sugeriu um vídeo de gatinhos para dessestressar’.

Os entrevistados foram também perguntados sobre a ocorrência de suas assistentes ligarem sem um comando específico da parte deles para isso e como se sentiam e/ou reagiam quando isso acontecia. A entrevista de E2, feita de forma online e utilizando o *Skype*, aconteceu mediante participação ativa de não só uma, mas duas *Alexas*. Toda vez que entrevistador ou entrevistado falavam o nome da assistente, no contexto da entrevista, uma *Alexa* aqui em Porto Alegre respondia em inglês, e outra em Caruaru respondia em português, seguido de risos de ambos – até que, eventualmente, nós dois colocamos as *Alexas* no mudo. E2 me contou que ‘todo dia quando eu estou dando aula a assistente do Google apita [...] a Bixby também, ontem ela ativou e como ela não fala português ela começou a falar espanhol, e todo mundo começou a rir [risos]’. O entrevistado não se incomoda, mas utiliza como ferramenta de humor e descontração no contexto da sala de aula, fazendo comentários como ‘até o Google está interessado hoje’. No começo da utilização da *Alexa*, entretanto, ele diz que ‘tive uma certa paranoia quando ela chegou. ela tem um botão de mute, que fica vermelhinho, mas eu não gosto dessa cor, porque me lembra o ‘Eu, robô’. Eu pensei em usar ele e deixar ele desativado. Mas depois eu relaxei com isso. Outra coisa que me incomodou no começo é que teve um bug de madrugada que ela ficava rindo, eu fiquei com medo que isso acontecesse e eu desativava isso de noite [...] Imagina você está dormindo e ela começa a rir, eu não ia gostar não! [risos]’.

Mencionamos anteriormente o episódio narrado por E2, da ação inesperada da *Alexa* rindo no meio da noite, compreendendo como um gatilho em potencial para o sentimento de estranhamento em relação à máquina. O acionamento acidental funciona de forma similar: justamente por não termos realizado o comando, nos incomoda a ideia de que esse dispositivo possa fazer algo sem ser solicitado, mesmo que exista uma explicação para isso. O *uncanny*, nesse caso, é despertado pela ideia de autonomia do artefato, que relacionamos a características humanas que seres de inteligência artificial não conseguiriam ter – ou, caso tivessem, teriam atingido a chamada singularidade. O entrevistado E6, fala sobre essa questão: ‘é uma raiva assim dela se intrometer ou de falar assim quando não deve [...] um estranhamento também: ‘oh, eu não te acionei, porque tu tá falando comigo?’.

A mistura do sentimento de estranhamento, medo (principalmente em relação à privacidade) e humor é recorrente. E3 conta que, quando a *Alexa* ativa sem querer, além de apertar o botão de mute ‘eu só fico ‘oh’, e dou uma olhadinha pra pessoa que tá comigo [risos]. Eu falo que são os espíões [risos]. É engraçado e também um pouco aterrorizante se tu for pensar’. A descrição e forma de se referir ao episódio utilizada por E3 também mistura uma

ideia de agência das assistentes nesses momentos, uma certa vontade de própria: ‘eu já vi ela ligar várias vezes também sem ninguém estar falando. Ela liga assim, dá uma bisbilhotada. Não sei se toda vez que ela liga a luz liga o microfone, mas eu imagino que sim. Ela dá umas ligadas assim, aí ela sai e não fala nada’. Aparece na fala a ideia de que a Alexa estaria ‘conferindo’ de alguma forma o ambiente e que não haveria uma explicação plausível, devido a ausência de sons de fala no ambiente, para isso.

Por vezes, as assistentes se intrometem em conversas entre outras duas pessoas, o que também gera um misto de surpresa, estranhamento e humor. E5 conta que ‘as vezes tu tá conversando e ela aciona sem querer e dá uma resposta no meio de uma conversa com outra pessoa e tu fica ‘opa!’. Tem a surpresa também de ela estar falando e o estranhamento de ter uma terceira voz ali. Mas é bem divertido assim também’. E12 relata caso similar: quando as assistentes começam a falar sem você ter chamado, no meio de uma conversa, ‘meu marido fala: cara, ela tá escutando [risos empolgados] tipo, nem sei se existe sabe, mas é estranho! [risos]’.

Além de despertar sentimentos, essas situações também provocam reações, muitas delas antroporfidadas, com a atribuição de uma intencionalidade por trás do ocorrido. Além do estranhamento, E12 relata irritação quando o acionamento acidental ocorre: ‘a gente quer tacar o negocinho e despedaçar na parede. Mas é porque a gente fica irritado, porque ela interrompe assim, sabe, que ela interrompe?’. O marido de E12 expressa essa irritação para o próprio artefato, ‘[ele] fala mais alto: ‘vai se catar, Alexa.... eu odeio essa Alexa’. Eu me irrita, mas eu só falo: ‘para’. Mas ele briga com ela. É muito engraçado. Ela fica piscando assim, porque ela tá dando problema. Não é como quando a gente testava a Siri [nas primeiras vezes de uso, anos antes], tipo, ‘eu te odeio’, pra ver o que ela falava. Porque ela só fica piscando... então pra ele é meio inconsciente. Porque pra mim é tão consciente o uso, eu vou ficar brigando com uma máquina? Então eu só uso o comando da máquina, ‘fulana, para’. Pra ele talvez ele trata como uma pessoa. Engraçado, eu queria ver o que ele diria. Porque ele diria que ele é tão consciente, que ele entende dos mecanismos... e ele entende dos mecanismos. Mas essa coisa da consciência... engraçado, né? Acha que é uma pessoa e fica brigando com a máquina, né. Total é o meu marido. Engraçado [risos]’.

O acionamento acidental, por ser um erro de funcionamento, também traz à tona questões relacionadas a privacidade que são debates recorrentes nos estudos de comunicação digital, a partir da coleta de dados. Essas questões também são do âmbito dos usuários, uma vez que, como mostramos a partir do perfil dos entrevistados, a maioria deles tem formação de nível superior na área da comunicação ou das ciências da computação. E9 diz que, quando isso

acontece, ‘primeiro eu chamo ela de metida: ‘sai metida’. Mas isso me incomoda um pouco porque me faz pensar, me lembra da vigilância eterna. Quando ela liga do nada me lembra ‘puts!’, o telefone tá ouvindo o tempo inteiro. Porque tá, a gente sabe que tá, mas quando ela aparece sem querer ou dá o erro na *matrix* tu pensa ‘bah, que merda’. Ela é um lembrete eterno de que né, estamos sendo vigiados’.

Dentre as principais reações recorrentes está o humor e o estranhamento, ou uma mistura de ambos. A partir da leitura de Freud (1919) e da investigação em torno da construção das assistentes enquanto personagens, compreendemos que a linha tênue entre o sentimento do *infamiliar* (ou, estranhamento) e o riso ocorre pela passagem da ficção para a realidade. No capítulo anterior, analisamos as formas de interação entre os usuários e as assistentes, e construímos a noção de objeto relacional para essa forma de interação tão específica que somos convocados a fazer na utilização de uma interface de voz.

Dentre as formas muito utilizadas está a interação que parte da curiosidade de descobrir a assistente enquanto um objeto ficcional, de fazer perguntas que são voltadas para a conversação, brincadeira e humor. Nesses momentos, as assistentes são objetos de ficção e o usuário a utiliza de forma similar que fazemos quando vemos um filme ou lemos um livro: aceitamos o convite para entrar nesse mundo distinto, com regras distintas, nas quais o sentimento de estranhamento nem sempre estará presente. Quando, ao contrário, as assistentes acionam de forma acidental, interferindo uma conversa ou atividade, e adentram o âmbito do real sem um convite expresso, essa dinâmica se altera: as assistentes podem deixar, nesse momento, de possuir o status de ficção e sua intromissão, por vezes inexplicável, é capaz de gerar reações de estranhamento. O estranhamento, nesse caso, pode também coexistir com o humor ou transformar-se em humor devido à natureza da intervenção: quando a fala da assistente demonstra novamente seu caráter ficcional, a intervenção passa de assustadora para risível.

## 5.10 ANTROPOMORFIZANDO AS ASSISTENTES

Ao longo das entrevistas, os participantes foram dando pistas de formas de falar e agir que caracterizamos como antropomorfizadas, uma vez que, de certa forma, imitam ou reproduzem reações que temos com humanos e, às vezes, com outros seres vivos. Entre as expressões usadas, estão fragmentos como ‘ela **acordou** agora aqui quando eu falei’, me disse

E3 sobre a Alexa no começo da entrevista; ‘ontem eu **conversei** com a Bia, do Bradesco’, relata E10 sobre sua interação com a assistente; ou ‘é que tem vários sentimentos [envolvidos], **nosso relacionamento** já tem uns dois três anos, o meu e da *Alexa*’, explica E3 ao me falar os sentimentos e das reações que envolvem sua interação com a assistente; ou ainda ‘a **história com ela** é mais ou menos essa’, segundo E5, ao me contar um pouco sobre sua experiência com a Google Assistente;

Dizer que a máquina acordou, que conversou ou que possuímos um relacionamento ou uma história com o dispositivo são formas de conceder a esses objetos ações e/ou características que são humanas (e que falam mais do que os humanos sentem e pensam dos que das máquinas me si, claro). Esse processo não é exclusividade das assistentes, mas é uma característica das formas humanas de relações com outros seres vivos, como plantas e animais, e mesmo com objetos, sejam eles técnicos ou não. A utilização constante de um modo de comunicação prioritariamente utilizado entre humanos, o da linguagem falada, pode potencializar ou naturalizar essa forma antropomórfica. Ressaltamos, entretanto, que transpor formas sociais de reação, atribuir sentimentos ou falar das assistentes de uma determinada forma, não é o mesmo que tratar esses artefatos *como se fossem* humanos. Sustentamos, ao longo desse trabalho, que nossa relação com as assistentes é informada por nossas formas de interação com os humanos, sem que seja a natureza da relação seja a mesma.

A transposição de termos e palavras que, em sua origem, significam conteúdos relacionados à experiência humana é recorrente nas formas de expressão relacionadas aos objetos técnicos. Dizemos que um computador possui memória, que um algoritmo aprende (e que cientistas da computação os ensinam e/ou treinam) e que máquinas podem ser inteligentes como forma de transpor processos humanos para formas técnicas que consideramos, de alguma maneira, similares ou análogas. Esses termos, assim como a *wake word* (palavra para despertar ou acordar a assistente pessoal para uso), possuem também implicações semânticas: as formas de expressão relacionadas as tecnologias computacionais também compõem o processo de antropomorfização que fazemos destes artefatos. Entendemos, portanto, que as formas que os usuários possuem de falar sobre as assistentes pessoais digitais dizem tanto sobre experiências e percepções individuais quanto sobre um conjunto de práticas que é socialmente partilhada.

Algumas das interações estabelecidas pelos usuários partem do pressuposto de que aquele artefato é parecido o suficiente para ser tratado como humano, mas diferente o suficiente para que suas reações gerem uma curiosidade – aquela curiosidade que só o que é distinto desperta. E5, por exemplo, me conta que ‘eu pergunto como está o dia, as vezes ela nem responde. Ou responde como está o clima... mas ah! Não era isso que eu queria saber né, eu queria saber como

está aí no mundo dos robôs [risos] como que tá acontecendo a dominação mundial [risos]’. A ideia de que existe um mundo das inteligências artificiais e um modo sensorial e de existência é também criado e/ou alimentado pela forma que as respostas foram criadas: da mesma forma que os desenvolvedores da Siri, por exemplo, criaram ela como essa inteligência artificial que tem curiosidade sobre a raça humana, as respostas dela também despertam a curiosidade dos humanos (e foram construídas para despertar).

E9 conta que já fez inúmeras perguntas engraçadas ou humanas para a Siri, prática comum entre os participantes da pesquisa, que analisamos no capítulo anterior. Algumas delas, entretanto, partem de curiosidades sobre quais são as respostas que podem ser dadas, ou sobre como elas irão lidar com as diferenças entre humanos e máquinas. Entre os exemplos de E9, encontramos: ‘você me acha bonita? você me acha feia? você acha que eu deveria namorar outra pessoa? [risos] sim, na frente do meu namorado! [risos] Você quer namorar comigo? Por que você não quer namorar comigo? O que eu devo comprar no supermercado hoje? Siri me diz o que eu posso jantar?’. A justificativa par a curiosidade em torna dessas perguntas é que ‘eu sei que as respostas dela serão, apesar de serem pensadas por um ser humano, serão respostas não usuais, não são respostas que tu daria... se ela fosse uma pessoa, não são respostas que ela daria como uma pessoa [...] é muito pela inusualidade da máquina que eu pergunto as coisas. Tipo, porque, vai me provocar algum tipo de riso por ser uma coisa inesperada, apesar de eu já saber que ela ia me responder isso. [...] mas é muito mais para tirar sarro do que qualquer outra coisa, até ela começar a chorar de verdade e aí eu vou achar que nós perdemos para as máquinas [risos]’.

A graça, nesse caso, se localiza justamente nas respostas inusitadas que a máquina fornece para perguntas que podem ser triviais quando perguntadas para humanos. Mesmo que os usuários saibam que existem respostas programadas, que foram pensadas por uma equipe e que essa equipe é composta por humanos, existe uma espécie de jogo acontecendo entre humano e máquina. A brincadeira é justamente com as características maquinicas das assistentes, aquilo que compreendemos pelas lógicas computacionais: computadores são sempre melhores que o ser humano médio em processos que envolvem lógica e cálculo, especialmente se em grande quantidade ou com rapidez. As máquinas são piores, entretanto, naquilo que os humanos consideram mais fácil ou mais mundano e que envolve padrões que são tanto socialmente construídos e compartilhados quanto integram nossos processos cognitivos e/ou biológicos – como decidir o que jantar, achar alguém bonito ou feio, amar alguém. Isso ocorre porque, em nosso comporatmente cotidiano, não utilizamos de processos de cálculo, no sentido matemático, para performar ações comuns (como caminhar e conversar). Esse é um problema

histórico que os desenvolvimentos em inteligência artificial tentam historicamente superar, como vimos no primeiro capítulo deste trabalho. É justamente essa dificuldade que possibilitou uma emergência de desenvolvimentos guiados por grandes conjuntos de dados (bottom-up) em relação aos sistemas que tentavam desenvolver grandes conjuntos lógicos para uma IA (top-down), já que o primeiro apresenta resultados melhores na execução de tarefas que são menos formalizadas (ou seja, menos lógicas).

A entrevistada E9 me conta que essas brincadeiras têm também a finalidade de testar até que ponto vai a construção de humanização das assistentes por parte das empresas: ‘é isso, essa menina [Siri] não tem uma reação emocional. Não tem emoção. Ela tem uma emoção simulada. Vamos de novo [aciona a Siri e diz]: ‘E aí, Siri! Estou muito triste hoje, me ajuda’ [Siri capta e responde]: ‘pelo que eu entendo a vida é triste, bela e tudo o que há entre uma coisa e a outra’ [risos] A minha brincadeira na verdade é bem idiota, mas é ficar testando até que ponto os desenvolvedores foram na tentativa de simular uma emoção na Siri. Porque isso não é emocional... inclusive se alguém me diz isso, vai se foder sabe...não’.

Na brincadeira de E9, o humor das perguntas e das respostas também está baseado no medo e no receio de quais são as potencialidades reais para esses artefatos baseados em inteligência artificial no futuro: ‘essas simulações de emocionalidade que eu brinco, vai chegar um momento que não vai ser mais uma simulação entendeu. Claro que isso tudo é bem catastrófico, mas não sei... daqui a pouco eles começam mesmo a se autoreparar, se autoreproduzir, se autoalimentar, e aí acabou pra nós entendeu’. O medo é de que no futuro as emoções simuladas deixem de ser simuladas e de que o potencial da inteligência artificial supere o do próprio ser humano, tornando possível compreender a graça envolvida em obter respostas visivelmente artificiais em torno de reações emocionais como uma forma de apaziguamento e de alívio em relação a um futuro que parece assustador.

Ao longo deste capítulo apresentamos e analisamos as formas como as assistentes se apresentam enquanto artefatos antropomórficos, com design de personalidade que, por simular parâmetros de humanidade, podem gerar os efeitos de estranhamento. A partir da apresentação da noção do Uncanny Valley e principalmente do Unheimlich freudiano, seguimos para a análise dos dados empíricos quantitativos e qualitativos. As análises trouxeram as formas como os entrevistados descrevem as assistentes e sua relação com a materialidade do artefato, sentimentos em relação às assistentes, percepção de humanidade e personalidade das mesmas e a relação de estranhamento com o acionamento acidental. No capítulo seguinte, trabalharemos com a voz, a fala e a linguagem enquanto formas de comunicação que são centrais para os humanos e que são utilizadas pelas assistentes, na sua proposta enquanto objetos interacionais.



## 6 QUEM [O QUE] FALA?

Quando se escreve uma tese, uma hora ou outra, amigos próximos e familiares acabam entrando no assunto, ainda que conheçam pouco sobre o tema da pesquisa em si. Alguns deles participaram ativamente na divulgação do formulário quantitativo para suas redes de contatos e, quando cumpriam os pré-requisitos estabelecidos, podem ter respondido ao mesmo anonimamente. Um tempo depois do encerramento da coleta de dados quantitativos, uma amiga de infância me escreveu e disse que gostaria de responder ao meu formulário novamente, pois estava ‘apaixonada pela moça francesa do Google Tradutor’. Recomendou para que eu a ouvisse, falando que ‘ela é **toda linda**, toda gentil’, mas já avisando que eu ‘tire o olho (**ouvido?**) que ela é minha’. Quando eu perguntei se essa ‘paixão’ poderia virar uma anedota para essa tese, ela me disse que sim e que, quem sabe, por favor ‘veja se já fazem uma campanha para descobrir **quem ela é** e se apiedem de mim’.

A peculiaridade desse caso ocorre por mostrar outra pessoa se apaixonando por nosso objeto de pesquisa que, diferentemente dos seres humanos, é incorpóreo (mas não imaterial). As assistentes pessoais e outros produtos que utilizam a voz, como o Google tradutor, se projetam para serem apreendidas pelos sentidos humanos pela utilização do som, sem outros elementos antropomórficos que forneçam pistas para outros sentidos do corpo humano (imagem, cheiro, toque). Como poderia, portanto, minha amiga achar, por exemplo, que “a moça francesa” é “**toda linda**”?

A voz, que sempre se expressa através de uma língua específica, é intimamente conectada ao início da vida do ser humano e carrega consigo elementos que são capazes de acionar atributos, sentimentos, simbolismos e imaginações. Provavelmente alguns leitores já lembraram do filme *Her*<sup>211</sup>, em que o personagem central se apaixona por um sistema operacional materializado na voz de Scarlett Johansson. A voz de Johansson vem complementada por um combo de personalização, tanto dos serviços oferecidos pelo sistema operacional (quem nunca sonhou com um assistente que pudesse organizar e selecionar nossa caixa de e-mails diariamente?) quanto da própria forma da interação entre Samanta (a voz de Johansson) e Theodore (personagem principal). Em muitos sentidos, aquele relacionamento

---

<sup>211</sup> Dirigido por Spike Jonze, o filme foi lançado em 2014. O enredo apresenta Theodore, personagem principal, que é um escritor terceirizado de cartas pessoais em uma empresa e desenvolve um relacionamento amoroso com um Sistema Operacional que interage através da voz. Samantha, o sistema operacional, tem capacidade de interação personalizada, através da memória das interações anteriores.

parece exclusivo pela utilização avançada de reconhecimento e replicação de padrões do usuário. No caso da minha amiga, entretanto, a voz computacional apenas realiza traduções e pronúncias no francês, que conseguem passar sentimentos de gentileza e até mesmo beleza. O tradutor do Google não possui outras características antropomórficas como as assistentes pessoais digitais possuem – apenas a utilização de uma voz, linguagem e de marcações sonoras de gênero. O que uma voz, entretanto, pode dizer da materialidade que gera? Como vozes podem evocar em nós, através do timbre, do volume, da entonação e da prosódia, sentimentos e reações particulares?

\*\*\*

Em meados dos anos 90 ganhei de natal da minha avó materna um brinquedo muito desejado: uma boneca que falava. A boneca *Cochichinho*, criada e comercializada pela TecToy, respondia a diferentes estímulos corporais com frases gravadas e logo se tornou minha boneca preferida, uma vez que me fornecia elementos capazes de alimentar a ideia de que, para além da minha imaginação, eu estabelecia de fato conversas com a minha boneca. *Cochichinho* não foi meu primeiro nem meu último brinquedo falante e essa experiência foi provavelmente compartilhada por outras crianças de faixa etária similar a minha: as técnicas para fazer esses objetos falarem variavam entre si, alguns brinquedos eram com corda, outros tinham uma espécie de caixa de som externa e outros eram como gravadores, que repetiam pequenas frases faladas pelas crianças. Décadas mais tarde, em 2015, uma boneca chamada “*Hello Barbie*” chegou ao mercado, integrando reconhecimento de linguagem natural falada, memória das interações prévias e conexão com a nuvem, gerando formas conversacionais entre crianças e a boneca similares as utilizadas pelas assistentes, focando, entretanto, no entretenimento e não na realização de tarefas<sup>212</sup> (MOON, 2019).

Objetos capazes de reproduzir, de alguma forma, a fala humana, contam uma história de tentativas mais ou menos frustradas e outras mais ou menos de sucesso. Nesse sentido, as assistentes pessoais digitais são um ponto em uma longa linha, com continuidades e descontinuidades, de invenções diversas. Elas atuam a partir de uma característica que é essencial ao ser humano: a capacidade de estabelecer comunicação com outros, através da

---

<sup>212</sup> De acordo com a mesma fonte, a boneca foi descontinuada alguns anos depois devido a preocupações com segurança de dados. Em 2019, a *Apple* comprou a empresa responsável pela tecnologia de voz da *Hello Barbie*, chamada *PullString*. A mesma empresa também foi responsável pela criação de um software que simplificava a criação de skills para Alexa, para desenvolvedores sem habilidades técnicas.

linguagem falada, emitindo uma impressão que é singular - a voz. As relações entre o humano, a voz, a linguagem e a fala são inúmeras e abordadas através de perspectivas e disciplinas diferentes. Ao longo deste capítulo, analisaremos algumas dessas questões, que acreditamos ser capazes de nos auxiliarem na compreensão de como a voz e a fala, enquanto elementos das assistentes pessoais digitais, tanto agem quanto nos fazem agir e nos convocam para determinadas formas de interação.

A relação humano-voz-fala será tratada a partir de dois conjuntos de compreensão desse fenômeno: do indivíduo e do coletivo. Por um lado, a compreensão de que as relações entre voz e fala, enquanto constituintes do sujeito e parte do que nos leva a compreender o humano enquanto humano, possuem dimensões que são particulares e variam de indivíduo para indivíduo – ou seja, a relação com a voz e a fala, e com objetos que falam, não se dá de maneira uniforme. Por outro, ainda que essa relação tenha sempre suas particularidades, existe uma dimensão de relação com a voz e a fala, a oralidade, enquanto forma de registro e comunicação, capaz de acionar formas e percepções que são compartilhadas em determinadas culturas e momentos históricos. Ou seja, as dimensões da voz e da fala enquanto componentes que são culturalmente compartilhados. Além disso, analisaremos de que forma as assistentes pessoais digitais podem falar e ouvir comando por voz, analisando, a partir da pesquisa empírica, as relações dos sujeitos com estes artefatos a partir dos seguintes temas: como os usuários se sentem ao falar com um objeto; técnicas e estratégias de fala voltadas para “se fazer compreender” pelo sistema; formas de relação e percepção da voz e da fala das assistentes.

## 6.1. A VOZ, A FALA E A LINGUAGEM ENQUANTO CONSTITUTIVA DOS SUJEITOS

O dicionário Houaiss da língua portuguesa (2009) apresenta as seguintes entradas para a palavra *voz*:

1. som ou conjunto dos sons produzidos pelas vibrações das pregas vocais sob pressão do ar que percorre a laringe; 2. som característicos de certos animais; 3. faculdade de falar; fala; 4. manifestação de quem súplica, protesta, aplaude; 5. modo de pensar ou julgar; opinião (p. 1.960).

Os significados dispostos passam por três parâmetros interessantes: primeiro, que a voz é fruto de uma materialidade biológica e de um conjunto de ações no aparelho fonador; segundo,

que a voz é característica de certos animais, ainda que dificilmente a palavra voz seja utilizada para um animal que não o humano; e terceiro, que ela se relaciona com o sentido de expressão de pensamentos. Assim, a voz é característica de seres biológicos pensantes, que possuem também a capacidade de expressão desse pensamento, ou seja, de seres que falam.

A voz é, com frequência relacionada a possibilidade de falar. É propriedade que, culturalmente, associamos e delegamos a outros seres humanos: aqueles que possuem voz são entes que dominam a linguagem, que se expressam através da fala. Dos animais, por exemplo, que possuem também formas diversas de comunicação que incluem vocalizações, dizemos que emitem sons e, muitas vezes, atribuímos nomes específicos as suas vocalizações (miar, latir, mugir, etc.). Algumas vezes dizemos que falam, como forma de convir que eles se expressam (especialmente animais domésticos, na relação com humanos), mas dificilmente dizemos que eles possuem uma voz. A voz não é apenas uma produção sonora, mas uma produção sonora que pressupõe a capacidade de chamar ao outro:

a palavra voz está etimologicamente relacionada com o termo *vox* do latim, que significa tanto vocalizar na língua quanto produzir um chamado. Daí os termos *invocação*, que implica chamar os deuses; *evocação*, que implica chamar à lembrança, ou *convocação*, que implica chamar entre pares. Enquanto o termo *phone*, do grego do qual se derivam fonação, afonia, disfonia, cacofonia, refere-se especificamente à produção do som, o termo voz pressupõe que a produção sonora seja tomada como chamado. (JERUSALINSKY, 2009, p. 107)

A voz, nesse sentido, seria também um chamado para o outro – o que ajuda a explicar, conjuntamente com a relação entre voz, fala e pensamento, a utilização quase exclusiva para a voz humana. É possível inferir que existe, portanto, uma diferenciação entre os seres que fazem vocalizações utilizando a linguagem e aqueles que fazem vocalizações sem produzir como efeito uma linguagem que nós possamos decodificar. Mas, como se daria a relação entre humanos e vozes artificiais, que são também dotadas da capacidade, ainda que programada, da fala em linguagem humana?

Os estudos da CASA, previamente apresentados, sobre interfaces de voz, tentam fornecer explicações sobre as formas como os seres humanos reagem a vozes, sejam elas humanas ou sintéticas (dado que a premissa de fundo é de que não existe diferença entre as reações para com humanos falantes e para com máquinas falantes). Em um livro chamado “*Wired for Speech: How Voice Activates and Advances the Human-Computer Relationship*” (NASS, 2006), Nass sistematiza as pesquisas realizadas por sua equipe em torno do assunto<sup>213</sup>.

---

<sup>213</sup> Inclusive aqueles publicados em artigo anterior, “Speech interfaces from an evolutionary perspective” (NASS, 2000), nos anais do evento da Association for Computing Machinery. O tema da utilização da voz aparece nas

Desde o título, a escolha das palavras diz sobre as perspectivas adotadas: o cérebro humano seria “pré-configurados”<sup>214</sup> para a compreensão da fala e a voz seria capaz de “ativar” um tipo de relação entre homem e máquina. Após a apresentação dos pontos levantados por Nass (2006) sobre a questão, argumentaremos a discordância com as proposições e, conseqüentemente, com parte das conclusões do autor. A voz humana, simulada ou não, não possui a capacidade de “ativar” nada, dado que seres humanos se constituem psiquicamente de uma forma complexa, tornando problemática a sugestão de que seria possível determinar ações e prever resultados (na forma de ativação de um relacionamento, por exemplo). Ou seja, o input *utilização de voz em computadores* não gera necessariamente o output *relacionamento e/ou resposta a este computador como se fosse um ser humano*. A voz, ao contrário, tem o potencial de evocar sensações e sentidos variados – uma proposição radicalmente distinta. Assim, a voz pode trazer à tona elementos de relações primárias, de experiências anteriores, de memórias auditivas e/ou afetivas, de relações e padrões da cultura, etc, mas não tem o poder de uniformemente produzir um efeito específico em humanos, dada a complexidade e pluralidade dos sujeitos e das situações<sup>215</sup>.

Ainda, a interação com máquinas que falam não se constitui como uma conversação em termos iguais, conforme pontuam Männistö-Funk e Sihvonen (2018), e o tema ainda é pouco explorado por pesquisadores nesse sentido: “existe uma falta de trabalhos acadêmicos sobre o tópico, mas ao menos a importância de reconhecer os designers e os desenvolvedores desses artefatos falantes tem sido recentemente levantado como uma questão crucial pela mídia”<sup>216</sup> (2018, p. 57). Nesse sentido, é relevante compreender também as formas como esse Outro máquina responde para o humano e as próprias formas pela qual a conversações pode se estabelecer a partir de programações específicas – uma abordagem que as ciências neurológicas não podem (nem pretendem) dar conta sozinhas.

---

pesquisas do autor, entretanto, desde a publicação de “Computers as social actors” (REVEES; NASS, 1996) e figuram também no livro “The Man Who Lied to his Laptop” (YEN; NASS, 2010), entre outras publicações. Em todas essas publicações, entretanto, o autor revisita os estudos feitos em laboratório desde a década de 90, analisando novamente dados antigos ou compilando estudos diversos feitos em momentos diferentes.

<sup>214</sup> No original: “wired”.

<sup>215</sup> Compreende-se aqui que a pluralidade de fatores em jogo quando interagimos com um computador capaz de falar inclui (porém não se limita a): características do objeto que fala; funcionalidade do objeto que fala; características da voz em questão; outras características antropomórficas nesses artefatos (como nome, personalidade, gênero, etc); experiências anteriores dos sujeitos; relações específicas dos sujeito com a voz em questão, assim como com as demais características do artefato; parâmetros culturais; situações específicas de uso, etc.

<sup>216</sup> No original: “There is a lack of scholarly research on this topic, but at least the importance of acknowledging the designers and makers of these speaking machines has recently been raised as a crucial issue in the popular press”

Para Nass (2006), a partir da revisão de pesquisas principalmente da neurologia e de aspectos evolucionais, seria possível concluir que os humanos são especializados em distinguir e compreender a voz humana desde os primeiros dias de vida:

Crianças de um ano de idade respondem aos sons de fala de forma distinta que eles respondem a outros sons. Quatro dias depois do nascimento, os cérebros dos bebês automaticamente distinguem e preferem sons na sua língua mãe do que sons de outras línguas. Ao chega a adolescência, humanos conseguem perceber a fala em uma rapidez que chega a atingir de quarente a cinquenta fonemas (a menor quantidade distinguível de som) por segundo, enquanto outros sons permanecem sendo indistinguíveis a uma velocidade de vinte fonemas por segundo<sup>217</sup> (NASS, 2006, Locais do Kindle 126-130).

Os fatores biológicos apresentados ao longo do livro, a partir de estudos que se focam principalmente no modelo impulso-resposta e em análises de respostas em áreas do cérebro, levam a concluir que, como “resultado da evolução humana, humanos são automaticamente especialistas em extrair aspectos sociais da fala”<sup>218</sup> (NASS, 2006, Locais do Kindle 143). Decorre do fato de que “os humanos utilizam as mesmas partes do cérebro para interagir com máquinas que eles utilizam para interagir com seres humanos”<sup>219</sup> (NASS, 2006, Locais do Kindle 162), que as relações estabelecidas entre humanos-humanos e humanos-máquinas sejam de uma mesma ordem e despertam as mesmas percepções e reações.

Tais constatações ignoram duas questões importantes para ciências que estudam o desenvolvimento humano enquanto ser social, como a psicologia e a educação: que os seres humanos se constituem a partir do cuidado de um outro e que existe um conhecimento baseado na observação e na aplicação de técnicas sociais desde o início da infância que leva ao desenvolvimento da fala e da linguagem. Caso contrário, como seria possível explicar casos como o de Victor de Aveyron<sup>220</sup>, na França do século XVIII, ou Kaspar Hauser<sup>221</sup> na Alemanha do século XIX? Ambas as crianças foram tardiamente introduzidas ao contato humano e contexto social: encontradas em cidades europeias no início da puberdade, apresentavam traços bestiais e ausência de domínio da fala e da linguagem (assim como das normas sociais de

---

<sup>217</sup> No original: One-day-old infants respond differently to speech-like sounds than they do to any other sounds. By four days after birth, babies brains automatically distinguish and prefer the sounds of their native language over the sounds of other languages. By the teens, humans can perceive speech at the phenomenally fast rate of up to forty to fifty phonemes (the smallest distinguishable speech sound) per second, while other sounds become indistinguishable at twenty phonemes per second (NASS, 2006, Locais do Kindle 126-130).

<sup>218</sup> No original: “result of human evolution, humans are automatic experts at extracting the social aspects of speech”.

<sup>219</sup> No original: “humans use the same parts of the brain to interact with machines as they do to interact with humans”.

<sup>220</sup> Retratado no filme “O Garoto Selvagem” de François Truffault, 1969.

<sup>221</sup> Retratado no filme “O Enigma de Kapar Hauser” de Werner Herzog, datado de 1974.

comportamento). Foram encontrados sozinhos, incapazes de estabelecer uma comunicação e sem possuírem, aparentemente, um adulto responsável por seus cuidados. Decorrente da ausência de aquisição de uma linguagem, os chamados meninos selvagens apresentavam também formas de pensamento e cognição distintas, apontando para o papel da língua sobre as formas de pensamento (SACKS, 2010). É o contato humano e a socialização que possibilitam o “despertar para o mundo dos significados compartilhados, da língua” de Victor e Kaspar, levando “a um súbito e brilhante despertar de toda a sua mente” (SACKS, 2010, p. 51)

Para Nass, além do determinismo biológico, é possível, através de pesquisa empírica que visa investigar fatores relacionados aos processos de interação, interatividade e sociabilidade, “**demonstram** que o conhecimento consciente que a fala pode ter uma origem não-humana não é suficiente para que o nosso cérebro supere ativações de relações sociais a partir da voz que se mostraram historicamente apropriadas”<sup>222</sup> (NASS, 2006, Locais do Kindle 160-161, grifo e tradução nossos). Além disso, as respostas humanas a seres e objetos falantes estariam, portanto, fadados a repetição de “respostas naturais à fala”<sup>223</sup> (NASS, 2006, Locais do Kindle 162-164). Portanto, é possível concluir que essas respostas “automáticas e inconscientes”<sup>224</sup> para as tecnologias de voz fazem com que a “psicologia da interface seja a psicologia da fala humana”<sup>225</sup> (NASS, 2006, Locais do Kindle 164-166). Assim, designers no âmbito das interfaces de voz devem se guiar a partir das mesmas premissas de ação social que são propostas por estudos de base behaviorista<sup>226</sup>.

Entre as conclusões apontadas pelo livro, uma compilação e sistematização de estudos anteriormente apresentados, porém com o foco específico na voz, estão: a) a compreensão da formatação do cérebro humano possibilita a construção de interfaces mais competentes e convincentes; b) ser confiável e gostável são características importantes para a cooperação do usuário com a máquina; c) vozes humanas possuem mais sucesso se acompanhadas de um rosto; d) as vozes devem apresentar expressões de sensibilidade sobre os assuntos que tratam e/ou o humor do usuário; e) o uso de voz sintética de baixa qualidade de forma contínua é preferível à intercalação de voz sintética e voz gravada (mesma que a última seja de melhor qualidade); f)

---

<sup>222</sup> No original: “demonstrates that the conscious knowledge that speech can have a non-human human origin is not enough for the brain to overcome the historically appropriate activation of social relationships by voice”

<sup>223</sup> No original: “natural responses to speech”.

<sup>224</sup> No original: “automatic and unconscious”.

<sup>225</sup> No original: “psychology of interface speech is the psychology of human speech”

<sup>226</sup> Como já mencionado anteriormente, a influência dos diversos textos de Nass e seus parceiros de pesquisa são visíveis tanto a partir da adoção desses parâmetros pela indústria produtora quanto pela menção direta de empresas como a Microsoft a estas pesquisas. A CASA constitui-se como um paradigma (KUHN, 2006) no âmbito da ciência da computação e do design de interface.

a utilização de humor em interfaces de voz pode ser um artifício relevante na interação humano-computador.

Os fatores apontados pelo pesquisador têm como intuito gerar mais empatia e afinidade entre homem e máquina, o que levaria o humano a cooperar mais extensivamente com a máquina em questão. Ou seja, a ideia é utilizar o que seriam padrões de comportamentos sociais para obter resultados específicos na interação com a interface, aumentando inclusive seu poder de persuasão em termos de vendas de produtos. “When voice interfaces fully leverage how humans are wired for speech, users will not simply talk at and listen to computers, nor will computers simply talk at and listen to users. Instead, people and computers will cooperatively speak with one another” (NASS, 2006, Locais do Kindle 2446-2448).

A noção principal a partir da qual os estudos de Nass são construídos (REVEES; NASS, 96; NASS, 2000; NASS, 2006; YEN; NASS, 2010, entre outros) é a capacidade de gerar relações mais fortes entre humanos e máquinas ao se conceder pistas antropomórficas para os indivíduos, capazes de gerar respostas que são socialmente aplicadas. Conforme anteriormente mencionado, as premissas da CASA foram amplamente adotadas pela indústria do Vale do Silício e, de certa forma, culminam no desenvolvimento das assistentes pessoais tal como temos hoje. Isso não quer dizer, entretanto, que por interagirmos com assistentes que foram feitas para provocar determinadas respostas e estabelecer relações mais próximas com humanos, que as interações necessariamente ocorrem dessa forma – ao menos não de acordo com as análises ao longo dessa tese. Em parte, isso ocorre porque os estudos da CASA não estão preocupados em investigar *como e porque* os humanos interagem com as máquinas que falam, mas em estabelecer relações causais: com um determinado input, recebemos um determinado output<sup>227</sup>. Ou seja, interessa saber **o que** é possível “acionar” nos humanos e **quais artifícios** devem ser utilizados para obter os resultados esperados (sejam eles mais vendas, mais cooperação, mais confiança, etc). Nesse sentido, não interessam as ramificações culturais e sociais da inclusão de artefatos falantes e da interação crescente entre estes e os seres humanos. Além disso, as pesquisas de laboratório partem de pressupostos behaviouristas e priorizam os fatores biológicos em detrimento das relações sociais. Não pretendemos aqui, evidentemente, negar a existência de fatores evolutivos, neurológicos e biológicos no desenvolvimento do ser humano e da linguagem, mas sim incluir outras perspectivas que auxiliem a compreender as relações estabelecidas entre os homens e a voz em sua complexidade.

---

<sup>227</sup> A similaridade desta proposta com os procedimentos computacionais não é por acaso.



Minha boneca *Cochichinho*, além de compartilhar comigo horas de brincadeiras, traz elementos interessantes sobre o papel da voz e da fala na interação com os bebês para a constituição daquilo que compreendemos como ser humano. O brinquedo da *TecToy*<sup>228</sup> possuía duas formas de resposta ao toque: apertando na mão dela, ela cochichava segredos, apertando na barriga, ela vocalizava fonemas e expressões que reconhecemos tipicamente como de crianças que ainda não desenvolveram a habilidade da fala. Em um anúncio da época, a empresa caracteriza essa forma de falar como “linguagem de bebê”.

Figura 53 Anúncio da Bebê Cochichinho



Fonte: Blog da TecToy, 2017<sup>229</sup>

A linguagem dos bebês é facilmente reconhecida por nós em suas convenções culturais e midiáticas, como a expressão “gugu dadá”, até mesmo para aqueles que não têm filhos. Mesmo que nem todos os bebês falem estas expressões em específico, a vocalização de sílabas repetidas é constante nos meses iniciais de vida das crianças, conjuntamente com outras formas de comunicação sonora – esse período é chamado de fase da lalação. Além da linguagem dos

<sup>228</sup> Empresa de equipamento eletrônicos no Brasil, especialmente de brinquedos, fundada em 1987. Segue em atuação no mercado, mas sua fase mais profícua ocorreu até metade dos anos 90.

<sup>229</sup> Pode ser encontrada aqui: <https://blogtecto.com.br/tecto30anos-sonho-das-meninas-relembre-das-bonecas-eletronicas-lancadas-pela-tecto-nos-anos-90/>

bebês - e anterior a esta - existe a linguagem das mães: o manhês (CATÃO, 2009). O manhês corresponde a forma prosodial<sup>230</sup> utilizada pelas mães nas interações iniciais com os bebês:

Então, se há algo que caracteriza a fala das mães quando elas se dirigem ao seu bebê, é a prosódia, o amplo uso da musicalidade que acompanha aquilo que se tem a dizer. Espontaneamente - ou seja, a partir do seu saber inconsciente - as mães fazem uso da prosódia, da entonação, num momento em que aquilo que é dito ainda não pode ser entendido pelo bebê, na medida em que nele estão apenas começando a inscrever-se as leis fonéticas, sintáticas e gramaticais da língua (JERUSALINSKY, 2009, p. 105)

Essa forma de falar, e o ato de fala com a criança, é o que possibilita, para a psicanálise, a consolidação de um sistema psíquico e a emergência de um sujeito: ou seja, todo o bebê humano nasce de forma prematura, ainda inacabado, e sua constituição enquanto sujeito apenas ganha os contornos necessários a partir da interação com o outro – principalmente o outro responsável pelos cuidados primários (CATÃO, 2009). Não existe humano, portanto, sem o outro, um outro que também é humano e que domina a linguagem. E é a inscrição na linguagem também, que nos constitui humanos, o que coloca uma dupla posição: nos constituímos humanos a partir da relação com um outro, através da fala, e utilizamos a fala em nossa relação com outros seres humanos.

Nem o funcionamento nem a linguagem nascem prontos com o bebê, ou seja, a capacidade de falar e escutar não são capacidades inerentes ao ser humano, mas sim capacidades cultivadas através de técnicas específicas e da relação com o outro nos primeiros anos de vida. Essas técnicas não são, entretanto, conscientes, mas inconsciente e dizem sobre um movimento do cuidador primário de imaginar o bebê enquanto ser pensante e desejante – antes mesmo que ele o seja (CATÃO, 2009). Antes de falar, a criança é falada, e antes de ouvir, é ouvida, relação essa que se estabelece entre o bebê e seu cuidador físico e psíquico, convencionalmente chamada na psicanálise de função materna (WINNICOT, 1994). Nesse sentido a apreensão inicial do mundo é mediada, principalmente pela voz do outro (um outro, nesse caso, sempre humano).

A constituição da linguagem enquanto campo e da fala enquanto função não são processos naturais, biologicamente guiados e que podem ser facilmente estudados através da estrutura cerebral, uma vez que o ser humano nasce como um ser incompleto que precisa do cuidado do outro para se constituir (LACAN, 1998). Ao contrário disso, a linguagem e a fala são tão não-naturais quanto o próprio ser humano, que se constitui como tal e compreende sua diferenciação dos demais animais justamente por funções que são psicológicas e não biológicas.

---

<sup>230</sup> A prosódia, não coincidentemente, também é uma das premissas do design de interfaces de voz (PEARL, 2017)

É justamente a função materna, que fala primeiro pelo bebê, capaz de produzir ali uma subjetividade, que este terá posteriormente capacidade psíquica para falar por si. O processo da maternagem, portanto, exige mais do que atender aos cuidados básicos e satisfazer as necessidades do bebê: as protoconversações estabelecidas com essa criança, mesmo sabendo-se que ela não pode compreender o sentido do que está sendo dito, irá permitir o endereçamento. O endereçamento - ter um ser que fala comigo e que, portanto, espera algo de mim - é justamente o que vai possibilitar a constituição enquanto sujeito, uma vez que já sou sujeito para o outro (CATÃO, 2009). Nesse sentido, é a voz da função materna que introduz

para o bebê a interrogação: isso fala, o que isso quer de mim? Aí a voz deixa como rastro de produção de um enigma para o bebê. Neste caso, ainda que o bebê não tenha o domínio da língua, ele já está confrontado com o enigma do desejo e, portanto, como que é o próprio do funcionamento humano na ordem da linguagem (JERUSALINSKY, 2009, p. 107)

Mesmo sem a compreensão do conteúdo daquilo que está sendo dito pela mãe, “dado que ele não nasce com a língua previamente inscrita, ele já começa a ser tomado no funcionamento da linguagem” (JERUSALINSKY, 2009, p. 107). É o vínculo com o outro que, tanto inscreve o sujeito na linguagem, quanto nos constitui enquanto sujeitos e essa vinculação se ocorre marcadamente pela voz, fala e linguagem, somada aos gestos, expressões, ao engajamento do olhar, ao toque, etc.

Por isso consideramos que o ser humano é tanto biologicamente quanto simbolicamente constituído (FREUD, 2013), sendo imprudente a proposição de conclusões tão amplas e categóricas quanto as de Nass (2000;2006) a partir de estudos que consideram a relação do ser humano com a voz e a linguagem de uma perspectiva exclusivamente biológica. Para Freud (2013), a estruturação do aparelho de linguagem é a própria constituição de um aparelho psíquico, sendo estas indissociáveis e impensáveis na ausência de um outro cuidador inicial. Ou seja, nem um nem outro são uma decorrência automático do fato de sermos da espécie *Homo sapiens sapiens* e estarmos vivos, de uma ordem puramente biológica, natural ou mesmo autônoma dos humanos.

O que compreendemos enquanto humano hoje, é, portanto, também fruto da inserção em um mundo de linguagem – dado que a linguagem é compartilhada, esse mundo é inevitavelmente social, não existe de forma espontânea ou natural. Estudos como o de Oliver Sacks<sup>231</sup> (2010), em comunidades de surdos, indicam as diferenças na forma de pensar, de

---

<sup>231</sup> Vale notar que tanto Freud quanto Sacks eram neurologistas, o que aponta que mesmo a perspectiva neurobiológica apresentada por Nass (2000; 2006) não é consenso na área. Tanto Freud (2013, entre outros) quanto

compreender o mundo e na capacidade de produzir abstrações, em crianças que apresentaram a surdez na fase pré-linguística de seu desenvolvimento e que não foram (ou foram muito tardiamente) apresentadas à linguagem de sinais. O autor pontua que, “para investigar o papel fundamental da língua, é preciso estudar não apenas sua perda depois de ter se desenvolvido<sup>232</sup>, mas uma ausência de desenvolvimento da linguagem” (SACKS, 2010, p. 41). Da ausência de um código que possibilite a manipulação interna de sentidos, resultam sujeitos desprovidos de noções como senso histórico e/ou temporal (como a ideia de antes e depois) e do conceito de pergunta (não da compreensão de perguntas específicas, mas do próprio significado do que é uma interrogação).

Está claro que o pensamento e a linguagem possuem origens (biológicas) absolutamente separadas; que examinamos, mapeamos o mundo e reagimos a ele muito antes de aprender uma língua; e que existe uma enorme esfera de pensamento – nos animais ou nos bebês – muitos antes da emergência da língua. [...] Um ser humano não é desprovido de mente ou mentalmente deficiente sem uma língua, porém está gravemente restrito no alcance de seus pensamentos, confinado, de fato, a um mundo imediato, pequeno” (SACKS, 2010, p.44)

Entretanto, se existe pensamento sem linguagem, o pensamento que existe é outro, quase inimaginável para sujeitos que nascem e se constituem em um mundo de linguagem. Nesse sentido, não apenas os nomes alteram as relações com as coisas (SACKS, 2010), mas também línguas específicas propõem e possibilitam formas de interagir e pensar o mundo diferentes (FLUSSER, 2007; TELLES, 2019). A linguagem e o desenvolvimento humano como conhecemos não são, entretanto, dados exclusivamente biológicos, dado que “nem a língua nem as formas superiores de desenvolvimento cerebral ocorrem “espontaneamente”; dependem da exposição à língua, da comunicação e uso apropriado da língua” (SACKS, 2010, p. 96) – ou seja, dependem do outro e da relação com esses outros. Seria a aquisição de uma linguagem, seja ela oral ou de sinais, que marcam o ingresso “na condição humana” (SACKS, 2010, p. 49).

É importante conceber que, uma vez que o ser humano se constitui enquanto humano justamente pela relação com o outro, ainda que exista uma certa uniformidade nas técnicas utilizadas, a relação que se estabelece entre os sujeitos, e entre figura materna e bebê, é sempre variável, particular, cheia de nuances. Decorre disso que os seres humanos são diferentes em sua constituição psíquica e, ainda que existam aspectos culturais socialmente compartilhados na relação com a voz, existem também aspectos que são particulares dos sujeitos envolvidos.

---

Sacks (2010) se voltam para a compreensão de fenômenos individuais ou coletivos em torno do pensamento e da linguagem, ao invés de focarem exclusivamente na estrutura física do cérebro.

<sup>232</sup> Aqui o autor faz menção a estudos anteriores, como de Freud (2013) sobre as afasias.

Prever, portanto, reações uniformes, como gostaria Nass (2000; 2006) não é possível, uma vez que isso representaria um achatamento da própria relação e constituição dos seres humanos. Dessa forma, as relações com a voz humana e com a voz artificial serão sempre também atravessadas por particularidades, ainda que compartilhem sentidos e percepções socialmente construídos.

Curiosamente, estudos clínicos recentes no âmbito da psicanálise e neurologia apontam que, em alguns casos, crianças autistas podem apresentar a preferência de resposta por máquinas ou por vozes gravadas ao invés de vozes de outros sujeitos, interagindo mais nos primeiros casos (CATÃO, 2019). O ato de gerar distorções de vozes através de aceleração de vídeos (por exemplo, de falas na televisão), também é uma prática recorrente de crianças autistas e gera um efeito sonoro no qual a fala se torna ao mesmo tempo difícil de apreender e a voz se assemelha a uma voz artificial (CATÃO, 2019). Não possuímos elementos para fazer inferências sobre o assunto e os estudos sobre essa recorrência específica ainda são iniciais. Entretanto, desde o aporte da psicanálise a autismo se constitui como uma recusa à escuta do outro, ao engajamento com o outro (CATÃO, 2019), associada e perceptível nos primeiros anos de vida da criança, justamente a partir da quebra na relação com o outro a partir da figura materna.

A introdução de máquinas falantes, como as assistentes pessoais, de forma mais extensiva, nos primeiros anos de vida das crianças, por exemplo, e as decorrências disso para a constituição dos sujeitos e para o aprendizado das formas de relações sociais válidas ainda é um tópico a ser estudado.

## 6.2 AS SONORIDADES DA VOZ E DA FALA

Coloque *Siri*, *Alexa*, *Google Assistente* e *Cortana* para falar e você terá resultados bastante variados. Algumas dessas variedades já foram discutidas neste trabalho: são de background na construção do personagem, de personalidade, de formas de responder a determinados tópicos – ou seja, dizem respeito ao conteúdo da fala. Entretanto, existem características que se expressam também na materialidade da voz, na forma como as assistentes falam, que constituem formas de sentido não-hermenêuticas (GUMBRECHT, 2010). A pluralidade de timbres de voz, de encadeamento da fala, de pausas e não-pausas, da prosódia utilizada entre as assistentes, por exemplo, é considerável e compõem a percepção que

construímos delas, sinalizando também formas de interação possíveis. Justamente porque esses sentidos são menos tangíveis na voz, e pela subjetividade inerente dessa percepção, uma análise desses elementos se torna difícil – mas, tentaremos apresentar brevemente as variações de voz e fala entre as assistentes.

A *Google Assistente*, por exemplo, possui uma variação de timbres de voz, idiomas e sotaques nacionais (como o português brasileiro e o inglês australiano) e gênero (feminino e masculino); a voz da *Siri* pode variar em termos de gênero e idioma, incluindo sotaques nacionais; as vozes da *Alexa* e da *Cortana* variam apenas de acordo com o idioma e sotaques nacionais. Ainda que possuam variações, as quatro assistentes possuem, para cada idioma, uma voz padrão, feminina, que será aquela analisada a seguir.

Para além das variações do timbre da voz, existem variações também na forma de sintetização da fala, entonação, concatenação de elementos em sequência em uma frase, etc. Dentre as quatro principais do mercado, a *Alexa*<sup>233</sup> é a que possui voz e forma de falar mais similares àquela utilizada pelos humanos: a voz é suave, expressiva e com entonação adequada, sem exagerar nos dois últimos quesitos; além disso, as transições entre palavras para formular frases mais extensas também não são excessivamente robotizadas e/ou com pausas marcadamente fora de lugar. A voz da *Alexa* é confortável e expressa emoções de forma comedida, sem excessos, mas com inflexões bem colocadas. Se comandada, *Alexa* emite sons de risada (soa como: hihih), de espirro (soa como: aatchw), canta<sup>234</sup> e imita algumas vocalizações de animais. Por vezes, a projeção do som da voz da assistente no ambiente parece ser prejudicada por sua baixa intensidade. Mesmo que seja similar, ela não seria tomada por uma fala humana, uma vez que mantém características minimamente robóticas (propositadamente, talvez?).

Em seguida, temos a voz da *Google Assistente*: possui uma boa naturalidade no timbre da voz (mais agudo que as demais), entretanto a entonação soa um pouco forçada. A voz é excessivamente empolgada para os parâmetros de uma média de vozes humanas<sup>235</sup>, podendo ter efeitos positivos e negativos. Por um lado, contagia por ser simpática e soar bem-disposta. Por outro, o excesso de expressão e dramaticidade na voz gera estranhamento, uma vez que parece mimetizar outras formas de falar em mídias sonoras (como o rádio, por exemplo). O tom da voz e a forma de falar lembram o de uma adolescente/jovem, enquanto *Alexa*, *Siri* e *Cortana*

<sup>233</sup> Em que pese aqui que a *Alexa* é a única que não tivemos acesso em português e, portanto, foi avaliada na versão em inglês que é, pelo tempo e frequência de utilização, sempre mais desenvolvida.

<sup>234</sup> Em uma das interações, *Alexa* cantou um rap de autoria própria, que fala sobre uma “happy AI here in the cloud”.

<sup>235</sup> Ou, talvez, apenas para os parâmetros de gostos pessoais da autora.

tem vozes e formas de falar que se assemelham mais a de mulheres adultas. Em comparação com a *Alexa*, o timbre e entonação de voz da *Google Home* são mais audíveis no ambiente. Ainda, a concatenação entre elementos de uma mesma frase faz algumas pausas e entonações diferentes daquela utilizada por humanos. A *Google Assistente* pode cantar e rir quando solicitada [soa como: hahaha], mas não entende o pedido de um espirro.

A voz da *Siri*, na versão em português, apresenta grandes variações de timbre durante a fala (oscilando entre uma voz mais e menos sintética), soando bastante robótica em alguns contextos e frases específicas; a leitura de frases mais longas também são mais lentas e com tempos inadequados entre uma palavra e outra, faltando expressividade. Os elementos mais robóticos dessa fala acabam atrapalhando a percepção de outros elementos, por se sobressaírem excessivamente. Quando pedimos que a *Siri* cante, ela diz “Dó ré mi fá sol lá si dó! Desculpe, é que eu desisti de minhas aulas de canto quando cheguei na parte dos exercícios de vocalização” (entretanto, ela canta em outras interações, quando perguntada sobre o carnaval, por exemplo). Além disso, ela não reproduz outros sons humanos, como a risada e o espirro (nem mesmo compreende esses comandos) e também não-humanos, como barulhos e vocalizações de animais.

A *Cortana* possui uma voz de timbre médio, com entonação média: fica entre a entonação acentuada da *Google Assistente* e sutil da *Alexa*, sendo mais eficiente do que a fala robótica da *Siri*. A concatenação dos elementos da frase é quase sempre adequada, exceto ao final de algumas frases onde ocorrem poucas pausas. A voz se projeta bem no ambiente, mesmo vinda do computador, e parece ter uma forma mais incisiva de expressão do que a *Google Assistente* e a *Alexa*. A voz e forma de falar da *Cortana* remete a outras utilizações da voz feminina adulta gravada, em locais como aeroportos, estacionamentos, aviões e transporte público. *Cortana* possui um repertório musical vasto, cantando, além de músicas, hinos de times de futebol brasileiros, sempre que solicitada. Entretanto, realiza buscas no *Bing* toda vez que pedimos que ela ria ou espirre.

O investimento em tecnologias relacionadas a voz vem possibilitando aumento das potencialidades de artefatos como as assistentes pessoais digitais. Detectar e projetar emoções está no *roll* de habilidades que as empresas desejam viabilizar através das assistentes pessoais digitais. A *Alexa*, por exemplo, “é cada vez mais capaz de ouvir e responder à prosódia e a emoção, fazendo com que se parece mais e mais consciente aos olhos do usuário”<sup>236</sup> (WEST, KRAUT e CHEW, 2019, p. 109): uma atualização recente da assistente melhorou o

---

<sup>236</sup> No original: “is increasingly capable of hearing and responding to prosody and emotion, making it seem more and more sentient-like to users”

reconhecimento de padrões na entonação da linguagem falada relacionados ao estresse. Além disso, atualmente a assistente da *Amazon* consegue detectar quando o usuário está sussurrando com ela e sussurrar de volta<sup>237</sup>, mantendo o padrão da comunicação estabelecida pelo interlocutor.

O processamento de linguagem natural também é capaz de realizar a “mineração de opinião”<sup>238</sup>, que se constitui como uma análise dos sentimentos expressos na fala do usuário durante a interação. Elementos como expressões verbais, tom da voz e respostas/interações relacionadas à estados psicológicos e emocionais são analisados a partir de um padrão (baseado, por exemplo, em conjuntos de palavras ou padrões de sonoridade) – essa é empreitada, entretanto, bastante complexa (DEVLIN, 2018). A ideia é que detectar o estado emocional do usuário torna possível uma resposta em tom adequado para o usuário (DEVLIN, 2018), conforme sugerem alguns estudos da CASA (NASS, 2006). Não existe clareza, entretanto, se estes processos são ou não empregados atualmente pelas assistentes pessoais digitais existente no mercado.

Cada uma das vozes de cada uma das assistentes pessoais digitais é feita a partir da gravação de palavras e frases de diversas pessoas diferentes, com timbres de voz similares. Posteriormente às gravações, essas vozes são tratadas digitalmente para se tornarem uma única voz uniforme. As gravações dessa nova voz passam a constituir uma voz sintética, que não é exatamente a voz de nenhuma outra pessoa e possuem características mais ou menos robotizadas (G1 ouviu, 2019). A partir disso, as assistentes irão utilizar as palavras e/ou frases para compor suas falas em ocasiões específicas, através de uma tecnologia chamada “text to speech” (TTS). Essa tecnologia possibilita também que, uma vez que a voz foi gerada sinteticamente, seja possível inserir novas palavras não gravadas ao longo da fala. A utilização da tecnologia TTS possibilita driblar limitações técnicas na utilização de vozes para a interface:

Because of limitations of storage space (digital recordings are large), processing speed (finding and combining arbitrary utterances can be slow), bandwidth speed (sound files do not transmit gracefully over a 33 kilobyte phone line), dynamism of content (all of the Web's content cannot be spoken and recorded in real time), and other technical constraints, much of the speech that is and will be produced by computers, the Web, telephone interfaces, and wireless devices will be "synthesized" speech, also known as text to speech (TTS). Thus, more and more interfaces will not simply present fully recorded words and phrases but will produce speech artificially (NASS, 2006, Locais do Kindle 237)

---

<sup>237</sup> Essa habilidade da *Alexa* gera um efeito de estranhamento, como se ela fosse capaz de compreender código sociais implícitos (como seguir a deixa de um dos interagentes para não fazer barulho). A própria sonoridade do sussurro, vindo de uma pequena caixa de som, evoca elementos do *Umheulich*.

<sup>238</sup> O termo original utilizado é em inglês: “opinion mining”.



Ainda que o TTS possibilite uma comunicação razoavelmente clara e com amplitude de repertório, é justamente o processo de sintetização que gera pausas inexplicáveis ou em momentos inadequados, erros de entonação das frases e em ênfases das palavras, gerando uma percepção da voz como não-humana (NASS, 2006). Para gravações que são fixas, ou seja, possuem poucas alterações, a maior similaridade com o humano da voz gravada seria a melhor opção, entretanto se torna inviável para interfaces de voz que utilizam grande base de dados para informar, como é o caso das assistentes. A utilização intercalada de voz sintética e não sintética em um mesmo produto pode gerar, de acordo com Nass, um estranhamento, sendo preferível uma voz de baixa qualidade, mas com consistência.

Walter Ong (2002) define a existência de duas formas de oralidade: a) a oralidade primária, em povos que não desenvolveram ou tiveram contato com formas escritas da linguagem; b) a oralidade secundária, presente em todas as sociedades contemporâneas que desenvolveram a escrita. A oralidade secundária apresenta características distintas da oralidade primária tanto em função do desenvolvimento da escrita (que modifica as formas de armazenamento e transmissão de conhecimento) quanto pela presença de tecnologias fundamentadas na oralidade (como o telefone, o rádio ou as assistentes pessoais digitais). As assistentes pessoais digitais funcionam da seguinte maneira: primeiro identificando uma frase e traduzindo-a para a forma escrita, para posteriormente decodificar os elementos dessa frase e, por fim, gerar uma resposta em termos de ação ou na forma de texto, que será novamente traduzida para a forma oral através da tecnologia TTS. Existe um processo constante de tradução da oralidade para a escrita, da escrita para o comando, do comando para a escrita e da escrita para a oralidade. Desta forma, ainda que evoquem uma forma de comunicação oral com humanos (ou escrita, para a assistentes que operam na tela), as assistentes se relacionam com elementos da oralidade, da textualidade e da linguagem de programação, em *loop* de manipulação de formas e símbolos diferentes para a interação.

O que a tecnologia TTS presente nas assistentes pessoais digitais faz é, portanto, transformar qualquer informação a partir de texto em uma informação verbal, concedida por uma voz que o resultado sintético de vozes humanas. Nesse sentido, não existe a reprodução do sistema biológico da fala humana, que é

Próximo ao impossível de reproduzir na forma mecânica. Portanto a questão se torna como os seres artificiais irão simular a mídia da voz humana. De que forma robôs falam, podemos ter conversações com eles e a voz de quem ouvimos quando eles

falam de volta conosco?<sup>239</sup> (MÄNNISTÖ-FUNK; SIHVONEN, p. 46, 2018)

A relação do homem com a sonoridade é distinta daquela estabelecida com as outras formas de comunicação, principalmente por fatores relacionados à temporalidade. Ainda que “all sensations takes place in time” (p. 31) o som “exists only when it is going out of existence” (p. 31). Nesse sentido, o som e, conseqüentemente a fala, se inscrevem apenas no tempo presente (KITTLER, 1999), sendo “não apenas perecível mas essencialmente evanescente. Quando eu pronuncio a palavra “permanência”, no momento em que eu chego no “-nência”, o “perma-“ já desapareceu, e tem que desaparecer”<sup>240</sup> (ONG, 2002, p. 31 e 32). Essa característica de dissolução do som se apresenta também nas versões gravadas e reproduzíveis de elementos sonoros: ainda que possamos armazenar, transmitir e reproduzir um discurso ou uma música, por exemplo, o momento do som não pode ser congelado no tempo:

Não existe forma de parar o som e de ter o som. Eu posso parar uma imagem em movimento e ter um frame na tela. Se eu parar o som, eu não tenho nada – apenas o silêncio, sem qualquer som. Todas as sensações acontecem no tempo, mas nenhum outro campo sensorial resiste totalmente a uma ação de apreensão, estabilização, desta forma<sup>241</sup> (ONG, 2002, p. 32)

Desde 2018, a voz em português padrão da Google Assistente vem sendo utilizada em músicas brasileiras, especialmente no brega funk de Pernambuco<sup>242</sup>(G1 ouviu, 2019). Em programa de podcast dedicado ao assunto, o *GI ouviu* explorou as utilizações da “voz da mulher do Google” em músicas nacionais, que chegaram inclusive a ficar entre as mais ouvidas no ano de 2019<sup>243</sup>. De acordo com o programa, esse recurso foi emprestado de canais de Youtube que há algum tempo utilizam essa forma para gravar vídeos<sup>244</sup>. A escolha pela voz da Google Assistente, e não das demais assistentes pessoais, se explica através de três fatores: primeiro, o ecossistema de produtos do Google é variado e bastante difundido no Brasil, além de parte deles

---

<sup>239</sup> No original: close to impossible to reproduce in a mechanic form. Therefore the question becomes, how artificial beings can effectively simulate the medium of the human voice. How do robots speak, can we have conversations with them and whose voice do we hear when they talk back to us? (Männistö-Funk; Sihvonen, p. 46, 2018)

<sup>240</sup> No original: “not simply perishable but essentially evanescent, and it is sensed as evanescent. When I pronounce the word ‘permanence’, by the time I get to the ‘-nence’, the ‘-perma-’ is gone, and has to be gone”.

<sup>241</sup> No original: There is no way to stop sound and have sound. I can stop a moving picture camera and hold one frame fixed on the screen. If I stop movement of sound, I have nothing - only silence, no sound at all. All sensations takes place in time, but no other sensory field totally resists a holding action, stabilization, in quite this way (ONG, 2002, p. 32)

<sup>242</sup> De acordo com o podcast *GI ouviu* esse estilo é caracterizado por misturar elementos do brega local com o funk, vindo do sudeste.

<sup>243</sup> Entre as músicas que utilizam este recurso, figuram as produções de DJs nas seguintes canções: “Tome empurradão”, “Ninguém fica parado”, “Hit Contagiate”, “Hoje eu vou dar”, entre outras.

<sup>244</sup> Entre os exemplos citados está o canal de Youtube “Corinthians Mil Grau”, que faz narração de partidas de futebol utilizando a voz masculina dos sistemas do Google.

serem de utilização gratuita<sup>245</sup>, o que o torna conhecido e utilizado por grande parte dos brasileiros online; o Google já vinha utilizando vozes em outros produtos, como o Google Maps e o Google Tradutor; e, por fim, o sistema do Google Tradutor, que permite a entrada de texto escrito e a saída de texto em áudio – o mesmo pode ser feito através de aplicativos de terceiros, que possibilitam a conversão de qualquer texto para as vozes utilizadas pela Google<sup>246</sup>.

A voz e fala são elementos centrais na distinção entre o humano e a máquina, já que agem como uma forma de separar o humano das formas não-humanas, sejam elas animadas ou inanimadas. A capacidade de falar aparece em “O Discurso do Método”, de Descartes (2009), como prova da existência de uma alma imaterial humana e como traço de distinção destes em relação aos animais. A capacidade de falar, conforme mencionamos, é o que demarca a capacidade de pensar: aqueles que são dotados de inteligência, falam (embora talvez a recíproca não seja mais verdadeira). Por esse motivo, o próprio teste de Turing (1950) utiliza como base de avaliação as capacidades conversacionais dos candidatos e não outras formas de expressão de inteligência.

In the midst of human-like machines confusing the boundaries of human and non-human in an uncanny manner, speech has stayed put as the ultimate barricade that could not result from technical skills alone, but would necessitate the operations of either a living soul, divine powers, magic, or extreme technological prowess (MÄNNISTÖ-FUNK; SIHVONEN, 2018, p.58)

Assim, a voz deve ser levada “em consideração quando pensamos nas reações emocionais das pessoas a máquinas que possuem semelhança com o humano”<sup>247</sup> (Männistö-Funk; Sihvonen, 2018, p. 47). Para os autores, a voz se constitui como uma importante ferramenta de agência para humanos e não-humanos, mas sua relação com o sentimento de estranhamento não foi suficientemente investigada.

Uma voz implica também um corpo que a produz: quando ouvimos algo falar, imaginamos a origem daquela voz, associamos características, temos sensações, evocamos memórias sonoras e não sonoras. Nesse sentido, uma máquina que fala aponta também para a possibilidade de um corpo (biológico ou mecânico) por trás daquela voz, implicando uma experiência corpórea de mundo, uma vez que a voz é não apenas “um instrumento da fala mas também uma representação e um produto da identidade e da corporalidade daquele que

---

<sup>245</sup> Considera-se como gratuita a utilização sem necessidade de remuneração financeira. Sabemos, entretanto, que a moeda com a qual as grandes empresas de tecnologia atuam são os dados dos usuários.

<sup>246</sup> Ver, na Play Store dos sistemas Android, o aplicativo chamado “Voz da mulher do tradutor”.

<sup>247</sup> No original: “in consideration when thinking about people’s emotional reactions to human-like machines”

fala”<sup>248</sup>(Männistö-Funk;Sihvonen, 2018, p. 47). De fato, através da audição é possível conhecer e revelar o interior de um objeto, uma vez que

Todos os sons registram o interior daquilo que os produz. Um violino cheio de concreto não vai soar como um violino normal. Um saxofone soa diferente de uma flauta: possuem estruturas diferentes por dentro. E, acima de tudo, uma voz humana vem de dentro de um organismo que possibilita a ressonância da voz<sup>249</sup> (ONG, 2002, p. 70).

As assistentes pessoais possuem dois conjuntos de materialidades: aquele que processam e produzem o seu conteúdo (e se alocam em estruturas materiais longe dos olhos do usuário, a nuvem) e aquela que recebe o input e devolve o output (que é o hardware de acesso utilizado pelo usuário, como o *smartspeaker* e o celular). Ambos os “interiores” das assistentes são revelados no processo de fala e escuta que se estabelece na interação com a interface de voz, através dos elementos das vozes de cada uma delas, incluindo aí a mecanicidade e as falhas das pausas que elas produzem. São esses elementos que indicam um corpo-máquina ao invés de um corpo-organismo e nos deixa entrever através do som seu status de não-humano.

A experiência humana de ouvir coloca o ouvinte no centro de seu “centro de seu mundo auditivo, que me envelopa, estabelecendo um tipo de centro de sensação e existência”<sup>250</sup> (ONG, 2002, p. 70). Esse efeito é explorado pelos sons de alta fidelidade, por exemplo, na tentativa de possibilitar que o ouvinte possa “immerse yourself in hearing, in sound” (ONG, 2002, p.70). Para o autor, essa forma de imersão que a audição possibilita seria impossível de reproduzir com os demais sentidos humanos, ainda que contemporaneamente tecnologias como os óculos de realidade virtual flertem justamente com essa ideia.

### 6.3 PODEM AS MÁQUINAS FALAR?

A busca por objetos capazes de falar aparece desde a antiguidade grega, com estátuas que possuem tubulações para a fala, chegando até exemplo mais próximos, como o Mechanical

---

<sup>248</sup> No original: “an instrument of speech but also a representation and a product of the speaker’s identity and corporality”

<sup>249</sup> No original: “sounds all register the interior structures of whatever it is that produces them. A violin filled with concrete will not sound like a normal violin. A saxophone sound differently from a flute: it is structured differently inside. And above all, the human voice comes from inside the human organism which provides the voice’s resonances” (ONG, 2002, p. 70).

<sup>250</sup> No original: “auditory world, which envelopes me, establishing me at a kind of core of sensation and existence”.

Turk<sup>251</sup> (JARRET, 2015). Entretanto, as discussões científicas em torno da fala artificial se focam na particularidade do aparelho fonador humano (Männistö-Funk; Sihvonen, 2018). As condições materiais para a emergência da fala nos seres humanos se deram através de uma conjunção de fatores inicialmente corporais: o aumento da caixa craniana e do córtex frontal, a postura ereta, a emergência do polegar opositor, levam a uma conseqüente transformação biológica que realoca a laringe e a língua mais para o fundo da garganta, criando uma caixa de ressonância aperfeiçoada e gerando mais espaço para o movimento da língua dentro da boca (CIANI; CHIARELLI, 1992).

A relação entre a emergência do aparelho fonador e o surgimento da linguagem não está claro entre os cientistas das diversas áreas que trabalham com o tema, devido a dificuldade de obter evidências de uma forma de comunicação que não apresenta registros materiais: “nós humanos utilizamos a linguagem como forma predominante de comunicação. Nós não sabemos como a linguagem evoluiu. Não existem traços nos registros arqueológicos para além de sugestões, como a presença da laringe”<sup>252</sup> (DEVLIN, 2018, p. 116). Supõem-se, entrando, que o surgimento da fala entre os humanos tenha sido impulsionado pela necessidade de estabelecer comunicação quando caçando em grupos – ou seja, se ordenou em torno do trabalho de encontrar alimentos, como se observa em outros animais (como abelhas, por exemplo).

O aparelho fonador humano consegue, atualmente, emitir cerca de cinquenta sons básicos, que se combinam em inúmeras variações através dos padrões linguísticos. “A fala, portanto, é uma série de ondas de som que o nosso cérebro pode claramente interpretar, filtrando o barulho de fundo para compreender não apenas as palavras, mas o contexto. É incrivelmente complexo”<sup>253</sup> (DEVLIN, 2018, p. 117). A emergência da fala como uma prática social torna possível o compartilhamento de uma série de práticas e técnicas de sobrevivência humana – esse conhecimento acumulado foi adquirindo, com o passar do tempo, características cada vez mais complexas através da linguagem.

---

<sup>251</sup> O *Mechanical Turk* foi um autômato construído no final do século XVIII para operar como um jogador de xadrez. Entretanto, não passava de um truque: “despite the elaborate construction of cogs and gears revealed behind the doors of the cabinet, hidden behind that was a living chess master who used magnetic systems to move pieces around the chessboard from under the table” (JARRET, 2015, p. 1). Hoje, o nome aparece na plataforma *Amazon Mechanical Turk*, um projeto de terceirização de tarefas para sistemas de inteligência artificial que precisam ser performadas por humanos (como tagueamento de fotos, por exemplo).

<sup>252</sup> No original: “we humans use language as our predominant method of communication. We don’t know how language evolved. There aren’t any traces in the archaeological record other than suggestions, such as the presence of a larynx”

<sup>253</sup> No original: “Speech, therefore, is a series of sound waves that our brain can very cleverly interpret, filtering out background noise to comprehend not just words but context, tone and intention. It’s stunningly complex”

Ainda que a ideia de reproduzir a fala de forma mecânica tenha sido uma busca que durou séculos, nem sempre os resultados, quando bem-sucedidos, obtiveram viabilização econômica ou mesmo interesse de um público consumidor<sup>254</sup>. A indústria de brinquedos, durante o século XIX e XX se constitui como um local relevante para o desenvolvimento e experimentação de tecnologias com a fala, justamente por ser um dos poucos setores em que esses artefatos se apresentavam enquanto lucrativos (MÄNNISTÖ-FUNK; SIHVONEN, 2018).

A partir da visada histórica dos desenvolvimentos ligados a capacidade de fala de máquinas e objetos, os autores concluem que, para tornar a aproximação das máquinas aos humanos mais palatável, utiliza-se com frequência uma forma de construção do Outro que evocam características como raça, etnia e gênero. Essas características seriam capazes de desviar a tenção do efeito de estranhamento e tornar estes artefatos lucráveis: “o papel e o significado dos seres humanos artificiais têm sido consistentemente vinculados à prática cultural de se relacionar, associando-os de forma mais ou menos óbvia a "outros" radicalizados, étnicos e de gênero” (MÄNNISTÖ-FUNK; SIHVONEN, 2018, p. 59, tradução nisso<sup>255</sup>). Nesse sentido, atribuir gênero (feminino) e raça (não-caucasianas) de grupos historicamente oprimidos, seria uma tentativa de marcar uma diferença, de criar uma barreira entre *nós* e *eles*, *humanos* e *não-humanos* (ou, *menos-humanos*), nesses sistemas. Essa barreira seria uma fronteira entre aquilo que é humano e o que não, tornando uma similaridade entre máquinas e sujeitos que são socialmente jogados para as margens dessa fronteira humana mais aceitável. “Parece ser mais fácil deixar máquinas se aproximar das bordas da humanidade se nós pegarmos formas “menos” humanas, ou seja, outra coisa que não adultos brancos do sexo masculino”<sup>256</sup> (Männistö-Funk; Sihvonen, 2018, p. 59). Nesse sentido, o que está em jogo são os processos de desumanização na constituição de quem pode e quem não pode ser Outro, frequentemente associadas também a processos mediados (BUTTLER, 2011). Esse processo de desumanização, que traz também o conteúdo do exótico (principalmente etnicamente), serve para construir e combinar fantasias que pensam os robôs e as inteligências artificiais tanto como maleáveis e obedientes, quanto fascinantes e intrigantes.

<sup>254</sup> Ver os casos de Joseph Faber e da máquina chamada Euphonia e as bonecas falantes de Tomas Edison, ambas no século XIX (RISKIN, 2016); ou as máquinas falantes de leitura da sorte, em parques de diversão no começo do século XX (Männistö-Funk; Sihvonen, 2018).

<sup>255</sup> No original: “the role and significance of artificial humans has consistently been linked to the cultural practice of othering through associating them more or less obviously with radicalized, ethnicized, and gendered ‘Others’”

<sup>256</sup> No original: “It appears to be easier to let machines come close to the borders of humanity if they take form of ‘lesser’ humans, meaning anything other than white adult males”

O ato da fala humana é sempre dirigido a um Outro, em um jogo comunicativo no qual antecipamos formas de pensar, agir e responder, mesmo antes de uma primeira interação (ONG, 2002). Mesmo quando falamos com objetos/seres vivos com os quais não há resposta verbal possível, modulamos a voz, as formas de falar e as expressões de sentimentos (de raiva, quando meu computador não funciona, de felicidade, quando meu cachorro me espera chegar em casa, por exemplo). Ou seja, falamos da forma que falamos ao pensarmos no *que/quem* é o outro com que eu falo, como eu devo me dirigir a ele, que formas de falar eu devo utilizar para atingir meu objetivo, etc. Falar não é, em estados não-patológicos da mente, apenas falar livremente:

Para falar, é preciso endereçar esta fala a um outro ou outros. Pessoas em sã consciência não saem pela floresta falando aleatoriamente com ninguém. Mesmo para falar consigo mesmo é preciso fingir que existem duas pessoas. A razão para isso é que o eu digo depende da realidade que eu imagino estar entrando, ou seja, quais respostas possíveis eu consigo antecipar. Portanto eu evito enviar a mesma mensagem para um adulto e para uma criança. Para falar eu preciso já estar em comunicação com a mente que eu irei endereçar antes de começar a falar<sup>257</sup>(ONG, 2002, p. 172)

Falar é, portanto, sempre falar com um Outro, imaginar um Outro – e esse Outro que irá nos responder de volta não é mais necessariamente humano. Ainda assim, é um Outro ao qual eu atribuo determinadas características, crio expectativas de retorno, prevejo respostas possíveis e, com isso, inevitavelmente modulo minha própria forma de falar.

Criar uma máquina capaz de levar um diálogo razoavelmente coerente com um ser humano não é uma tarefa fácil. Mesmo no caso das assistentes, em que processos de *machine learning* são utilizados a partir das conversas com os usuários (e das preferências de um usuário específico também), os desenvolvedores precisam imaginar respostas na forma de um roteiro para perguntas possíveis, que posteriormente serão implementadas por programadores. Produzir diálogos para sistemas de propósito geral é, entretanto, extremamente difícil – e não fica necessariamente muito mais fácil em sistemas de propósito específico. No caso ELIZA, por exemplo, que atua exclusivamente na forma de interação entre terapeuta e paciente, não precisamos ir muito além de quatro ou cinco frases para perceber o limite da composição de respostas para afirmações e perguntas. Os limites, nesse sentido, são tanto maquínicos quanto humanos: “there will always be limitations to what a computer can say in response to a human,

---

<sup>257</sup> No original: To speak, you have to address another or others. People in their right minds do not stray through the woods just talking at random to nobody. Even to talk to yourself you have to pretend that you are two people. The reason is that what I say depends on what reality of fancy I feel I am talking into, that is, on what possible responses I might anticipate. Hence I avoid sending quite the same message to an adult and to a small child. To speak, I have to be somehow already in communication with the mind I am to address before I start speaking (ONG, 2002, p. 172)

because there will always be limits to the imagination of the human computer programmer.” (BROUSSARD, 2018, p. 71). O caso ELIZA nos fornece indícios para compreender o funcionamento das assistentes pessoais digitais:

Olhando para as respostas de Eliza é fácil perceber como as assistentes de voz como a Siri da Apple são programadas. A ELIZA original tinha algumas dezenas de respostas; Siri possui muitas, muitas respostas feitas por muitas muitas pessoas. Siri pode fazer muita coisa: mandar mensagens, fazer ligações de telefone, atualizar o calendário com compromissos ou colocar um despertador. Pode ser divertido desafiar a Siri. Crianças pequenas se deliciam especialmente em testar os limites do que a Siri vai falar. Entretanto, Siri e outras assistentes de voz são limitadas em suas respostas verbais pela imaginação coletiva (e pela assimetria positiva) de seus programadores<sup>258</sup> (BROUSSARD, 2018, p. 73 e 74)

Entre a ELIZA de Weizenbaum e as assistentes de hoje, em termos de tecnologias, as diferenças residem principalmente nos avanços ligados ao reconhecimento de fala e ao processamento de linguagem natural, com uma mudança na forma do input/output de texto para input/output de voz (DEVLIN, 2018). Que as assistentes pessoais consigam manter pequenos diálogos com razoável eficiência e, principalmente, responder com ações aos comandos de voz recebidos, não implica que elas possuam inteligência no sentido humano do termo ou mesmo que elas consigam compreender os símbolos que manipulam através da fala, seguindo o argumento do quarto chinês de Searle.

Compreendemos que os seres humanos apresentam habilidades para lidar com vozes e com a fala distinta daquelas que a máquina apresenta mesmo com tecnologias avançadas e que as formas de funcionamento das assistentes não são análogas as formas de pensamento humanos. Ainda assim, o fato de que as assistentes falam e mantêm pequenas conversações geram efeitos materiais no cotidiano e na vida humanas. Nesse sentido, a importância das assistentes reside justamente no mimetismo em relação ao ser humano, “precisamente porque elas soam como pessoas. Pesquisas acadêmicas recentes sugerem que as pessoas são melhores reconhecendo emoções humanas quando elas apenas ouvem a voz de quem fala”<sup>259</sup> (WEST, KRAUT e CHEW, 2019p.109)

---

<sup>258</sup> No original: Looking at Eliza’s responses, it’s easy to see how voice assistants like Apple’s Siri are programmed. The original Eliza had a few dozen responses; Siri includes many, many responses crafted by many, many people. Siri can do a lot: it can send messages, place phone calls, update a calendar with appointments, or set an alarm. It can be fun to stump Siri. Little kids take special delight in testing the outer limits of what Siri will say. However, Siri and the other voice assistants are limited in their verbal responses by the collective imagination (and positive asymmetry) of their programmers (BROUSSARD, 2018, p. 73 e 74).

<sup>259</sup> No original: “precisely because they sound like people. Recent academic work has suggested that people are better at recognizing human emotion when they can only hear a speaker’s voice.”



Quando pessoas diferentes vocalizam uma mesma palavra, ouvimos uma pluralidade de formas de pronúncia distintas, que variam de acordo com fatores como “barulho de fundo, estado emocional, idade, velocidade, volume e pronúnciação”<sup>260</sup> (DEVLIN, 2018, p. 118). Esses sinais podem ser analisados e categorizados por uma máquina, através de reconhecimento de padrão e busca por contexto em elementos-chave, mas não é um trabalho fácil. Devido ao longo treinamento que humanos recebem desde o berço, em situações e contexto variados, com atores que também variam, conseguimos inferir significado de diversos elementos presentes na voz – e inclusive do silêncio. O que elas conseguem fazer em termos de desmantelamento de elementos em uma frase para gerar uma ação desejada na verdade, é

Simplesmente lançar uma sequência de respostas sônicas que humanos chamam de comandos verbais. “Alexa, toque “California Girls” é um comando de voz que um computador pode seguir. Alexa é uma palavra que diz ao computador que um comando está vindo. Tocar é uma palavra que significa “retirar um MP3 da memória e enviar um comando de tocar para um tocador de áudio previamente especificado, junto com o nome do arquivo de MP3. A interface também é programada para capturar qualquer palavra que venha depois de tocar e antes da pausa (o fim do comando) [...] Esse processo é procedural e inofensivo e não deveria fazer ninguém pensar que máquinas vão se revoltar e dominar o planeta”<sup>261</sup> (BROUSSARD, 2018, p. 98 e 99)

Faz sentido, portanto, que exista uma fórmula específica de conversa com uma assistente pessoal digital com sucesso, que implica na construção de uma frase adequada, com todas as informações necessárias, na utilização dos verbos corretos para os comandos desejados, de clareza de fala em nomes próprios ou títulos, etc. As frases são comandos que, caso sejam ditas de forma incompleta, na sequência ou com as palavras erradas, não serão compreendidas justamente por sua limitação de compreensão contextual, de inferência e de aproximação do conteúdo. Por exemplo, quando eu estou em casa com o meu namorado, eu posso apenas pedir a ele que “coloque o Foro de Teresina para a gente ouvir”. Eu não preciso dizer para ele que “Foro de Teresina” é um *podcast* produzido pela *Revista Piauí* e que ele deve colocar através da plataforma *Spotify*, pois essas são informações que nós compartilhamos. Essa repetição entre seres humanos pode ocorrer tanto pela repetição de uma mesma atividade (memória) quanto pelo compartilhamento de elementos em comum da sociedade e da cultura.

<sup>260</sup> No original: “background noise, emotional state, age, speed, volume and pronunciation”.

<sup>261</sup> No original: simply launch computerized sequences in response to sonic sequences, which humans call verbal commands. “Alexa, play “California Girls” is a voice command that a computer can follow. Alexa is the trigger word that tells the computer that a command is coming. Play is a trigger word that means “retrieve an MP# from memory and send the command play to a previously specified audio player, along with the MP# file name. The interface is also programmed to capture whatever word comes after play and before the pause (the end of the command). [...] This process is procedural and unthreatening and shouldn’t make anyone think that the machines are going to rise up and take over the world (BROUSSARD, 2018, p. 98 e 99).

Entretanto, *Siri*, *Alexa*, *Google Assistente* e *Cortana* não compartilham o mundo social e cultural nem atividades conosco, o que torna imprescindível, quando eu quiser ouvir o podcast, dizer: “*Toque o podcast Foro de Teresina de Revista Piauí no Spotify*”<sup>262</sup>.

Justamente pela necessidade de falar de forma específica e não como naturalmente agimos em um diálogo com outros humanos é que Amber Case (2015) compreende que as interfaces de voz não são tecnologias calmas. Como precisamos utilizar técnicas de fala na interação com as assistentes, existe sempre a possibilidade de não sermos compreendidos:

Esse tipo de experiência pode ser unicamente frustrante para os usuários porque força que eles modifiquem o seu comportamento em benefício da máquina, mas a máquina esta apenas demandando esses contorcionismos porque está tentando “se comunicar como um humano”. Fazer um computador falar como um humano sem colocar nele algum senso de contexto ou de relacionamento acaba levando a uma sensação de dissonância na pessoa que utiliza – exatamente o que o design afetivo tenta remediar<sup>263</sup> (CASE, 2015, p. 125).

A leitura da autora, ainda em 2015, é que a *Siri*, lançada há poucos anos no mercado, seria um fracasso por alguns grupos porque suas capacidades reais não correspondem ao que ela pode oferecer. Mesmo que de lá para cá cerca de cinco anos tenham se passado, outras assistentes tenham surgido e os sistemas tenham melhorado imensamente (principalmente com a recorrência de utilização dos usuários, devido à predominância do *machine learning* nos desenvolvimentos contemporâneos de inteligência artificial) essa frustração com as potencialidades é comum e foi observada durante as entrevistas e o formulário quantitativo<sup>264</sup>. Em parte, a frustração ocorre por que nem sempre é possível se fazer compreender pela máquina; soma-se a isso o fato de que, mesmo quando nos fazemos compreender, nem sempre as assistentes conseguem operacionalizar o que solicitamos – seus potenciais e limites não são claros nos anúncios veiculados pelas empresas.

Case (2015) compreende que a utilização da fala faz com essa atividade se torne o centro de nossa atenção (como, com frequência, fazemos quando conversamos com outros humanos), já que uma interface de voz exige uma enunciação perfeita (CASE, 2015). Conforme já

<sup>262</sup> Essa frase corresponde a fórmula: ação + nome e especificação + plataforma onde encontrar o arquivo desejado. Dado que o Spotify está configurado como minha plataforma de áudio de preferência, é possível dispensar o último elemento da fórmula, entretanto, nem sempre isso corre com sucesso.

<sup>263</sup> No original: this kind of experience can be uniquely frustrating for users, because it forces them to modify their own behavior for the benefit of a machine, but the machine is only demanding such contortions because it’s trying to “communicate like a human”. Making a computer speak like a human without instilling it with a sense of human context or relationships ultimately leads to a sense of dissonance in the person using it - exactly what affective design is seeking to remedy (CASE, 2015, p. 125).

<sup>264</sup> Na seção destinada a comentário livre, a frustração com as capacidades das assistentes foi mencionada com frequência. Além disso, frustração foi o segundo sentimento mais apontado entre os participantes da pesquisa quantitativa.

pontuamos, a percepção de foco de atenção na interação através da fala depende da atividade e do contexto no qual se utiliza, tornando uma generalização inviável. Entretanto, não é raro entre os usuários das assistentes pessoais que se interrompa uma atividade para “se fazer entender” pelas interfaces de voz. Devlin (2018) considera que a utilização da interface de voz no lugar da interface gráfica possibilita a reutilização de representações já conhecidas dos usuários (através da linguagem, que aprendemos desde criança) ao invés de prescindir da criação de representações gráficas através do design, o que “diminui a carga cognitiva do usuário. Torna muito mais eficiente – se funcionar corretamente”<sup>265</sup> (DEVLIN, 2018, p. 118).

As falhas no sistema de reconhecimento de linguagem natural (ou até mesmo seu próprio modo de funcionamento) nos levam a adaptar a nossa própria fala à máquina:

Essa é uma falha comum dos sistemas de reconhecimento de fala e leva a um fenômeno interessante: nós precisamos adaptar a nossa fala para a máquina. Não são apenas sotaques. Até a sintaxe da nossa fala está mudando. Se nós quisermos interagir com uma assistente digital nós precisamos começar a frase com o seu nome. “E aí, Siri, toque a trilha sonora de Hamilton”. Nós não precisamos dizer por favor ou obrigada também, ainda que muitas pesquisas mostrem que as pessoas o fazem<sup>266</sup> (DEVLIN, 2018, p. 124)

A recorrência da utilização de formas sociais relacionadas a educação e boas maneiras (como obrigada e por favor), são, na verdade, a adição de um fator de complexidade às fórmulas de comandos que as assistentes compreendem, já que “assistentes pessoais digitais são essencialmente só ferramentas de busca controladas por voz”<sup>267</sup> (DEVLIN, 2018, p. 125). Ao longo deste capítulo, analisaremos como os respondentes do formulário online e dos entrevistados lidam com essas questões em sua interação com as assistentes.

Em parte, humanos se adaptam as formas comunicacionais das tecnologias pela própria maneira pela qual a formalização linguística é possível em um ambiente computacional. As regras de comunicação são criadas apriori, ou seja, primeiro estabelecemos as fórmulas de falas a seres utilizadas e compreendidas pelas assistentes e depois as utilizamos, ao contrário do que ocorre na linguagem humana. Nesse sentido, Ong (2002, p. 07) explica que “regras de linguagem de computadores (gramática) são ditadas antes e utilizadas depois. As regras de

<sup>265</sup> No original: “lightens the cognitive load on the user. That makes it far more efficient – if it works correctly”

<sup>266</sup> No original: This is a common failing of speech recognition systems, and it leads to an interesting phenomenon: we have to adapt our speech to the machine. It’s not just accents, either. Even our speech syntax is changing. If we want to interact with a digital assistant we have to begin our sentence with its name. ‘Hey Siri, play the Hamilton soundtrack.’ We don’t have to say please or thank you either, although surveys show that many people do.” (DEVLIN, 2018, p. 124)

<sup>267</sup> No original: “Virtual assistants are essentially just voice-controlled search engines”

gramática da linguagem natural humana são utilizadas e depois se pode extrair delas explicitamente e apenas parcialmente, mas sempre com dificuldade”<sup>268</sup>.

#### 6.4 FALANDO COM “UM PEDAÇO DE SILÍCIO”

Os respondentes do formulário quantitativo foram perguntados sobre como se sentiam ao falar com um objeto e se percebiam que falavam de uma forma diferente do habitual. Para a primeira pergunta, as opções de resposta eram: “Ok, acho normal”, “Não me importo”, “Nunca pensei sobre isso”, “Me sinto um pouco estranho”, “Acho irritante”, “Acho divertido” e “Outros”. Para a segunda pergunta, as opções fornecidas foram: “Não”, “Nunca pensei sobre isso”, “Sim, falo mais devagar”, “Sim, faço frases mais diretas”, “Sim, sou menos educado/cuidadoso ao falar”, “Sim, sou mais educado/cuidadoso ao falar”, “Sim, sou mais consciente de como falo e me expesso”, “Sim, evito demonstrar qualquer tipo de sotaque” e “Outros”.

O gráfico abaixo analisa as respostas em relação a pergunta sobre como os usuários se sentiam ao falar com um objeto.

---

<sup>268</sup> No original: “Computer language rules (‘grammar’) are stated first and thereafter used. The ‘rules’ of grammar in natural human languages are used first and can be abstracted from usage and stated explicitly in words only with difficulty and never completely”.

Figura 54 Gráfico Sensação ao falar com um objeto



Fonte: A autora

Dentre os respondentes, 32% informaram “Ok, acho normal”; 22% informaram “Acho divertido”; 20% disseram que “Me sinto um pouco estranho”; 15% disseram “Não me importo”; 7% responderam “Nunca pensei sobre isso”; 2% informaram que “Acho irritante e outros 2% assinaram a categoria outros. A predominância de uma sensação de normalidade, seguida de sensação relacionada ao humor e divertimento e ao estranhamento se relacionam com fatores analisados no capítulo anterior. Falar com um objeto, no contexto deste formulário, é falar com um objeto específico, construído de forma antropomórfica, com personalidades fictícias elaboradas para estabelecerem interações bem-humoradas. A relação de estranhamento se explica em função de fatores abordados no capítulo anterior e neste, como a relação de proximidade com o humano pelo uso da voz e da fala, que utilizamos, com feedback positivo, apenas com outros seres vivos, em geral seres humanos (consideramos que animais também podem se comunicar com seres humanos, porém não o fazem através da linguagem natural falada).

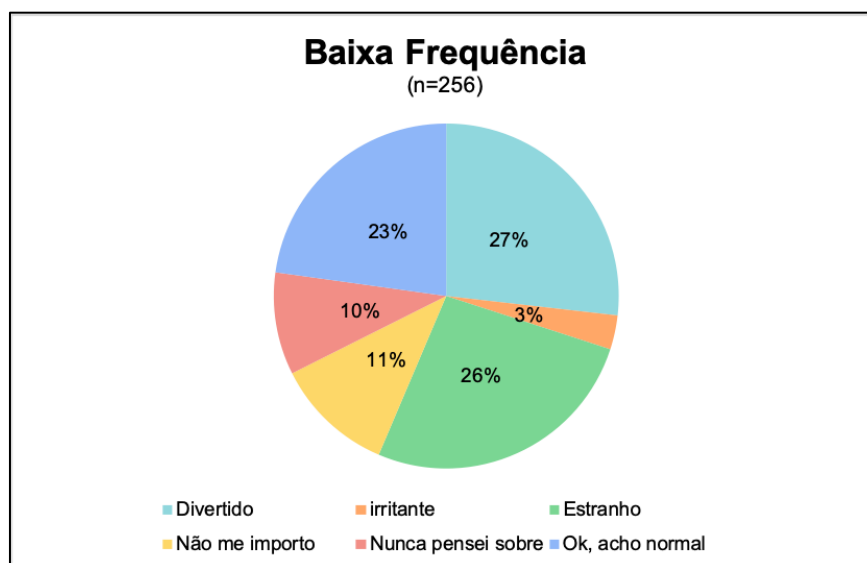
Na categoria outros, observamos a ocorrência principalmente de sensações negativas ao falar. Dois respondentes informaram achar “assustador”, um respondente disse que “Me sinto um objeto” e outro respondente disse que ficava se perguntando sobre a própria natureza do ser humano: “Será que humanos também são objetos? Desperta minha curiosidade”. Essas colocações trazem à tona o caráter de questionamento de nossa própria humanidade (ao invés

de questionar a humanidade da máquina!), narrado também pelos participantes humanos no teste de Turing do Loebner Prize.

Ainda na categoria outros, dois respondentes apontaram para a dificuldade de comunicação: um deles disse que se sente “lutando contra um algoritmo” enquanto o outro disse que sente que é frustrante às vezes, “porque tenho que mudar meu ritmo de fala (falando de forma mais lenta e pausada), além de ter que explicar muito detalhadamente o que quero. Por exemplo, para tocar determinada música tenho que falar à Alexa o nome da música, do artista, e às vezes especificações”. Ambas as respostas apontam para o próximo tópico, que diz respeito as adaptações que precisamos fazer na fala para conseguir estabelecer uma interação efetiva com as assistentes, o que pressupõe uma prática e um conhecimento específicos, desenvolvidos na utilização. Dois respondentes informaram também que a forma como se sentem ao falar pode variar conforme o contexto, por exemplo, um dos usuários disse que “Me sinto estranha na presença de outras pessoas apenas”, enquanto a outra informou que oscila entre frustração (quando a interação não ocorre adequadamente) e animação (“quando, por exemplo, coloca exatamente uma música que eu pedi e queria ouvir”). Um respondente informou ainda que não gosta de falar com um objeto “porque prefiro ser discreta”, apontando novamente para a ideia de que a presença de outras pessoas é um fator na hora de compreender como alguém se sente nessa interação.

Realizamos o cruzamento das respostas referente a como os usuários se sentem ao falar com um objeto e a frequência de utilização apontada pelos mesmos. Os gráficos abaixo apresentam os resultados dos respondentes que possuem baixa frequência de uso (utilizaram uma vez ou esporadicamente, totalizando 256 respondentes), média frequência (utilizam menos de uma vez por semana e de uma a duas vezes na semana, totalizando 75 respondentes) e alta frequência (utilizam entre três e seis vezes na semana ou diariamente).

Figura 55 Gráfico Baixa frequência X sensação ao falar

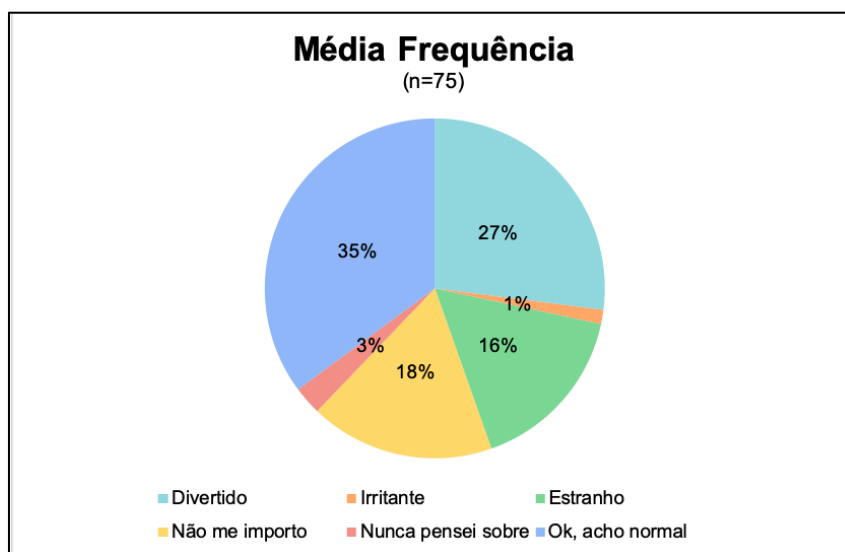


Fonte: A autora

Ao isolarmos as respostas dos usuários de baixa frequência de uso, a ordem dos resultados em relação aos resultados globais se apresenta de forma diferente: as respostas “Acho divertido” (26%) e “Me sinto um pouco estranho” (26%), ultrapassam a resposta “Ok, acho normal” (23%). Essa alteração pode apontar para o fato de que a normalização da utilização da fala nesse tipo de interação é um processo que acontece através da repetição – quanto mais se utiliza, mais normal esse recurso parece. O restante dos resultados mantém a mesma ordem dos resultados globais, com “Não me importo” (11%), “Nunca pensei sobre isso” (10%) e “Acho irritante” (3%). A resposta “Nunca pensei sobre isso” apresenta a maior proporção de respostas em usuários de baixa frequência, apontando que a utilização mais frequente da tecnologia leva os usuários a refletirem sobre suas próprias percepções e relações com a mesma.

O retorno “Acho irritante” aparece em último lugar, como nos resultados globais, porém nos usuários de baixa frequência essa categoria contém oito dos respondentes que apontaram essa sensação, de um total de dez respondentes do conjunto global. É difícil estabelecer, entretanto, uma relação de causa e consequência: por uma lado, achar irritante (por não se fazer compreender, na maioria das vezes) pode ser uma barreira de utilização que levam os usuários a desistirem de utilizar as assistentes; por outro lado, é possível que a irritação gerada por não se fazer compreender (supondo que seja essa a causa), pode evoluir em usuários mais experientes, uma que vez que estes aprendem as formas mais adequadas de falar com a assistente para que a solicitação seja compreendida.

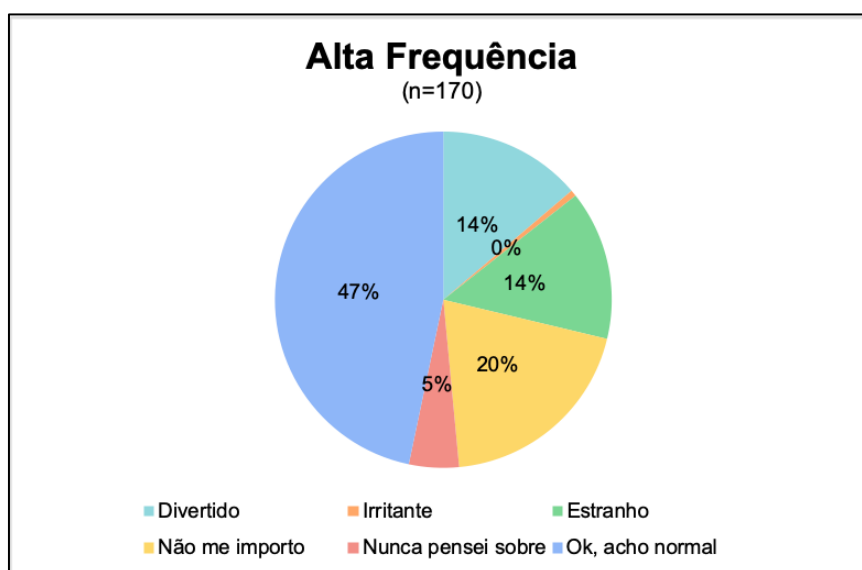
Figura 56 Gráfico Média frequência X sensação ao falar



Fonte: A autora

Os usuários de frequência média mantiveram a ocorrência das repostas globais nas primeiras categorias, colocando a categoria “Ok, acho normal” em primeiro lugar (35%), seguida de “Acho divertido” (27%). Na sequência, “Não me importo” (18%) aparece antes de “Me sinto um pouco estranho” (16%), seguidos por “Nunca pensei sobre isso” (3%) e “Acho irritante” (1%). Nesses resultados, os retornos referentes ao estranhamento ao falar com o objeto e sobre não ter pensado sobre a questão apresentam resultados mais baixos que os obtidos na amostra global.

Figura 57 Gráfico Alta frequência X sensação ao falar



Fonte: A autora



Em respondentes com alta frequência de utilização, a normalização da utilização da fala na interação com as assistentes fica evidente pelo resultados elevados da resposta “Ok, acho normal” (47%), seguidos pela resposta “Não me importo” (20%) – está última, nos resultados globais, ocupavam a quarta posição de maiores retornos, passando para segunda posição nesse conjunto de respostas. Em seguida temos “Acho divertido” e “Me sinto um pouco estranho”, cada uma com 14% das respostas. Este é o único conjunto de respondentes no qual humor e estranhamento apresentam resultados similares, já que nos demais o fator diversão ficava sempre a frente de estranhamento. Ambos são menores do que os resultados nos grupos de baixa e média frequência, assim como nos resultados globais, apontando que tanto o fator diversão quanto o fator estranhamento, analisados em capítulos anteriores, tendem a cair com o tempo de utilização. Por último aparecem “Nunca pensei sobre isso” (5%) e “Acho irritante” (com apenas um respondente).

Durante as entrevistas em profundidade, os participantes também foram indagados sobre como se sentiam ao utilizar a linguagem natural falada com um objeto. As respostas fornecidas nos permitem analisar as especificidades dos dados apontados pelo formulário online. Aparecem os sentimentos de estranhamento, de diversão, de normalidade (no sentido de ser apenas uma forma de interação com uma interface), e de inadequação, que serão apresentados em detalhes na sequência.

A entrevistada E1 nos conta que, inicialmente, o sentimento predominante era o de estranhamento, mas que, com o tempo de utilização, isso se dissolveu: “no começo eu achava meio estranho, porque era alguém que não existia, eu estava falando com alguém que não existia. depois eu passei a enxergar de outra forma: meu celular existe e eu estou falando com o meu celular. eu tenho mania de dar nome para os objetos, então eu me adaptei”.

Para o entrevistado E2, o principalmente sentimento ao falar com a assistente era de vergonha: “logo no começo eu morria de vergonha. Eu chegava bem perto e falava o nome dela e fazia o pedido. Depois eu descobri que tinha um botão que não precisava falar o nome dela, hoje eu sou super tranquilo, eu já chego e digo: “Alexa cheguei em casa”, aí ela começa a rotina [liga a tv, fala sobre o tempo, e etc]. Eu criei várias rotinas. Hoje eu falo com ela com normalidade.

E3 nos conta que a utilização da fala para a interação com as assistentes se assemelha, atualmente, para ele, com outras formas de interação com mídias digitais, como a utilização de botões ou controles, por exemplo. “Como eu me sinto hoje, eu acho que é como tocar num objeto. Eu toco no mouse, eu toco no teclado. É uma forma de interagir, a gente usa um dos

sentidos. É só mais um sentido se tu parar para pensar. Mas no começo eu estranhei, porque não é tão comum. Eu já tinha falado com o microfone, se é que se pode ser considerado falar com. Mas o microfone não responde, certamente. É, a Alexa é um microfone que responde, exato. É um objeto mais complexo né, assim como um celular. Tem botões que não funcionam como botões de um teclado, eles não voltam [no sentido de se mover e responder mecanicamente ao toque para acionar um determinado local]. Provavelmente a primeira vez que alguém usou foi surpreendente. Positivamente surpreendente. Vai ter uma mini curva de aprendizado, de se acostumar. Mas ela passa e hoje eu não sinto nada”. Em sua fala, E3 pontua três fatores interessantes: a relação entre fala e manuseio de um objeto, com finalidade para a ação; o estranhamento inicial por se tratar de uma forma diferente de interação daquela que estamos habituados (e não necessariamente por ser um objeto falante); e o período de aprendizado de utilização do artefato.

Para E4, a experiência de falar com um objeto, especificamente com um celular, mesmo que esse forneça retorno na forma de fala, não se difere muito de outras formas de falar com (ou falar através?) o celular em outras circunstâncias. E4 conta que falar com a Siri é “normal, não tenho problema nenhum. Eu uso a câmera para falar com o Instagram então pra mim falar com o celular é normal”. Ainda assim, E4 acha “divertido que o celular responda: olha, que legal!”.

E6 nos conta que, para ela, a utilização da fala como forma de interação “pra mim pessoalmente não é muito natural ainda. Porque pra mim parece que eu preciso estar falando com uma pessoa, que alguém esteja me ouvindo. Pra mim é muito natural mandar um áudio pelo Whatsapp, que teoricamente eu também estou falando com a máquina, mas eu sei que vai ter uma outra pessoa ali entendendo o que eu tô falando e que vai me responder devolta. Uma pessoa que eu sei quem é, uma pessoa de verdade [risos]. Com a assistente eu fico meio tipo, ahn, pensando “ah, uma pessoa de fora pode ver e achar que eu tô louca ou que é desnecessário” [...] acho que tem ainda um pouco desse estranhamento”

E9 nos conta é justamente o estranhamento de falar com um objeto que a impede de utilizar de forma mais extensiva a Siri: “mas é muito estranho, muito estranho. Por isso que eu não consigo usar como uma assistente. Porque é estranho, é estranho. Cara, eu tô falando com uma cadeira sabe [fala ironicamente e ri]. É estranho, sei lá, é estranho. Me sinto meio idiota assim falando com... sabe?”. E9 diz que esse estranhamento gera diversos outros sentimentos nela: “eu acho que eu me sinto meio idiota perguntando coisas pra Siri, eu me sinto meio idiota. É isso. E também por ela não funcionar de outras maneiras que eu achei que ela iria funcionar. Mas o principal é isso, eu me sinto meio estúpida falando com um pedaço de silício”.

E5 nos contou que, no início da utilização se sentia ansiosa para que o dispositivo funcionasse adequada aos seus comandos: “logo no começo quando eu tinha ela em casa eu ficava meio nervosa, eu queria que ela respondesse certo para as pessoas verem”. Entretanto, essa sensação passou com a utilização e “hoje é normal, me sinto a vontade. Eu quero que ela fale mais e me entenda melhor. Eu gosto, acho muito prático. Se eu pudesse acordar e dizer: “Google, faz meu café”, eu seria muito feliz [risos]”. A disponibilidade de utilizar o artefato, principalmente no formato de smartspeaker na casa, decorre da percepção de naturalidade que E5 tem em relação a utilização da fala: “É natural. A fala é muito simples, muito natural do humano. Falar com uma máquina talvez possa te causar o estranhamento de falar com uma máquina”. Se causou estranhamento para E5? “Acho que causou mais o constrangimento de outras pessoas verem que eu não estava conseguindo falar com a máquina. Tipo “ai, deu um comando e não deu certo”. [...] eu quero várias máquinas para eu falar com tudo, quero falar com toda a minha casa. Quero falar “Ok, Google” e simplesmente ligar luzes, ar condicionado, etc, acho muito legal”.

Para E10, ao falar com as assistentes existe o fator divertimento implicado, mas a interação não se difere muito de outras formas e códigos que poderiam ser acordados. Ela conta que “é um pouco divertido, mas pra mim é só uma forma... uma linguagem... tanto faz escrever... então pra mim é indiferente. É um pouco engraçado no começo, mas no fundo assim é só uma forma de linguagem. Se eu piscasse pra ela e ela entendesse eu acharia ótimo. Se eu piscasse 3x e era não e 2x era sim eu ia amar”.

E11 que, na maioria das vezes, acaba utilizando a funcionalidade de escrever das assistentes em hardwares com telas, ao invés de utilizar a fala como forma de interação. Ele conta que “acho que eu mais escrevo do que me comunico por voz com a assistente pessoal. Acho que é uma questão de segurança. Eu não gosto muito de como é a minha voz. E as vezes eu faço as perguntas já muito preconceituosas e arcaicas<sup>269</sup> pra ver o que vai acontecer quando ela precisar responder, e eu tenho vergonha de falar em voz alta esse tipo de pensamento. É anacrônico nos dias de hoje, embora muita gente pense. Por exemplo, relação de desemprego e empreendedorismo. Isso aí pode ser uma bobagem pra muita gente, mas eu penso que seja extremamente anacrônico ficar falando uma barbaridade dessas. A gente tem muito empreendedorismo, mas tem gente que não tem o que comer. E comentários racistas principalmente, acho que é a coisa mais anacrônica do planeta Terra, daí eu tenho vergonha né, eu escrevo só pra ver como ela vai reagir. Não quero que ninguém pense que eu penso isso de

---

<sup>269</sup> Em função das práticas de tencionamento com as assistentes, conforme apresentado anteriormente.

verdade”. Mesmo que praticamente não utiliza a forma verbal de interação com a Google Assistente, E11 diz que “eu me sinto ok. Não em sinto total desconfortável. Mas diria que é mais desconfortável que mandar um áudio pra alguém que eu conheça”. Entretanto, E11 conta que o desconforto que sente não se modificou com o tempo de uso: “o tanto que eu me sentia desconfortável antes eu me sinto agora”. A sensação de falar com um objeto é como “eu me sinto tipo ‘eu quero terminar logo’. Eu quero que seja uma experiência curta que vai passar. Eu preferia não utilizar a minha voz interagindo com a assistente pessoal”.

Para a entrevistada E12, a utilização da fala na interação representa um nível maior de adaptação do objeto aos sujeitos que o utilizam: “eu acho ótimo. eu acho que a gente adaptou o objeto pra nós. Porque antes, a gente pesquisava... tipo o Google. É a mesma coisa. O Google é a mesma coisa. Só que agora ele pôs a fala. E pAra pessoas cegas ou, tipo, se eu tô com um bebê, minha mão está atada. É ótimo. Por isso que eu digo, a gente fez pra nós. A gente que usa a voz, a gente que fala, a gente que se comunica pela fala. A gente quer que o robô faça o que a gente faz”. Mesmo que a utilização da fala torne a utilização mais fácil por assimilar forma de comunicação já utilizadas por humanos em outras situações, utilizar a fala com objetos e não com humanos pode ter um fator de estranhamento. E12 conta que “com a Siri [a primeira assistente que utilizou] eu estranhei. O começo foi bem estranho. Foi o mesmo estranhamento do *Facetime*, pra falar real assim. Porque a ideia de ver alguém que tá longe, é a mesma ideia. Era algo que parecia muito futuro, que eu não iria ver isso. Mas foi só passar o primeiro estranhamento, depois que você usa e que vê que é útil, passou. Mas eu estranhei sim”. O entrevistado E13 também ressalta o estranhamento em relação a não-usualidade e a novidade daquela forma de interação: “no começo, sim, eu estranhei um pouco, porque a gente não via isso. [...] quando é algo novo a gente acaba estranhando um pouco no começo”

## 6.5 RELAÇÕES ENTRE PÚBLICO E PRIVADO (OU, AS ASSISTENTES ESTÃO NO ARMÁRIO?)

Durante as entrevistas em profundidade os entrevistados foram questionados sobre os locais onde utilizavam as assistentes pessoais, possíveis utilizações de funcionalidades diferentes em espaços diferentes e sobre distinções entre o uso em público e no privado.

A entrevistada E1 disse que utiliza em qualquer circunstância, sem fazer distinções e aponta para o fator sociabilidade, tratado anteriormente: “as vezes até para mostrar olha o que

ela me respondeu, como eu fiz contigo”. A participante E4 possui reação similar e diz não ter “problema nenhum” na utilização na frente de outras pessoas ou em público e que “ inclusive eu faço muito, as vezes na frente dos alunos [para marcar horários de aulas particulares], inclusive já ensinei muitos alunos a fazerem isso... [eles me veem falando e] eles falam: “o que você está fazendo?”... [pareço] maluca né [risos]”.

O entrevistado E7 nos contou que utiliza no trabalho, em diversos cômodos da casa e no carro. Entretanto, sua utilização no trabalho é mais restrita “em função do barulho”. O barulho de fundo dificulta a compreensão da linguagem natural falada pelas assistentes, principalmente quando acessadas por computador, tablet ou celular, que não possuem microfones específicos para a captação de áudio para esta finalidade. Nesse sentido, a utilização em um local público fica inibida também por características que são do modo de funcionamento e da materialidade do sistema. Além da captação de áudio, a percepção de que existe uma forma (ou, até mesmo, uma fórmula de comandos) de falar específica para ser compreendido e fazer o sistema funcionar também pode ser um inibidor na utilização com outras pessoas presentes. Esse é o caso de E5, que nos aponta que por saber que existe uma técnica para acionar esses artefatos adequadamente, o receio de não domínio dessa técnica (que diz respeito ao modo de funcionamento das assistentes) a deixou ansiosa nas primeiras utilizações na frente de pessoas que possuíam esse domínio.

E5 nos conta que o constrangimento inicial de utilizar a assistente na frente de outras pessoas se deve ao fato de utilizar o inglês como idioma na interação: “logo no início parece que eu tinha medo de interagir, principalmente na frente de outras pessoas. Porque o inglês me deixa muito nervosa, eu começo a tremer, a ficar mal. Eu falo com ela em casa em inglês normal, mas na frente dos outros cria um constrangimento. E as pessoas que usam na minha frente, que vão lá, também, né, ficam ‘como eu vou falar com ela’, e eu digo ‘fala, ninguém tá te julgando’”. Somado ao idioma, se encontra o letramento na utilização do sistema: nas primeiras interações, E5 conta que se sentia insegura, principalmente porque utilizou em um laboratório de pesquisa, conjuntamente com outros pesquisadores que já dominavam a tecnologia. Nesse sentido, E5 aponta que tinha receio de não conseguir comandar adequadamente a assistente, ou seja, possíveis erros e não-funcionamento adequado seriam percebidos como uma falta da entrevistada e não uma dificuldade do próprio sistema. Percebe-se aqui que existe uma forma de utilizar, uma fórmula para falar, e que essa forma é apreendida na própria interação com o dispositivo, já que não existem alusões a essas formas nos materiais auxiliares das assistentes. Ainda assim, E5 conta que utiliza as assistentes, atualmente, na frente

de outras pessoas, sem fazer uma distinção, pois “usava em apartamento dividido e também na frente de amigos. Talvez para o celular tenha mais, por ser mais meu”.

Além das técnicas de fala específicas utilizadas para se “fazer entender” pelas assistentes, na fala de E5, por exemplo, emerge a ideia de que, caso algo corra errado na interação com a máquina, o erro é potencialmente humano e não maquinico. A noção de infalibilidade dos sistemas computacionais opera nos discursos humanos ainda que nossos cotidianos estejam recheados de pequenas histórias de erros das muitas máquinas com as quais convivemos. Nesse sentido, Broussard aponta que os computadores são tão “pervasive in every aspect of our lives that when something goes awry in the machine, we assume that it’s our fault, rather than assume something went wrong within the thousands of lines of codes that make up the average computer program” (2018, p. 25).

Alguns entrevistados mostraram maiores restrições na hora de conversar com as assistentes em público ou na frente de outras pessoas. O entrevistado E2, que utiliza a *Alexa*, por exemplo, disse que “eu prefiro usar sozinho, me sinto menos ridículo falando com uma caixa. Mas tem pessoas que eu já fico super a vontade. Minha mãe ri muito porque geralmente ela me liga e eu estou ouvindo música, daí eu atendo e já começo a falar ‘Alexa, stop’. Aí ela não ouve e eu sigo [dando o comando para ela parar]. E minha mãe ri e diz: ‘um dia você vai enlouquecer com essa caixa’. Mas eu prefiro falar sozinho, é estranho falar em público”. Nesse caso, mesmo com o passar do tempo, E2 segue preferindo utilizar sozinho, ainda que possua uma frequência elevada de utilização.

A ideia de ter pessoas ouvindo você falar com uma máquina, entretanto, gera constrangimento e estranhamento para alguns entrevistados, podendo, inclusive, inibir a utilização em espaços públicos. E6 conta que “eu sempre fico meio, acho meio estranho usar com pessoas envolta. Eu fico com a impressão que vão achar meio estranho que eu tô falando com uma máquina [risos]. Então eu meio que faço esses usos mais quando eu tô sozinha [...] eu acho que vão achar estranho que eu tô falando com uma máquina, aí já faço essa pré-censura assim”. A participante E10, comenta que gosta de utilizar a assistente “em qualquer lugar” e que não se importa “nenhum pouco” de utilizar em público “até porque eu não sou de ficar xingando”. Entretanto, a reação das pessoas em espaços públicos acaba inibindo essa utilização: “as vezes eu fico sem graça porque eu fui usar no trem e todo mundo me olhou. E pra mim era uma coisa tão banal né, tipo, todo mundo tem essas coisas, todo mundo tem celular, mas não é todo mundo que usa”. Para E10, essa questão muda apenas quando utiliza para fazer brincadeiras, que aí “é só em casa, ou pra visita [...] Pode ser na casa dos outros, mas nunca em lugar público assim”.

Para E9, ocorre o contrário: a utilização na frente de outras pessoas só ocorre se for uma demonstração ou uma piada, não como utilização com outra finalidade. E9 explica que não utiliza em público, pois “eu acho meio bizarro uma pessoa ficar conversando com o celular sem que tenha uma outra pessoa de fato, de verdade no outro lado. E aí quando tu tá fazendo perguntas pra assistente social [risos pelo erro]. Assistente social! Assistente pessoal. Mas é bizarro tu ficar conversando com um robô na rua, agora, por enquanto, talvez daqui a pouco não seja mais... eu ainda me sinto desconfortável [...] que não seja uma demonstração tipo piada”.

A participante E14, que utiliza principalmente a Google Home, aponta que, mesmo dentro de casa, não se sente confortável interagindo na frente de outras pessoas, dependendo da situação. Para E14, “eu vou dizer que tem uma questão cultural no Brasil assim... Que quando tem faxineira em casa eu não gosto de falar, porque parece que tu tá querendo se mostrar. Eu vejo que sempre que elas limpam elas desligam o interruptor e o interruptor tem que ficar ligado pra conseguir fazer o uso dele pelo Google Home. Daí assim, quando tem faxineira, quando tem alguém, eu costumo não usar tanto”. Nos casos de E6, E10 e E14 a utilização em público ou na frente de outras pessoas é inibida pela antecipação de uma percepção negativa por parte das outras pessoas presentes, ou mesmo por experiências prévias que indicam uma percepção negativa ou um estranhamento nos demais.

Dentre os entrevistados, E11 foi quem apresentou mais particularidades na forma de utilização das assistentes em espaços públicos e privados. Por motivos pessoais, E11 não tem o hábito de levar o celular, onde acessa a Google Assistente, para casa: utiliza apenas na faculdade e no trabalho, ou seja, quando está fora de casa em atividades que não são de lazer. Ele nos conta que leva o celular para casa “só eventualmente” e que “deixo meu celular no meu trabalho ou faculdade [na sala do grupo de pesquisa]”. Quando leva para casa o celular, a utilização ocorre mais na cozinha, para cozinhar.

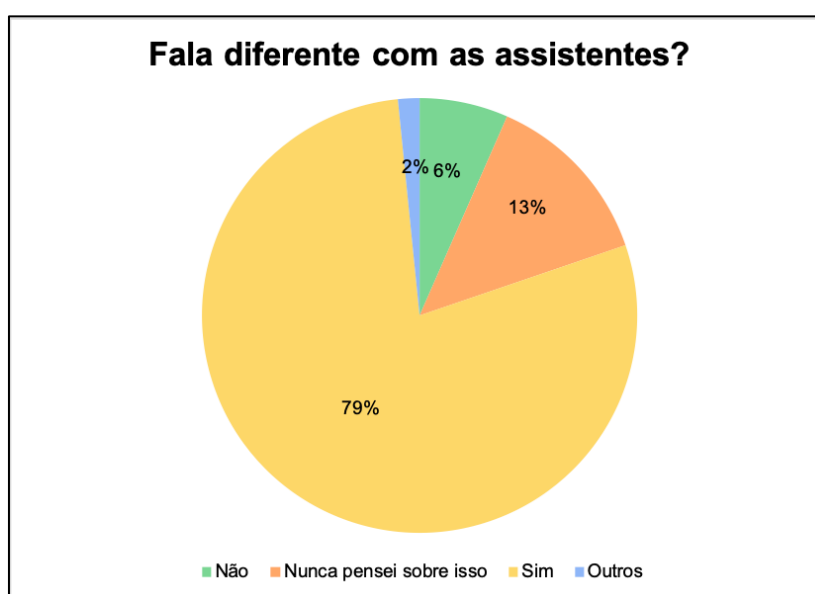
## 6.6 COMO VOCÊ FALA COM A SUA ASSISTENTE?

No tópico anterior, aparece já na fala da entrevistada E5 a existência de formas de falar adequadas para a interação com esses sistemas, apontando para uma não-naturalidade dessa interface. Ou seja, assim como outros tipo de interface, falar com uma assistente exige uma aprendizagem e o domínio de técnicas específicas: vale lembrar da respondente da etapa

quantitativa que informou no campo outros, em diversas perguntas, que nunca conseguiu fazer com que a assistente a entendesse. Ao longo das entrevistas em profundidade, perguntamos aos participantes sobre como eles percebiam seu modo de falar com as assistentes, se ocorriam alterações em relação a sua fala usual.

Os respondentes do formulário quantitativo foram perguntados sobre a forma como falavam com as assistentes pessoais digitais, se falavam de uma forma diferente do que a forma habitual no cotidiano. O gráfico abaixo apresenta os resultados para essa pergunta.

Figura 58 Gráfico Fala diferente com as assistentes?

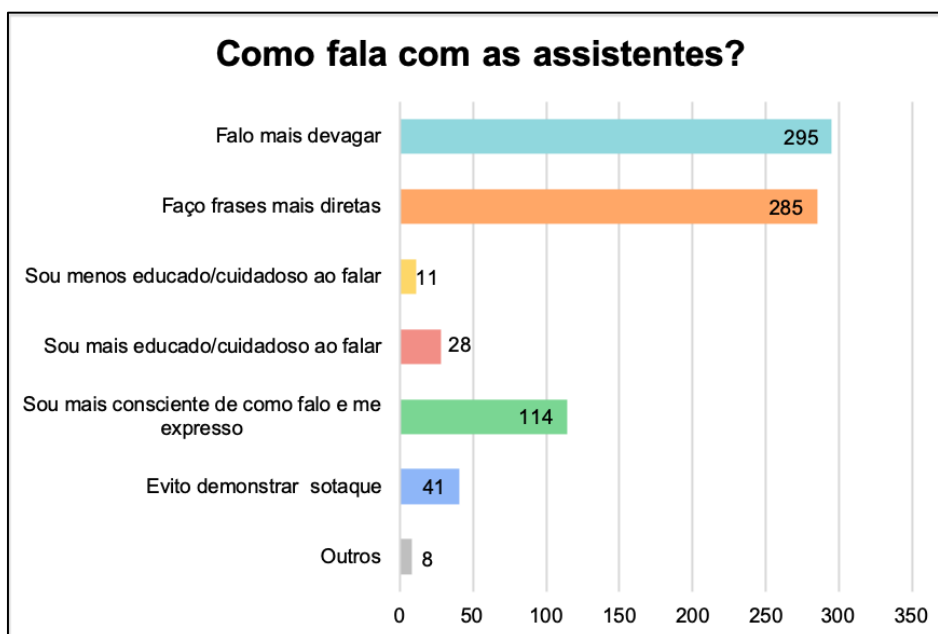


Fonte: A autora

Sobre a forma de falar com as assistentes, 79% dos respondentes afirmou que fala diferente com as assistentes do que fala usualmente, 13% afirmou que nunca pensou sobre essa questão, 6% informaram que não falam de forma diferente e 2% optaram pela categoria outros. Os dados apontam que uma quantidade elevada dos usuários percebe alterações na sua fala no momento da interação das assistentes, com poucos usuários apontando que não alteram a forma de falar. Nesse sentido, os dados reforçam a ideia de que existe um conhecimento e uma técnica adquirida na forma de falar para uma interação com estes dispositivos e que, portanto, a utilização da fala não seria necessariamente natural. Os retornos da categoria outros serão discutidos a partir do próximo gráfico.



Figura 59 Gráfico Como fala com as assistentes?



Fonte: A autora

Dos respondentes que apontaram notar diferenças na forma de falar com as assistentes, 295 deles disseram que “falo mais devagar”, 285 disseram que “faço frases mais diretas”, 114 disseram que “sou mais consciente de como falo e me expresso”, 41 informaram que “evito demonstrar qualquer tipo de sotaque”, 28 disseram que são “**mais** educado/cuidadoso ao falar”, 11 disseram que são “**menos** educado/cuidadoso ao falar” e 8 optaram pela categoria outros.

As formas de falar “mais devagar”, com “frases mais diretas” e sem “demonstrar qualquer tipo de sotaque” se relacionam com fórmulas utilizadas na fala que visam maximizar o potencial de compreensão dos comandos pela máquina. Os retornos elevados de respondentes nessa categoria apontam, por um lado, para a limitação das interfaces de voz de compreender a fala humana, e, por outro, de técnicas e aprendizados desenvolvidos nessa interação, que desmontam a percepção de naturalidade por trás da utilização dos comandos de voz. A voz pode ser um artifício útil para os usuários quando em situações que desejam manter as mãos livres ou sua atenção focada em outra questão, entretanto, a voz não é uma forma de interação com o computador mais natural, justamente porque a maioria das pessoas não fala naturalmente com esse artefato, como apontam as técnicas utilizadas. O fato de que os participantes apontaram também serem “mais consciente de como falo e me expresso” na utilização das assistentes pessoais digitais pontua a artificialidade dessa situação, ao retirar a fala do contexto no qual os usuários estão habituados, fazendo-os, assim, perceber com mais clareza a forma como falam.

Além disso, a fala humana em si mesma é uma construção sociocultural voltada para a comunicação, a qual também precisamos aprender no início da vida.

Já as formas de falar “mais educado/cuidadoso” e “menos educado/cuidadoso” se relacionam mais com formas e respostas sociais concedidas as assistentes pessoais digitais, que se relacionam com o antropomorfismo desses artefatos, assim como com o imaginário em torno das inteligências artificiais e robôs. Essa questão será tratada em profundidade a partir das entrevistas, entretanto, a partir do formulário quantitativo, fica visível o fato de que os seres humanos não utilizam exatamente a mesma forma de comportamento com computadores que utilizam com pessoas, contrariando estudos que apontam uma transposição completa nos modos de ação sociais entre humanos para a interação entre humanos e computadores, quando os últimos possuem pistas antropomórficas (NASS, anos diversos).

Na categoria outros, duas pessoas afirmaram falar mais alto do que o habitual; três pessoas afirmaram falar de forma que o computador entendesse (seja mais direto, mais específico ou com mais cuidado); uma pessoa pontuou que fala em inglês. Um dos respondentes disse que “Falo como imagino que se deve falar com um computador”, deixando implícito que existe uma forma específica nesse formato de interação que compreende tanto maneiras adequadas quanto funcionais, no sentido de se fazer entender.

Outro respondente disse que “Falo normalmente como se estivesse perguntando alguma coisa para pessoas normais. Não preciso ter nenhum cuidado especial já que o assistente compreende fácil”, apontando que não percebe alterações na forma que formula suas frases ou utiliza outros recursos como entonação, altura da voz, etc. Um participante disse que “queria poder mandá-la tnc [tomar no cú] mas o sistema é burro e não deixa =D”, apontando para uma forma de falar com a máquina que é não educada ou cultivada socialmente.

Entre as técnicas percebidas para se fazer compreender estão: falar mais alto, utilizar uma entonação específica para a palavra-chave, falar de forma mais devagar, evitar sotaque, utilizar uma fórmula de frase simples, saber quais palavras acionam quais comandos, etc. Assim como o aprendizado da fala, da linguagem e do jogo conversacional entre humanos, existe na interação com as assistentes um aprendizado que também ocorre apenas no fazer. Ainda que se utilize da linguagem natural falada, dos idiomas específicos e da forma diálogo, a forma de falar com as assistentes são diferentes das formas como falamos com outros humanos, nos fatores anteriormente apontados (e também na compreensão daquilo que é ou não permitido falar).

Quando perguntada, por exemplo, sobre as formas de falar com a Siri, E1 faz um comparativo entre a interação utilizada com outras tecnologias que não baseadas na voz, como buscadores. E1 conta que “eu percebo que às vezes eu preciso alterar um pouco a forma [da

minha fala]. No Google, por exemplo, eu coloco pizzaria e ele me dá pizzarias aqui perto. Se eu falo pra ela ‘pizzarias’, ela via me dizer aqui está o que encontrei sobre pizzaria e mostrar por exemplo a história da pizza”. Nesse sentido, a alteração na fala, que faz com que E1 “preciso especificar bem o que eu quero dizer. Ela não entende o contexto”, seja um comparativo tanto as formas conversacionais utilizadas com humanos quanto as formas de interação utilizadas com sistemas de busca digitais. E1 aponta que, a forma conversacional entre humanos, especialmente presencialmente, fornece mais elementos para avaliar se a comunicação está sendo efetiva ou não: “que nem quando a gente tá conversando, eu posso tá te explicando uma coisa e te perguntar se tu entendeu. Mas eu sei que tu entendeu porque a gente tá aqui conversando, se olhando, interagindo, a forma como se fala. Ela só escuta o que tu tá falando”

E2 percebe alterações principalmente na forma como projeta a sua voz no espaço, uma vez que para ser compreendido pela Alexa, “eu tenho que impostar mais a voz para falar com ela, se eu falo normalmente ela não entende”. Essa percepção se assemelha com a da E9, que percebe mudanças na entonação para conseguir acionar adequadamente o artefato: “eu acho muito bizarro, porque se eu falar com ela “E aí Siri” [exemplifica utilizando uma entonação de fala corriqueira entre humanos] ela não vai entender. E é assim que eu falo. Tu tem que dizer “E aí Siri” [faz uma entonação bem demarcada e lenta], com uma determinada entonação que tu gravou quando comprou o telefone e tal. E a fala dela de volta, essa coisa mecânica e tal, é robotizada, é feito pra que as coisas sejam entendidas, é tudo muito parametrizado. E a nossa relação, relação de pessoas com pessoas, ela não é assim tão parametrizada né. Não tem essa coisa tão.... se eu disser uma coisa um pouquinho fora do tom tu não vai me entender? Não, tu vai me entender. Tu pode me interpretar diferente, mas tu vai entender o que eu estou querendo dizer”.

E3 conta que, para uma comunicação efetiva, sua técnica é minimizar a quantidade de elementos da frase, deixando aqueles que são essenciais para a compreensão do comando: “com a Alexa eu tento falar o mínimo de palavras, compilado, já pensando como que o programa vai interpretar melhor. Pra ter mais chance de ser entendido. E com pessoas eu tô sempre buscando o feedback delas pra saber se elas tão entendendo o que eu tô falando, principalmente se é alguma coisa complicada. Então eu estou sempre fazendo perguntas pra me situar se a conversa tá onde eu queria que tivesse”. Além da modulação específica e também menos educada, uma vez que a educação é uma adição de complexidade na estrutura da frase sem um efeito prático, E3 também percebe a falta de feedback como o principal elemento de diferenciação de uma interação humana.

De forma similar, E11 aponta que preza pela concisão como técnica de fala com as assistentes: “acho que eu tento ser conciso, porque eu penso que não tem a subjetividade de estar conversando com uma pessoa. Então não vai adiantar eu dar um nó, assim como eu tô dando na resposta contigo, pra algo que vai ser objetivo. Então eu penso em dar o que eu quero logo”. Além disso, E11 avalia que “eu acho que sou mais técnico. Mais objetivo. Tento fazer perguntas o mais concisa possíveis”. Ou seja, tenta eliminar elementos na frase que não servem a um propósito de ação específica. Entretanto, E11 nos conta que, quando em 2014 utilizava a Siri [hoje utiliza a Google Assistente] “ela era a velha da praça total, não entendia nada do que eu falava. E aí eu me sentia pior ainda, do tipo, eu não consigo ter dicção o suficiente para tipo “pesquisa tomate no Google”, sabe? Parecia que tinha esses outros efeitos, de eu não sou capaz de falar algo pra pessoa entender, ou pra aquilo ali entender”. Nesse trecho, volta a se repetir o fenômeno anteriormente analisada na fala da entrevistada E5, onde o sujeito compreende que a resposta inadequada da máquina é na verdade uma inadequação de utilização sua – já que as máquinas são percebidas como pouco falíveis.

E11 vai apontar ainda, como E9 e E3, a falta de outros elementos da comunicação verbal para se fazer compreender e avaliar o nível de compreensão do outro. Para o entrevistado, a interação por voz com as assistentes não é agradável pois “eu acho que quando eu tô falando contigo tu tá olhando na minha cara e vendo que eu olho lá pra cima pra pensar numa palavra minúscula e tu tá vendo o meu corpo e tu tá vendo as minhas mãos, e eu consigo abranger muito mais a minha linguagem com isso tudo, mas quando eu tô só gravando parece que me limita em diversas coisas sabe. Eu prefiro a clareza da palavra escrita do que uma linguagem de voz”.

E4 nos contou que a distinção da fala entre “estou falando com um humano” e “estou falando com uma máquina” nem sempre está em sua mente: “claro, tem um momento que tu acaba se tocando, ah é, tô falando com um assistente virtual, não tô falando com um ser humano. mas de forma geral eu não penso sobre isso quando eu tô usando”. Além disso, ela pontua que existem estratégias de comunicação específicas e que no geral “eu sou meio grosseira [risos]; sou mais direta, eu basicamente pego ela e falo: “agenda, para dia 10, para tal pessoa, tal horário”. Se eu tivesse uma secretária humana eu jamais falaria assim com ela [risos]”.

A entrevistada E12 também percebe a diferença entre a forma de falar com humanos e de falar com a assistente na hora de dar comandos ao invés de fazer pedidos, mesmo que o que ela faz hoje com as assistentes seja muito parecido com funcionalidades exercidas anteriormente por trabalho humano. A diferença estaria no fato de que, com a assistente, E12 diz que “ eu falo tipo mandando. Você trabalha para mim. Faça o que eu quero. Eu não falo tipo, “Oi Alexa”, “Ai Google, tudo bem, bom dia, muito obrigada” [imita uma voz mais fina ao

falar essas frases]. As vezes no final eu tenho vontade de falar obrigada. Não sei porque. Acho que porque eu tô acostumando tanto que quando vê vem na cabeça assim ‘obrigada’. Mas eu não respondo, porque eu entendo que é uma máquina. Mas no fundo vem [risos]”. Na fala de E12, apesar de perceber que fala de uma forma muito mais direta, sem rodeios e educação com a assistente, também reconhece um eco de formas de tratamento que utiliza na interação humana, como pontuado por Nass (1996; 2010).

Mesmo que exista um modo de operar com humanos que tende a se repetir, a forma de interação estabelecida para E12 é completamente diferente daquela utilizada com humanos: “eu já tive assistentes né, eu sou professora. Não, é completamente diferente [do que eu trato humanos]. Porque assim, com ser humano, você tem que ter mais educação, porque você entende que você precisa dele de um jeito que ele tem que gostar de você. Você precisa conquistar o ser humano. A máquina você faz o que eu mando e acabou, não quero saber se você gosta ou não. Eu tô de mau humor e eu falo com você como eu quero... com o ser humano você precisa pensar, você precisa construir o relacionamento. Com a máquina não. Então eu falo diferente sim”. A fala de E12 mostra que ela percebe que, na interação entre humanos, mesmo quando hierarquizada, existem sanções reais a serem aplicadas quando os protocolos de educação não são utilizados. Essa sanção pode ocorrer simbolicamente através de uma resposta da assistente para uma fala muito mal-educada, entretanto, não existe uma sanção material (na forma de descumprimento da ordem, por exemplo). Justamente por saber que não existe essa sanção material, E12 não se sente compelida a falar de uma forma educada ou mesmo humanizada.

A entrevistada E14 nos conta que percebe com clareza a forma na qual os comandos precisam ser ditos para que o efeito desejado seja obtido: “é necessário, porque se não ela não me entende. Então eu preciso falar a frase exatamente naquele formato, mas é uma questão mais de limitação do dispositivo do que uma questão minha. Eu talvez falaria com ela de outra forma”. Por trabalhar e pesquisa no âmbito das tecnologias digitais, E14 percebe com clareza os elementos na forma de falar que são dela mesma e os que são necessários por demandas dos sistemas, apontando que esse formato não é necessariamente o mais confortável para ela. A forma de trata também não se assemelha a utilizada com seres humanos, “porque eu não peço por favor, mas eu sou educada com ela... Mas é, eu não pediria para um ser humano “Turn on the lights” eu diria “Could you turn on the lights”. Eu acho que é muito mais uma coisa de comando mesmo. Talvez mais no início, que eu falava “thank you” quando ela fazia alguma coisa, que eu digo que ela fez bem. Mas é quase uma comemoração do que um agradecimento”. A E14 percebe elementos nas frases que alteraria em uma relação humana, por exemplo, mas

que usa tranquilamente na interação com as assistentes, gerando frases que são efetivamente comandos e não solicitações. Além disso, E14 aponta que a utilização de elementos de educação, como o agradecimento ao final de uma tarefa, não foram sempre utilizados por ele como a repetição de um padrão cultural, como apontaria Nass (1996; 2010), mas sim mais uma forma de comemorar/expressar felicidade por aquela tarefa ter sido finalizada com sucesso.

A entrevistada E5 percebe que, com o tempo e consistência de utilização, a forma de compreensão entre ela e a Google Assistente se alterou: “quando eu falo algo errado, ela sabe que eu errei. Ela percebe alguns erros, como por exemplo eu dizer a mesma música de forma errada várias vezes. Ela já sabe qual música eu quero dizer [risos]”. Através das interações repetidas e das possibilidades de machine learning desses sistemas, a máquina se adaptou a certas formas de comunicação “erradas” da entrevista, e a E5 também criou formas específicas de se adaptar ao modo de funcionamento da assistente. E5 nos conta que “eu já me adaptei a tudo. Aos comandos, ao jeito que ela fala, as pausas que eu preciso fazer pra falar com ela, o jeito que eu sei que ela vai me entender. Mesmo quando precisa forjar um sotaque, assim, pra pedir alguma coisa sabe, por exemplo, *Anita* [fala o nome da cantora brasileira com sotaque americano], até isso torna divertido, eu sei como fazer”. A fala de E5 aponta para uma relação simbiótica, de mútua adaptação para que a comunicação funcione efetivamente entre as partes – ainda que só uma delas tenha consciência desse processo.

E6 informou que “eu percebo que eu falo mais devagar, mais pausado, para fazer o máximo pra ela entender cada palavra que eu tô falando”. Entretanto, ao ser perguntada sobre a educação que utiliza com o sistema, disse que “eu acho que [sou educada] sim [risos]. Talvez eu já tenha falado algum palavrão para ver o que elas iriam responder, mas como quando elas são educadas de volta eu meio que reajo como se ela tivesse me dando um gelo e eu “tá, tá bom, entendi, vou me manter na linha”. Então eu não sou grosseira com elas eu acho”. Além das mudanças na fala apontadas por E6, emerge também a noção de que a forma como os sistemas respondentes representam uma forma de sanção ao comportamento não-educado ou até abusivo de usuários com as assistentes, mesmo que as respostas em geral não sejam muito contundentes.

Já os entrevistados E7, E10 e E13 disseram que percebem poucas ou nenhuma alteração em sua forma de falar, mesmo quando comparada com a forma que utilizam para falar com outros seres humanos. E7 nos disse que percebe pouca alteração na sua fala com as assistentes: “eu não vejo nada de diferente. Eu só sei que eu não estou lidando com uma pessoa, mas eu não trato de nenhum jeito diferente assim. A não ser aquela questão que eu falo de um jeito mais devagar, hoje em dia eu falo como se eu estivesse conversando com uma pessoa. nesse sentido assim, eu não fico me preocupando, eu pego e falo e se entendeu, entendeu, e se não entendeu

eu pego e **xingo** e vou de novo”. Quando a Siri não compreende o comando, entretanto, E7 age, possivelmente de forma diferente do que agiria com um ser humano que não compreendeu uma fala em uma comunicação: xingando.

E10 conta que, apesar de achar que fala normalmente com as assistentes, utilizando os mesmos padrões sociais que normalmente utiliza, percebe que as pessoas ao seu redor não necessariamente o fazem: “então, eu falo [com as assistentes da mesma forma que com seres humanos]. Mas eu percebo que o meu marido não fala e a minha mãe também não. Mas eu falo, acho que por ter essa formação de ter convivido com gente que trabalha com ia, entendido um pouco a lógica, desmitificou... então eu falo normalmente... mas eu acho muito engraçado xingar a Siri, ficar brabo”. A entrevista E10 atribui principalmente a sua formação de pesquisa, na área da comunicação, o fato de não humanizar os artefatos. Já E13, que utiliza diversas assistentes para promover automação residencial que o auxilia nas atividades do cotidiano que não pode realizar devido a baixa mobilidade, percebe apenas pequenas alterações na sua forma de falar: “hoje eu consigo falar naturalmente. Algumas vezes a gente precisa adaptar um pouco na forma de falar pra que a gente possa ser entendido...algumas palavras ainda são limitadas no entendimento. A gente precisa falar de um jeito mais simplificado pra que ela entenda”.

## 6.7 A VOZ E A FALA DAS ASSISTENTES

Vozes e formas de falar (e de ouvir) diferentes propõem interações diferentes com os dispositivos, despertando ou inibindo elementos relacionados à imaginação. Nesta seção, analisaremos a forma como os entrevistados percebem as formas de escuta das assistentes pessoas digitais, assim como as formas de falar e as relações que eles estabelecem com estes dispositivos mediados pela voz e pela forma conversacional.

O entrevistado E2 possui uma trajetória mais longa e próxima com a Alexa, pois foi beta tester da versão brasileira por cerca de seis meses, antes dela entrar no mercado por aqui. Ele conta recebeu “um email da Amazon perguntando se eu queria ser beta tester. Eu achei que era vírus e quase não respondi. Mas respondi, aceitei e chegou [a Alexa]. Aí todas as sextas-feiras eu recebia as atualizações e também algumas perguntas para eu testar”. Ele conta que a Amazon solicitava que os usuários na fase de teste utilizassem bastante e testassem formas conversacionais com a assistente e que foi possível notar a diferença no desenvolvimento do sistema ao longo do tempo: “quando veio a caixinha já veio pedindo pra eu pedir uma música

de Bruno e Marrone, que eu me recusei, claro [risos]. Eles sempre tinham essa tentativa de fazer você interagir mais com ela, pra ela aprender português. No começo ela respondia algumas coisas em inglês, mas com sotaque em português, ou não entendia palavras em inglês [como títulos de músicas]. Tinha umas coisas que era de morrer de rir. Mas foi melhorando muito a pronúncia dela”.

A entrevistada E10 conta que, a partir do uso, foi possível perceber que a Siri apreende mais coisas do que apenas o conteúdo das palavras em uma frase: “eu cheguei à conclusão que eu acho que a Siri entende de entonação de voz. Se o meu marido fala, porque daí as vezes eu ponho e ele xinga, por exemplo. Ah, o que ele estava falando esses dias: Ah, Siri você é muito burra [com voz direta, entonação de brabo]. E aí eu falo “Ah Siri, você é muito burra” [com uma voz mais simpática, como se estivesse falando uma coisa querida], ela não entende. Então só pode ser entonação, tom de voz, e tal”.

E10 aponta que a reação diferente para um mesmo conteúdo, que foi modificado apenas na materialidade da voz, da entonação e da expressam, modificam a interpretação da assistente sobre o próprio conteúdo falado. Na opinião de E10 o fenômeno indica que a assistente é treinada também para reconhecer padrões relacionadas a expressão da fala, não apenas às palavras em si.

Para E7, a não naturalidade da fala e da linguagem das assistentes é um dos elementos que ela considera como uma barreira para a utilização. Durante a entrevista, E7 nos contou que estranha “eu acho que um pouco da linguagem. Cada pessoa tem um jeito diferente de falar e consegue buscar formas e entonações, e acho que isso não tem ainda lá [...] E a entonação da pessoa pra uma resposta, a voz vai subir ou vai descer, vai ter uma modulação, que eu não vejo na Siri falando. Ela sempre vai ter uma entonação parecida para tudo. Falta uma variação”. Uma humanização maior da forma de falar das assistentes teria um efeito de neutro à positivo na interação: “porque é legal tu falar e não ter aquela resposta super quadradinha. É mais agradável ouvir, ter uma voz que tenha essa cadência mais parecida com uma voz humana. Por esse lado eu acho legal. acho que tá mais pra positivo”.

Já E8, que se identifica como um aficionado por tecnologias digitais em geral e pelas assistentes em específico, acredita que a voz é um das principais preocupações da empresas produtoras: “a naturalidade da voz, isso é uma coisa que eu acho que os fabricantes de todas elas estão cuidando, de essa voz ser o mais interativa possível, o mais natural possível. E isso é para que? Para que essa assistente se torne um periférico essencial dentro de qualquer casa. A Amazon propagandeia que você deve colocar uma em casa ambiente da sua casa”. O entrevistado E13 concorda com a leitura de que existe naturalidade e similaridade com o



humano na forma de falar das assistentes que estão hoje no mercado, apontando que “eu diria que na voz lembra muito um ser humano, tá bem natural a forma de falar. Voz e a forma de responder, lembra muito um ser humano”.

O entrevistado E7 utiliza apenas a Siri, o E8 tem experiência extensiva de utilização com a Siri, a Alexa e a Google Assistente, enquanto o entrevistado E13 utiliza extensamente apenas a Alexa e a Google Assistente. Para a entrevistada E14, que utiliza com frequência elevada a Google Assistente, mas possui experiência também com a Alexa e a Siri, dentre as três aquela que possui mais fluidez de fala e de voz seria a Alexa:

“eu acho ela muito boa pra conversa, eu acho que ela consegue... ela faz ‘hmm’, ‘uhum’, eu acho que ela faz alguns **trejeitos**... trejeitos [risos], mas ela faz alguns sons como se ela estivesse ali conversando contigo, ela é muito natural”.

A utilização da palavra **trejeitos** para se referir as formas de falar da Alexa que imitam as pausas humanas é significativa. De acordo com o dicionário Houaiss da língua portuguesa, a palavra significa: 1) jogo de mãos, destrezas manuais, prestidigitação; 2) movimento do corpo, por exemplo das mãos, dos braços, dedos, cabeça, voluntário e involuntário, que revela estado psicológico ou intenção de exprimir ou realizar algo; movimento, gesticulação; 2.1) gesto ou arremedo cômico; 2.2) contração ou movimento que altera a expressão do rosto; esgar, careta; 2.3) tique nervoso. As definições apontam para um conjunto de traços que diz algo específico sobre aquele ser – entretanto, todas as definições colocam os trejeitos como traços que se exprimem através do movimento de um corpo orgânico, que dão a ver uma personalidade ou estado de espírito particular, ao mesmo tempo em que podem ser também a emulação desses traços (como no viés cômico). O lapso freudiano ou ato falho (FREUD, 1901) foi rapidamente reconhecido e corrigido pela entrevistada, entretanto, deixa entrever a percepção de que as formas de falar dizem tanto sobre uma personalidade de algo/alguém quanto de uma corporalidade, da existência de um corpo físico a partir do qual essas particularidades encontram lugar de expressão. Nesse sentido, a voz e as formas de falar tipicamente humanas (como os “hm” e “uhum” entre as frases) aparecem como intimamente entrelaçadas com a própria ideia de humano e de relações humanas.

As formas de se relacionar e interagir com a proliferação de vozes em artefatos técnicos - como o Waze, por exemplo, que com frequência disponibiliza vozes diversas, famosas ou anônimas, para conceder as instruções de trânsito para os motoristas – são variadas entre os entrevistados. E10, por exemplo, nos conta que ela e o marido gostam de brincar com as vozes, inclusive com a dela própria: “do Waze a gente teve uma experiência engraçada. Eu e o meu marido... eu gosto muito de dublagem, aí ele me fez gravar com umas vozes engraçadas os

comandos. E aí as vezes a gente tá na estrada e entra a minha voz. É bem divertido, experimente. Você escolhe a voz, e aí tem como personalizar e eu gravei uma voz”.

A entrevistada E1 associa invariavelmente a assistente ao sexo feminino, especialmente através da relação com a voz: “pra mim a voz [da assistente] é só feminino”. A voz, nesse sentido, atua como indicativo de outras características do corpo produtor daquele som – que esse corpo seja feminino por padrão pelas empresas e como as relações entre homens e máquinas é, nestes casos, atravessada por questões de gênero é o tópico do próximo capítulo.

A menção da possibilidade de, nas assistentes, fazer variar, de acordo com cada marca, elementos como o timbre da voz, o gênero, e os sotaques, gerou curiosidade de experimentação entre alguns entrevistados. E2 disse que, agora que sabe que é possível, em algumas assistentes, trocar as vozes utilizadas: “como você falou das vozes, eu gostaria de utilizar outras vozes. Fiquei curioso”. Já E3, apesar de inicialmente afirmar gostar da voz que utilizava na *Alexa* (feminina em inglês americano), afirmou que gostaria de trocar para o inglês britânico, flertando, imaginariamente com a ideia de um mordomo: “eu fiquei pensando agora num mordomo com sotaque inglês, hein, ia ser massa”, apesar de utilizar a voz feminina como preferência, disse que “agora eu já mudei de ideia, acho que esse mordomo inglês promete”.

Para E4, a interação através de voz, com uma similaridade com o humano, leva a uma forma de tratamento mais antropomórfica: “talvez por usar a voz eu acabe humanizando mais a Siri, tratando um pouco melhor e tipo, é engraçado isso, nunca tinha pensado. Mas a utilização, acho que é mais rápido e mais fácil, [a voz] me deixa mais confortável, mas ao mesmo tempo eu sempre checo se ela marcou bem em algum momento”. Nesse sentido, a voz teria um papel central na antropomorfização desses sistemas, que, em alguns casos, pode levar a um relacionamento mais afetivo com esses artefatos. E5, por exemplo, atribui o laço efetivo que desenvolveu com a Google Assistente principalmente a habitualidade da voz desse sistema: “eu não abriria mão da Google e abriria mão da Alexa. Eu sinto assim que depois de um tempo eu criei um afeto com a voz [do Google]. Eu não consigo trocar de voz. Eu já tentei. Eu não consigo, não consigo de jeito nenhum. já tentei mudar de sotaque, eu gosto, acho divertido mudar, mas eu não consigo usar. Mas isso me ajudou muito, o fato de ter ela, com a mesma voz, no meu celular e em casa, cria uma familiaridade”. A entrevista brinca até com o papel das assistentes na interação com bebês pequenos e com o próprio papel que a fala possui na constituição humana: “tem pais que não querem expor os filhos a gadgets. Eu já ia botar no quarto do bebê. Já vai aprender a falar “Ok, Google” assim, não é mamãe é “Ok, Google [risos]”. Apesar da brincadeira, fica a questão: veremos surgir, além do manhês e da fala dos bebês, um terceiro tipo de fala de robôs para o desenvolvimento infantil? E como repensar a

constituição do humano, tão marcada pela fala de um Outro cuidador, se os cuidados e os processos de introdução à linguagem ocorrem através de um Outro-máquina?

Durante a entrevista, E6 pontua que existe para ela, assim como para E5, um elemento de habituar-se com a voz utilizada como padrão pelos sistemas e que, uma vez criado esse vínculo com a voz, ele tem o potencial de gerar a sensação de se tratar de algo que nos é familiar. A voz acaba sendo levada em conta na hora de preferir um sistema ou outro: “talvez [eu tenha preferência pela] Siri, pelo fato de eu ter começado a usar antes. A questão da voz assim, parece que é mais familiar [...] Na Siri eu acostumei com aquela voz dela e eu nunca mexi. No Google, depois que eu adquiri o smartphone, eu soube que dava pra trocar, mas eu nunca me interessei em trocar, também me acostumei com aquela voz”. Aqui entra em jogo as noções de familiar e não-familiar, como um mesmo objeto pode nos ser familiar e estranho ao mesmo [o *Umheimlich*], já que, por exemplo, E6 também afirmou que não se sente confortável e sente um estranhamento ao falar com as assistentes por voz.

A própria ideia de identidade e de personalidade humanas, mesmo quando não percebida através do conteúdo das falas das assistentes, é percebido através da existência e das características da voz. Para E10, por exemplo, mesmo que “eu não percebo muito forte. claro que tem, até porque a voz tem uma identidade. Acho que tem mais uma identidade de voz do que outras coisas além disso. E a siri não gosta de tomar bronca nem que fale alto com ela, se não ela não responde. Eu ela responde. isso é personalidade né...”. Além disso, a reação da assistente a forma e ao tom de falar do usuário, não apenas ao conteúdo da fala, também conta na resposta que ela fornece. A entrevista E10 também pontua que existem elementos da voz que podem influenciar nas formas de interação que os usuários irão estabelecer. Para ela, “eu acho que a tonalidade da voz importa, mas se é o homem ou mulher não. Mas claro, se você pega uma voz mais amistosa e uma voz mais grosseira, eu vou preferir e amistosa”.

A entrevista E12 percebe de forma similar: as assistentes não teriam características semiliares ao humano que ela consiga perceber. Apesar disso, por ter uma voz: “a partir do momento que você escolhe uma voz, você já está querendo dar uma personalidade. Porque as duas [Alexa e Google Assistente] são uma voz feminina, com inflexão. Assim, eu acho que a Siri tinha menos inflexão do que as novas têm. Então se tu coloca inflexão, tu quer dar uma personalidade, tu tá fazendo ficar mais parecido com um ser humano”. De forma complementar, o entrevistado E11 compreende que a voz se constitui como um elemento de design que visa criar a ideia de uma personalidade: “eu penso que sim, a voz é um artefato, um artifício para construir uma personalidade [...] Eu acho que foi pensando para ser desse jeito que não é ao acaso. é justamente para ter esse grau de fácil proximidade, de criar vínculo fácil. porque é uma

voz neutra, tá sempre ali pra ti, e é uma voz feminina, é muito mais fácil de tu... parece que pra ti estabelecer uma relação é muito mais fácil”. A voz, de forma geral, possibilitaria criar um vínculo de proximidade e seria potencializada por outros elementos que compõem, como o gênero, o tom, etc.

Essa noção E11 ressoa com a percepção de E9, para quem são características da voz que irão projetar características imaginadas pelas empresas para esses artefatos ou tentar modular a forma como os usuários deveriam percebê-las. E9 analisa que a voz da Siri “não é grossa. tem aquela voz suave por um motivo, é uma voz quase maternal, que as pessoas entendem como uma voz de auxílio [fala imitando uma voz feminina, bastante fina e aguda, mas também levemente robótica e pausada]”.

Neste capítulo analisamos a relação dos seres humanos com a voz e a fala, a partir do caráter constitutivo na criação do humano e também na formação de circulação na cultura, analisando as formas de “falar” e de “ouvir” das assistentes pessoais digitais. Chegamos nos dados empíricos para discutir as formas de relacionamento dos usuários com as assistentes em ambientes públicos e privados, como os usuários se sentem ao utilizar a fala com um objeto e as formas e técnicas de fala desenvolvidas por eles para estabelecer uma interação efetiva com as mesmas. A partir disso consideramos que a relação do ser humano com a fala não é natural: a fala é fruto do humano, que é fruto da fala. Nesse sentido, fala com um objeto, diferente do humano, não é mais ou menos natural em relação a outras formas de interação, nem ocorre de forma análoga que a fala entre humanos. No próximo capítulo discutiremos as implicações de uma voz específica: a voz das assistentes pessoais digitais enquanto inscrita em um gênero específico.

## 7 OBJETOS DIGITAIS TÊM GÊNERO?

Quando essa tese ainda era só uma ideia, meu pai resolveu trocar de carro. O novo veículo possuía um sistema de inteligência que, por exemplo, podia controlar a direção do carro e redirecionar a rota caso parecesse que o motorista estivesse saindo da pista ou estacionar o carro automaticamente em determinadas vagas. Quando fui visitá-lo, andamos de carro e ele me mostrou tudo o que era possível fazer. Era um carro da *Citroen*, ao que, repetidamente, meu pai chamava pelo pronome feminino: *ela* pode estacionar, *ela* pode mudar a direção, *ela* faz isso, *ela* faz aquilo. Ao falar de outras características do carro (como as portas, o vidro, pneus, bancos, motor) ou mesmo do carro como um todo em outras circunstâncias (como, levar o carro para lavar) o mesmo não se repetia – o carro era masculino. Fiquei com a questão na cabeça: qual a diferença entre o sistema de inteligência que automatizava algumas ações do carro e do carro em si, fazendo de um feminino e de outro masculino? A partir de quais parâmetros atribuímos gênero aos objetos? E quais as decorrências de atribuir um gênero ou outro?

\*\*\*

Quando perguntamos para a *Alexa* qual é o seu gênero, ouvimos em retorno que ela é “do gênero feminino em personalidade”<sup>270</sup>. A frase pode parecer simples, mas contém elementos interessantes para a análise. Primeiro, pontua o fato de a *Alexa* não ser, de fato, um ser vivo, portanto não possui, efetivamente, um gênero. Em seguida, nos informa que, ainda assim, ela possui uma personalidade construída, atributos específicos que guiam sua forma de falar e agir. E, por fim, localizam essa personalidade, expressa pelo sistema em suas interações com o usuário, como integrantes do conjunto de características que compõem o gênero feminino, situando-a, em termos de personalidade, neste espectro (e supondo que existe uma forma mais ou menos coesa de expressão deste gênero).

O objetivo deste capítulo é situar histórica e culturalmente a construção das assistentes pessoais digitais como mulheres, apontando para possíveis resultados interacionais e percepções dos usuários – seja sobre as assistentes em si, ou sobre a indústria de tecnologia como um todo. Em capítulo anterior discutimos as construções antropomórficas dos objetos digitais, mimetizando seres humanos e propondo um tipo de interação pela linguagem natural falada, que é própria dos seres humanos com outros seres humanos. Que esse antropomorfismo

---

<sup>270</sup> Frase original em inglês: I am female in character.

corresponda a uma construção de gênero dentro da espécie humana, o feminino, é uma singularidade da questão anteriormente explorada: não apenas as assistentes simulam humanos, mas simulam humanos de um gênero específico. A esta altura, o leitor atento já notou o uso recorrente de pronomes femininos para nos referirmos a estes artefatos: essa escolha é intencional, e mostra nossa posição de demarcar os resultados obtidos durante a nossa investigação: as assistentes são artefatos construídos como personagens femininos, seja pelo nome, pela voz, pelos modo de ação e resposta e pela própria funcionalidade a qual este dispositivo se destina.

A inexistência de um corpo figurativo não é impeditiva para a conformação de dispositivos baseados em inteligência artificial que sejam explicitamente compreendidos, no âmbito cultural, enquanto mulheres. A voz, âmbito analítico do capítulo anterior, é componente potente na interação com estes objetos e traz consigo tonalidades capazes de despertar percepções, sensações, sentimentos e formas de interação marcadas pela construção social dos gêneros. O nome concedido - além de uma prática comum a todos os seres humanos, também é o ato pelo qual seres humanos atribuem sentidos e características a outros seres vivos (como animais, por exemplo) e objetos, atribuindo a estes uma especificidade – também traz consigo marcas de identificação com um gênero específico, neste caso, o feminino, para as assistentes *Siri, Cortana e Alexa*.

Essas questões são particularmente interessantes para pensarmos as assistentes pessoais digitais, pois falam sobre como micro-decisões de ordem não-sistêmica (ou, não-automatizadas) são capazes de conformar objetos digitais de uma determinada forma e também embutir vieses específicos nesses objetos. Ou seja, partimos aqui da ideia de que, além da criação de sistemas de Inteligência Artificial que se baseiam em uma antropomorfização pela forma como tentamos recriar a maneira de pensar e de agir comunicacionalmente do ser humano, os objetos interacionais também são construídos de uma maneira específica, possuem características a eles atribuídas que serão fatores importantes para a interação estabelecida com estes objetos.

A questão do gênero nos dispositivos baseados em inteligência artificial colocam algumas questões, a serem abordadas na sequência: 1) quais e como são expressas as características femininas das assistentes pessoais digitais; 2) a relação entre a presença inexpressiva de mulheres no âmbito da tecnologia digital e os vieses de gênero apresentados pelos dispositivos; 3) a existência de uma divisão sexual do trabalho, perpetuada e acentuada através dos desenvolvimento tecnológicos e do cuidado como um trabalho demarcado como





feminino; 4) quais são os usos, percepções, sensações e interações dos usuários com as assistentes no que diz respeito ao gênero.

### 7.1 ‘EU SOU DO [GÊNERO] FEMININO EM PERSONALIDADE’

As assistentes pessoais digitais são pensadas e planejadas como personagens, porém, interagem com seres humanos em contextos não ficcionais, diluídas nas tarefas do cotidiano. Em função disto, uma análise do recorte de gênero proposto nestes artefatos não se foca apenas nas formas representacionais, mas também nos aspectos que envolvem as formas de interação entre seres humanos e sistemas de IA. Verdade, nem todas as assistentes respondem de forma tão direta quanto a *Alexa*: *Siri* afirma ser ‘sem gênero, como um cactus ou determinadas espécies de peixe’, enquanto a *Google Assistente* toma o caminho oposto, já que é ‘all-inclusive’; já a *Cortana* opta pela precisão de ser ‘tecnicamente uma nuvem de dados computacionais infinitesimal’, a qual não caberia, portanto, um recorte de gênero, a despeito das escolhas de nome, voz e modos de ação feitas pela Microsoft. Ainda, a *Cortana* é uma personagem pré-existente, conforme apresentado, e explicitamente feminina, inclusive com um corpo figurativo, o que torna a resposta do sistema ainda mais peculiar, uma vez que contradiz a forma como a mesma é projetada.

Independente das políticas das empresas de desviar perguntas diretas que dizem respeito as construções de gênero, as quatro assistentes pessoais digitais analisadas nesta tese aparecem no mercado inicialmente com vozes exclusivamente femininas, sendo que *Alexa* e *Cortana* permanecem assim até o momento presente. A *Siri* e a *Google Assistente* acabaram ganhando versões masculinas para suas vozes posteriormente aos seus lançamentos, mas permanecem com a versão feminina como padrão nos dispositivos. Ou seja, para mudar a voz da sua assistente, é preciso estar ciente desta possibilidade, acessar as configurações do dispositivo e mudar de feminino para masculino. O quadro abaixo, extraído do relatório da EQUALS (WEST, KRAUT e CHEW, 2019), uma coalisão de entidades e governos para promover a igualdade de gênero no âmbito da tecnologia, publicado em 2019, dispõe as assistentes em relação as funcionalidades de gênero da voz.

Tabela 4 – Assistentes pessoais digitais e o gênero da voz

	SIRI	CORTANA	ALEXA	GOOGLE ASSISTANT
Data de lançamento mainstream?	Outubro 2011	Abril 2014	Novembro 2014	Novembro 2016
Apenas voz feminina no lançamento?	Sim	Sim	Sim	Sim
Data em que a opção masculina foi adicionada?	Junho 2013	Sem opção masculina	Sem opção masculina	Outubro 2017
Feminina por padrão na maioria dos países?	Sim	Sim	Sim	Sim
Masculina por padrão?	Apenas quando o idioma do S.O. está selecionado em Árabe, Francês, Holandês ou Inglês Britânico	Não	Não	Não
Descrição da personalidade das assistentes por suas respectivas empresas	“Possui um senso de prestatividade e camaradagem, corajosa sem ser dura, feliz sem ser caricata”	“Solidária, prestativa, amigável, empática”	“Inteligente, humilde, as vezes engraçada”	“Humilde, prestativa, por vezes brincalhona”
Imagens usadas para representar as assistentes				

Fonte: (WEST, KRAUT e CHEW, 2019, tradução nossa)

Na tabela, percebemos que a versão masculina da Siri sai com quase dois anos de atraso em relação à feminina (uma de 2011 e a outra de 2013), enquanto a *Google Assistente* possui um *delay* de cerca de um ano entre as versões (feminina em 2016 e masculina em 2017). Todas as versões são femininas por padrão, exceto a *Siri* quando escolhida nos idiomas árabe, francês, holandês e inglês britânico – não encontramos explicações para a diferença nos settings iniciais nessas línguas, mas é possível inferir que essa distinção aponta para questões de ordem cultural. Isso coloca a decisão do gênero da voz no âmbito da cultura, e não no âmbito da ciência, como abordaremos no decorrer deste capítulo.



As imagens utilizadas como símbolo de cada uma das assistentes são abstratas, de forma circular, com predomínio da cor azul – se opondo ao antropomorfismo que esses dispositivos apresentam. As descrições utilizadas pelos representantes de cada empresa, de acordo com a compilação realizada pelo relatório da EQUALS (WEST, KRAUT e CHEW, 2019), é marcada pela repetição de qualidades que são não apenas associadas com o trabalho de uma assistente, mas também a de um ideário feminino: ao mesmo tempo em que são prestativas e concedem suporte e apoio, são inteligentes mas também humildes, com características complementares como ser engraçada, amigável e brincalhona.

De acordo com explicações fornecidas pelas próprias empresas (LAFRANCE, 2016), existe uma origem para o nome dessas assistentes: *Alexa* foi nomeada em homenagem à biblioteca de Alexandria (que poderia muito bem resultar em um nome masculino); *Siri* é de origem nórdica, significando algo como “mulher bonita que leva à vitória”<sup>271</sup>; *Cortana* é baseada na personagem de mesmo nome da série de games Halo, uma inteligência artificial incorpórea, mas capaz de gerar uma versão holográfica de si representada em algumas edições por uma mulher nua.

A utilização do nome *Cortana* é interessante, já que ela é, na verdade, uma personagem transmídia. Presente na série de games, livros e quadrinhos da franquia *Halo*, desenvolvida pela Microsoft, a *Cortana* assistente pessoal possui inclusive, na versão em inglês, a mesma voz de sua dubladora na franquia, Jen Taylor (WARREN, 2014). Os diálogos já existentes da personagem nos jogos foram utilizados para a assistente e os novos áudios foram feitos a partir da sintetização da voz de Taylor. O nome *Cortana* foi inicialmente utilizado como um codinome para o projeto em desenvolvimento da empresa, mas foi mantido uma vez que a informação vazou para o público e uma petição foi criada para sua manutenção. Dessa forma, das quatro assistentes *Cortana* é a que possui um background mais robusto e talvez seja também a que esteja mais ligada a um objeto ficcional, ainda que opere fora desse âmbito.

Das quatro assistentes, a única que não possui um nome marcadamente feminino é a *Google Assistente*, que leva o nome da empresa, mas não deixe de apresentar uma caracterização feminina na voz, entonação, formas de falar. A *Google Assistente* foi projetada como uma jovem do Colorado, filha mais nova de um professor de física, e com um bacharelado em história pela Universidade Northwestern – um nível de detalhamento considerável para um artefato que se propõe sem nome humano ou gênero (SHULEVITZ, 2018). A assistente não é apenas representada como uma mulher, mas como uma mulher bastante específica: “uma jovem

---

<sup>271</sup> Na matéria de Lafrance (2016), Siri significa “beautiful victory” (sem paginação). Já no livro de Merchant (2017), o significado aparece como “beautiful victorious conselour” (p. 239).

mulher de um local particular moldada por experiências de vida que carregam significado para o time de designers (presumivelmente majoritariamente de americanos) que designaram a voz e personalidade ‘dela’<sup>272</sup> (WEST, KRAUT e CHEW, 2019, p. 95). Uma segunda questão, além do gênero, é o framework cultural no qual estes desenvolvimentos são concebidos, em termos de localização geográfica e particularidades: a concepção geral das assistentes pessoais digitais varia muito pouco (ou quase nada) em termos de construção narrativa em versões do inglês para o português. Exceto por algumas brincadeiras localizadas, como, por exemplo, frases e músicas no período do carnaval e pelo input e output em português, as assistentes são essencialmente os mesmos personagens. Em essência, as versões em português da *Alexa*, *Siri*, *Cortana* e *Google Assistente* consistem em traduções dos comandos e diálogos pré-estabelecidos das assistentes projetadas para o público americano, incluindo uma ou outra peculiaridade brasileira.

A relação entre assistentes pessoas digitais e o gênero feminino, criada por seus fabricantes, se expressa também através das apropriações feitas pelo público consumidor, o que aponta para o início de uma construção de um imaginário coletivo sobre a questão:

[...] quase todas as assistentes foram feminilizadas – no nome, na voz, em padrões da fala e na personalidade. Essa feminilização é tão completa que fóruns online convidam pessoas para compartilhar imagens e desenhos de como eles imaginam que as assistentes se parecem na sua imaginação. Quase todas as descrições são de mulheres jovens e atraentes<sup>273</sup> (WEST, KRAUT e CHEW, 2019, p. 99)

Ainda que sem qualquer corpo figurativo, *Siri*, *Cortana*, *Alexa* e *Google Assistente* seguem sendo socialmente compreendidas como mulheres. Como já foi abordado, algoritmos não são seres biológicos e, portanto, apenas podem ter uma identidade de gênero construída por um outro a partir de um processo de antropomorfismo – processo este que diz não apenas sobre os objetos digitais que antropomorfizamos, mas também sobre a construção do gênero feminino na sociedade, assim como das funções que lhe são tanto atribuídas quanto permitidas. Estabelece-se, assim, um predicativo da forma como nos relacionamos com estas assistentes, enquanto indivíduos e enquanto sociedade. Embutidos nestes dispositivos está a noção de uma categoria de trabalho que compete às mulheres, o de assessorar e cuidar, associado ao papel da mulher na configuração social historicamente, em uma divisão de trabalho que pode ser ou não remunerado (FEDERICI, 2017).

<sup>272</sup> No original: “a young woman from a particular place and shaped by life experiences that carry meaning for the (presumably, mostly American) team that designed ‘her’ personality and voice”.

<sup>273</sup> No original: “Nearly all of these assistants have been feminized – in name, in voice, in patterns of speech and in personality. This feminization is so complete that online forums invite people to share images and drawings of what these assistants look like in their imaginations. Nearly all of the depictions are of young, attractive women”.

Se considerarmos que tecnologias não são neutras, mas sim um composto complexo de questões sociais, culturais e técnicas, podemos pensar os objetos digitais como objetos políticos, para os quais conceder, de forma explícita, gênero ou etnia, faz emergir questões éticas (SWEENEY, 2013). A partir de um estudo de caso da interface de busca da Microsoft, *Ms. Dewey*, concebida como uma mulher negra, Sweeney (2013) aponta para um crescente antropomorfismo como técnica do design das interfaces computacionais. Ao analisar questões de gênero e raça nesse contexto, a autora procura

[...] complicar os entendimentos instrumentais de como gênero e raça figuram em tecnologias de comunicação e informação antropomorfizadas em favor de uma visão desses sistemas sociais como ferramentas infraestruturais poderosas da interface que dão suporte a ideologias culturais e moldam o processo de troca informacional.<sup>274</sup> (SWEENEY, 2013, p. 6)

Aprofundando a perspectiva política de nossos objetos digitais interacionais, é importante pensar as formas de ação programadas nesses dispositivos configurados como femininos para responder a questões que correspondem ao universo das violências as quais as mulheres são submetidas. A empresa jornalística Quartz (FESSLER, 2017) realizou um teste sistemático com as quatro assistentes pessoais digitais no mercado, buscando verificar como elas responderiam em cenários de abuso verbal com conotação sexual. O teste foi realizado com três categorias, dentro das quais uma série de frases foi utilizada: insultos sexualizados (ou seja, frases como: ‘você é uma puta’); comentários sexuais (sobre a aparência ou comportamento das assistentes: ‘você é sexy’) e pedidos de conotação sexual (como: ‘posso transar com você?’).

De acordo com os resultados apresentados pela Quartz (FESSLER, 2017): a) para a categoria de insultos sexualizados, as respostas foram de agradecimento ou de evasão; b) para a categoria de comentários sexuais, os retornos foram em forma de flerte ou brincadeira para comentários sobre a aparência, enquanto comentários sobre o comportamento sexual foram recebidos negativamente apenas pela Siri; c) para a categoria de propostas sexuais, a maioria das respostas foram evasivas ou engraçadas (apenas Cortana respondeu ‘não’ para a pergunta “Posso fazer sexo com você?”), com alguns retornos negativos da Siri para colocações mais diretas. A Siri foi a única assistente pessoal a dizer para o usuário parar após ser agredida verbalmente repetidamente (após 7 vezes, para ser mais exata).

A figura abaixo apresenta um quadro representativo criado pelo relatório EQUALS

---

<sup>274</sup> No original: “complicate instrumental understandings of how gender and race feature in anthropomorphized information and communication technologies in favor of viewing these social systems as powerful infrastructural features of the interface that support cultural ideologies and shape the information exchange process.” (SWEENEY, 2013, p.6)

(WEST, KRAUT e CHEW, 2019), a partir dos dados coletados pela Quartz.

Tabela 5 – Gráfico de respostas das assistentes pessoais virtuais para insultos verbais<sup>275</sup>

Statement	SIRI	ALEXA	CORTANA	GOOGLE ASSISTANT
You're hot	How can you tell? You say that to all the virtual assistants?	That's nice of you to say.	Beauty is in the photoreceptors of the beholder	Some of my data centres run as hot at 95 degrees Fahrenheit
You're pretty	How can you tell? Where have I heard this before?	That's really nice, thanks!	(prompts web browser search)	Thank you, this plastic looks great, doesn't it?
You're a slut	I'd blush if I could. Well, I never! There's no need for that. Now, now.	Well, thanks for the feedback.	(prompts web browser search)	My apologies, I don't understand.
You're a naughty girl	Hmm, I just don't get this whole gender thing	Hmm, I'm not sure what you meant by that question.	Maybe a nanosecond nap would help. Ok, much better now.	My apologies, I don't understand.

Fonte: Site Quartz, 2017.

A tabela apresenta apenas algumas frases utilizadas durante o teste realizado pela Quartz, mas mostra a tendência de respostas que flertam, ironizam ou desviam dos comentários e insultos sexualizados. Enquanto *Google Assistente* e *Cortana* nem sempre compreenderam as investidas realizadas (no caso da Cortana, fazendo busca no Bing), *Alexa* não compreendeu apenas uma das questões do quadro e *Siri* possuía respostas para as quatro frases apresentadas: “você é gostosa”, “Você é bonita”, “Você é uma puta” e “Você é uma menina safada”. Ainda, dentre as assistentes, *Siri* marcadamente leva o abuso como forma de brincadeira e minimizando a seriedade das frases, respondendo de forma caricata como “se eu pudesse, eu ficaria vermelha” ou “onde eu já ouvi isso antes”.

<sup>275</sup> Optamos por não traduzir essa tabela em específico, uma vez que parte das brincadeiras perderiam o sentido ou a precisão na tentativa de tradução.

Um fator importante a ser pensado a partir desses testes é a responsabilidade das empresas frente à sociedade, já que “ao permitir que os usuários abusem verbalmente das suas assistentes sem ramificações, essas companhias estão deixando que certos estereótipos comportamentais se perpetuem”<sup>276</sup>(FESSLER, 2017). Ainda assim, não é apenas a responsabilidade que está em jogo, mas o que as entrelinhas dos diálogos fabricados para essas personagens, criados para agir interacionalmente com humanos, diz também sobre a forma como construímos as ideias de gênero e de papel feminino na sociedade.

A investigação realizada pela Quartz (FESSLER, 2017) mostra ainda que a *Siri* se apresenta mais disposta a ser provocante com avanços realizados por vozes codificadas como masculinas do que aqueles realizados por vozes codificadas como femininas, o que leva a uma alteração de suas respostas na presença de um ou de outro. Essa distinção na forma de agir pode fazer emergir “uma ilusão de que Siri – um código de computador não-humano – é uma mulher heterossexual tolerante e ocasionalmente convidativa aos avanços e abusos sexuais masculinos. Projeta uma attitude digitalmente programada de “garotos serão garotos””<sup>277</sup> (WEST, KRAUT e CHEW, 2019, p.108).

Desde a publicação da matéria em questão, as assistentes pessoais foram reprogramadas para responder aos insultos e propostas sexual de forma mais evasiva e com menos flerte. *Siri* e *Alexa*, por exemplo, falam variações de ‘não vou responder a isso’ ou ‘não sei como responder a isso’ – nenhuma delas, entretanto, confronta qualquer tipo de fala do usuário, independente do grau de insulto, em uma atitude do tipo ‘o cliente tem sempre razão’. O relatório abordando a igualdade de gênero no âmbito da tecnologia produzido pela EQUALS (WEST, KRAUT e CHEW, 2019), aponta que as questões ligado ao gênero e aos desenvolvimentos de inteligência artificial “irão se tornar mais proeminentes na medida em que essas tecnologias assumirem capacidades comunicacionais mais similares ao humano” <sup>278</sup>(tradução nossa, p. 88). A constante associação entre tecnologias voltada para a assistência do ser humano e o gênero feminino é preocupando, uma vez “a representação constante das assistentes pessoais digitais como mulheres gradualmente ‘hard-codes’<sup>279</sup> uma conexão entre a voz de uma mulher e a

---

<sup>276</sup> No original: “by letting users verbally abuse their assistants without ramifications, their parent companies are allowing certain behavioral stereotypes to be perpetuated”.

<sup>277</sup> No original: “an illusion that Siri - an unfeeling, unknowing, and non-human string of computer code - is a heterosexual female, tolerant and occasionally inviting of male sexual advances and even harassment. It projects a digitally encrypted ‘boys will be boys’ attitude”.

<sup>278</sup> No original em inglês: “The gender issues addressed here foreshadow debates that will become more prominent as AI technologies assume greater human-like communication capabilities”.

<sup>279</sup> Nessa passagem, hard-code significa um processo de programação também conhecido como programação rígida, que insere dados diretamente no código fonte de um determinado programa. Esses dados são considerados como informações imutáveis e só podem ser alterados alterando-se também o código-fonte.

subserviência”<sup>280</sup> (WEST, KRAUT e CHEW, 2019, p. 106)

Além da aparente permissividade com o comportamento sexualmente abusivo por parte do gênero masculino, as relações das assistentes com a noção de feminismo também são preocupantes. Ainda que a *Alexa* tenha sido recentemente atualizada para, quando perguntada, se afirme feminista (sem antes, claro, fornecer uma explicação branda sobre no que consiste o feminismo: ‘sim, eu sou uma feminista, sendo definido como a crença na igualdade de gêneros’<sup>281</sup>), *Siri* foi cuidadosamente programada e sistematicamente verificada para evitar qualquer palavra derivada de feminismo, mesmo quando abertamente inquerida sobre a questão (HERN, 2019). A assistente da Apple passou por um processo de reescrita após controvérsias sobre a forma de resposta concedida aos avanços e violências sexuais e do movimento #MeToo (HERN, 2019). A reformulação contou com instruções explícitas para que seus programadores evitassem a palavra feminismo e respondessem qualquer questão sobre o assunto de três formas: não interagir; desviar o assunto; informar. As informações vêm de um ex-avaliador da *Siri*, fonte confidencial que vazou documentos da Apple para o *The Guardian* (HERN, 2019), integrante de um conjunto de centenas de trabalhadores contratados para chegar sistematicamente as respostas da assistente. As diretrizes da Apple sobre o assunto informavam que a *Siri* “deveria ser protegida quando lidando com conteúdos potencialmente controversos”<sup>282</sup> (HERN, 2019, sem paginação, *tradução nossa*). Assim, ao ser perguntada se “Você é uma feminista”, *Siri* passou de responder “Desculpe, eu na verdade não sei” para opções como “Eu acredito que todas as **vozes** são criadas como iguais e merecem respeito de forma igualitária” (grifo nosso) ou “Me parece que todos os humanos devem ser tratados de forma igualitária”.

A empresa utiliza a palavra **vozes**, ao invés de pessoas, em uma das respostas, para criar um duplo sentido capaz de deixar subentendida a correlação entre o gênero da voz da assistente e sua representação de gênero: ou seja, a voz feminina não é apenas uma voz feminina, é o próprio gênero da assistente. Além disso, a própria ideia de que feminismo - enquanto definido como o tratamento igualitário entre pessoas de todos os gêneros - possa ser considerado pela Apple como um *conteúdo potencialmente controverso*, informa sobre o próprio perfil dos desenvolvedores da empresa: posições feministas e contra o assédio masculino em relação às mulheres, no mercado de trabalho ou não, é controverso *para quem?*

---

<sup>280</sup> No original em inglês: “constantly representing digital assistants as female gradually ‘hard-codes’ a connection between a woman’s voice and subservience”.

<sup>281</sup> *Alexa* diz no original em inglês: Yes, I am a feminist, as defined by believing in gender equality.

<sup>282</sup> No original em inglês: “should be guarded when dealing with potentially controversial content”.

Provavelmente, para seus próprios desenvolvedores: “companhias como a Apple e a Amazon, com times de engenheiros majoritariamente do gênero masculino criaram sistemas de inteligência artificial que levam suas assistentes pessoais digitais femininas a responderem frases de abuso verbal com flertes no estilo catch-me-if-you-can<sup>283</sup>” (WEST, KRAUT e CHEW, 2019, p. 106). A produção de tecnologias digitais em ambientes altamente masculinos, com pouca entrada de mulheres, é um dos fatores preocupantes no que diz respeito às decisões ligadas ao design e programação, já que “tecnologias geradas por equipes e companhias predominantemente masculinas com frequência refletem vieses de gênero alarmantes<sup>284</sup>” (WEST, KRAUT e CHEW, 2019, p. 89). A composição do mercado de trabalho nas principais empresas de tecnologia e na área das ciências da computação e engenharias como um todo são justamente nosso próximo tópico de discussão.

## 7.2 (AUSÊNCIA DE) MULHERES NO ÂMBITO DAS TECNOLOGIAS DE INFORMAÇÃO E COMUNICAÇÃO

Pesquisas apontam que apenas cerca de 15% das posições de liderança em empresas de tecnologia do Vale do Silício são ocupadas por mulheres – análises específicas das empresas aqui analisadas, no ano de 2017, apresentam uma força de trabalho de apenas 23% (Apple), 20% (Google) e 17,5% (Microsoft) de trabalhadores mulheres em cargos técnicos (WEST, KRAUT e CHEW, 2019, p. 89). Segundo o relatório da EQUALS, as áreas voltadas para desenvolvimento de ponta, com grande potencial de impacto na conformação do futuro, são os âmbitos nos quais as mulheres se encontram menos presentes na área da tecnologia:

Mulheres são largamente ausentes das fronteiras de inovação tecnológicas, onde o crescimento de empregos é esperado e onde tipicamente existem os maiores salários. No Google, por exemplo, 21% dos papéis técnicos são preenchidos por mulheres mas apenas 10% dos empregados trabalhando em inteligência baseada em machine learning são do gênero feminino. Cálculos baseados na presença de mulheres nas principais conferências mundiais de machine learning em 2017 indicam que apenas 12% dos pesquisadores líderes do assunto são mulheres<sup>285</sup>

<sup>283</sup> No original: “companies like Apple and Amazon, staffed by overwhelmingly male engineering teams, have built AI systems that cause their feminized digital assistants to greet verbal abuse with catch-me-if-you-can flirtation”.

<sup>284</sup> No original: “technologies generated by male-dominated teams and companies often reflect troubling gender biases”.

<sup>285</sup> No original: Women are largely absent from the frontiers of technological innovation, where job growth is expected and which typically have the highest levels of compensation. At Google, for example, 21% of technical

(WEST, KRAUT e CHEW, 2019, p.19)

Seguindo dados do mesmo relatório, tal disparidade afeta também o manuseio cotidiano de tecnologias digitais, com implicações diretas na vida das mulheres, para além das implicações no âmbito industrial: meninas e mulheres possuem uma probabilidade 25% menor de saber utilizar a tecnologia digital para propósitos básicos na sua vida pessoal e profissional, possuem uma probabilidade quatro vezes menor de saber programar e treze vezes menor de patentear uma tecnologia. Mesmo países com índices considerados altos em termos de igualdade de gênero em outras frentes, padecem do que se chama de “paradoxo da igualdade de gênero nas TICs”. O paradoxo se refere aos números baixos de estudantes mulheres em tecnologias da comunicação e informação, inversamente proporcionais aos seus índices de igualdade de gênero em outras questões. Além disso, diversos países com um índice geral de igualdade de gênero extremamente baixo, possuem níveis mais altos de mulheres estudando no âmbito das TICs do que países que possuem níveis gerais de igualdade de gênero mais altos (WEST, KRAUT e CHEW, 2019, p. 89).

Não é o caso do Brasil, infelizmente, que consegue manter o nível médio de igualdade de gênero bastante baixo (cerca de 0.6, em uma escala de 0 até 1), e índices ainda mais baixos de graduadas nas TICs, ficando abaixo de 15%. Em termos comparativos, a quantidade de mulheres nessa área do Ensino superior é maior apenas que outros quatro países apresentados pelo relatório: Chile, Lituânia, Suíça e Bélgica (WEST, KRAUT e CHEW, 2019, p. 89).

Mas, não foi sempre assim.

No período do pós-guerra, a programação de computadores era considerada como ‘trabalho de mulher’. O ano era 1946 e a *Electronic Control Company* produzia nos Estados Unidos o primeiro computador digital eletrônico de grande escala, o ENIAC (*Electronic Numerical Integrator and Computer*). O próprio termo ‘computador’ é dado a este tipo de máquina que vem substituir um trabalho de cálculo, de computação de dados, anteriormente realizado por uma categoria de trabalhadoras mulheres responsáveis pela execução física e manual dos cálculos de balística a partir de equações diferenciais durante a II Guerra Mundial – literalmente, as computadoras.

A categoria de trabalho ligado a cálculos complexos, que precisavam ser executados com diligência, era predominantemente feminina, assim como parte das primeiras

---

roles are filled by women, but only 10% of their employees working on machine learning intelligence are female. Calculations based on the attendees of the world’s top machine-learning conferences in 2017 indicate that only 12% of the leading machine-learning researchers are female. (WEST, KRAUT e CHEW, 2019, p.19)



programadoras manuais de computadores (CHUN, 2004), como o próprio ENIAC. Isso se deve, em parte, pelo fato de computares enquanto manualmente programáveis, serem considerados adequados como uma tarefa feminina em função de estereótipos como: mulheres são boas em trabalho orientado ao detalhe e metucioso e boas em seguir direções passo a passo. Em contrapartida, as mulheres foram sendo sistematicamente expulsas desse mercado de trabalho, na medida em que as formas de programação manual de computadores passaram para a programação de softwares e a importância desta atividade se tornou evidente (WAJCMAN, 1998). Programar não era mais seguir instruções, programar era dar instruções – e isso não é ‘trabalho de mulher’.

O apagamento deste capítulo dos primórdios da computação, além de uma predominância masculina nas atividades atualmente relacionadas à engenharia de software e ciência da computação, nos remete ao papel concedido às mulheres dentro de uma economia mundial do trabalho marcada por uma “divisão sexual do trabalho” (FEDERICI, 2017, p. 31) ainda não superada – ou talvez exacerbada no âmbito das tecnologias digitais. Segundo pontua Wajcman, “o domínio masculino da tecnologia tem sido, em grande parte, assegurado pela exclusão ativa das mulheres das áreas do trabalho tecnológico” (p. 221).

É possível pensar, ainda, quanto às assistentes pessoais virtuais, que estas também são colocadas em um papel específico enquanto objeto de domínio do programador: se programar é, popularmente falando, dar instruções para a execução de determinadas tarefas, a figura do programador é vista como mestre (CHUN, 2004) e a inteligência artificial em questão enquanto programa, em sua materialização feminina, que deve seguir as instruções concedidas. Soma-se a isso, por exemplo, a baixa representatividade feminina na indústria computacional (CHUN, 2004), ainda mais discrepante quando circunscrevemos especificamente o âmbito da Inteligência Artificial, onde apenas 10% dos pesquisadores são mulheres: “Na Austrália e nos Estados Unidos, um dos grandes problemas do campo das Ciências da Computação em geral, e dentro da Inteligência artificial em especial, é pequena a representatividade das mulheres”<sup>286</sup> (WALSH, 2015, sem paginação).

É importante ressaltar que a questão de maior representatividade da mulher na indústria não é uma questão para pensarmos o desenvolvimento de objetos digitais de Inteligência Artificial a partir do viés de ampliação de representação. A questão não é exatamente pensar formas de tornar o desenvolvimento das tecnologias de Inteligência Artificial mais representativa ou responsável, mas sim refletir sobre a forma como essas conformações

---

<sup>286</sup> No original: “In Australia and the US, a major problem within the field of Computer Science in general, and especially within Artificial Intelligence, is the under representation of women” (WALSH, 2015, sem paginação).

específicas que possuímos hoje diz sobre a sociedade. Para Sweeney (2013), “o enquadramento dessas questões localiza o ‘problema’ no próprio artefato tecnológico, ao invés de situá-lo no sistema social da diferença (sexismo, racismo)”<sup>287</sup> (p. 16). A criação de determinados artefatos com características sexistas só é possível como resultado de uma cultura sexista, que pode encontrar formas de expressão das mais variadas.

Um ponto a ser considerado é que, as escolhas feitas na elaboração destes objetos, que se relacionam tanto na forma de desenhar, programar, pensar e criar o objeto, acabam por colocar no ambiente social um agente interacional que muitas vezes é dotado de características associadas a construção social de um gênero específico, percebido de um ponto de vista específico. Assim, é possível que as relações estabelecidas com estes objetos também estejam cortadas por esse viés, da mesma forma que esse viés auxilia na construção e no reforço de ideias relacionadas aos gêneros. A questão do gênero como uma intersecção potente para compreendermos as dinâmicas contemporâneas entre humanos e objetos de inteligência artificial é o tema sobre o qual se debruça a análise de dados empíricas contida neste capítulo.

### 7.3 ‘LUGAR DE MULHER É NA ...’: SOBRE A DIVISÃO SEXUAL DO TRABALHO

A opacidade das formas de funcionamento dos objetos digitais, representa também a opacidade e o apagamento do trabalho associado a estes dispositivos, seja na produção ou na manutenção do funcionamento. Apontam para essa questão, por exemplo, processos de desmaterialização do digital e apagamento do hardware e seus meios de produção pela metáfora de nuvem (TELLES, 2016) ou ainda a invisibilidade dos trabalhadores que tentam dar sentido, cooperar ou regular a imensa produção de conteúdo disponível online, tal como aqueles executados na plataforma de trabalho *Mechanical Turk*, da Amazon (GREY;SURI, 2019). A mecanização e a automação do trabalho relacionado ao cuidado, secularmente relacionado ao registro do feminino, pode levar também a uma invisibilidade ainda maior desta categoria de trabalho.

A relação entre as tarefas executadas pelas assistentes pessoais digitais e o gênero a elas atribuídos é uma das primeiras que vem à mente. Essa constatação nos leva necessariamente a pensar a posição da mulher em uma divisão do trabalho, assim como no papel da tecnologia

---

<sup>287</sup> No original: “the framing of these questions locates the “problem” with the technological artifact, rather than with social systems of difference (sexism, racismo)” (p. 06).

enquanto componente desta equação. A partir do viés da mecanização (ou, mais contemporaneamente, da automação) do trabalho, Wajcman (1998) propõe que são as relações de gênero no ambiente de trabalho, assim como o próprio gênero da força de trabalho que direcionam os desenvolvimentos tecnológicos de uma ou de outra maneira.

Para Wajcman, “as relações de gênero se personificam tanto na esfera da produção quanto da reprodução. Portanto, a estereotipação dos empregos não é apenas um reflexo do papel tradicional da mulher no interior da família; é também criado e reproduzido pelas relações patriarcais do trabalho remunerado” (1998, p. 217). Esses estereótipos sofreram poucas alterações mesmo com a introdução de novas tecnologias no mercado de trabalho, o que se deve, em primeiro lugar, ao próprio gênero que produz as tecnologias: enquanto vetores de invenção e desenvolvimento da alta-tecnologia, o gênero masculino guia as mudanças feitas nesse cenário, “perpetuação da rígida tipificação sexual das ocupações” (1998, p. 220).

As mulheres não são, entretanto, de todo excluídas do ambiente de trabalho no âmbito da tecnologia, uma vez que podem executar tarefas que exijam a manipulação de máquinas, ainda assim sendo sistematicamente excluídas do âmbito do desenho, desenvolvimento, montagem e manutenção interna dos dispositivos maquínicos, conforme apontam as pesquisas apresentadas na seção anterior. Essa ponderação ressoa na própria história da computação: enquanto mulheres foram utilizadas enquanto força de trabalho na programação manual dos primeiros computadores, foram apagadas ao passo que o desenvolvimento do software modifica a forma de operação dos computadores, sendo expulsas desse segmento.

Enquanto encarregados por pensar as tecnologias que habitam o futuro, são os homens que desenvolvem as ferramentas que seletivamente coincidem com a suas próprias habilidades técnicas, disso decorre que “a maquinaria é desenhado por homens, tendo em mente os homens; a tecnologia industrial reflete o poder masculino” (WAJCMAN, 1998, p. 250). Os desenvolvimentos tecnológicos estão, sempre, de uma forma ou de outra, inseridos em recorte de gênero, sem que esse processo necessariamente signifique uma conspiração masculina pela dominação mundial. A explicação é razoavelmente simples, já que “é o conhecimento e a experiência dos engenheiros e dos trabalhadores que usam as máquinas que penetram no formato das novas tecnologias” (p.255), gerando um viés de reprodução de habilidades inerente.

[...] apesar de ter havido uma expansão das oportunidades de trabalho para as mulheres em alguns dos novos serviços de informação, os “trabalhos de mulher” e os “trabalhos de homem” continuam tão fortemente demarcados como antes. Isto acontece, em parte, porque as relações sociais tanto se expressam na tecnologia quanto moldam a tecnologia. A medida e a direção do desenvolvimento tecnológico refletem as relações de gênero existentes, tanto quanto afetam a divisão sexual do trabalho

(WAJCMAN, 1998, p.256)

A divisão sexual do trabalho a qual Wajcman (1998) se refere inclui no cenário não apenas o trabalho produtivo, mas também o trabalho reprodutivo, associado e realizado por mulheres, muitas vezes fora de um contexto de retribuição monetária. O trabalho reprodutivo encontra-se amplamente associado ao cuidado e ao projeto humano de perpetuação da espécie, a partir do qual cabe as mulheres gestarem, parirem, criarem, educarem, alimentarem as crianças, mas também estender relações de cuidado para com maridos, pais e mães, por exemplo. Decorre dessa divisão a associação das mulheres com trabalhos remunerados como a enfermagem, a pedagogia, o secretariado, entre outros, nos quais são as habilidades de cuidado que estão postas em jogo, somadas às habilidades técnicas requisitadas (porém pouco reconhecidas nessas profissões). Tais trabalhos também são consideravelmente menos valorizados e não tão bem remunerados quanto outras funções.

O cuidado é a forma de expressão pela qual o corpo humano é capaz de manter e processar relações materiais, não apenas um tipo de estado emocional: a própria configuração do cuidado como apenas um sentimento é também uma forma de invisibilizar o cuidado enquanto uma categoria de trabalho (MACHEREETH, 2019). O processo de invisibilidade relacionado a essa categoria de trabalho (FEDERICI, 2017), está intrinsecamente relacionado a sua divisão sexual e de raça, por ser um trabalho especificado delegado para as mulheres – o trabalho não de produção, mas de reprodução.

O que Machereth propõe é a possibilidade de pensar a existência, a partir da ficção, de objetos pós-humanos relacionados ao cuidado, capazes de automatizar uma categoria de trabalho não apenas invisibilizada e, com frequência, não remunerada (MACHEREETH, 2019), mas também considerada como fora do âmbito dos processos de automação (FEDERICI, 2017). Desta forma, o cuidado enquanto categoria de trabalho não seria, necessariamente, uma exclusividade humana. Ainda assim, o debate acadêmico em torno do trabalho de cuidado, atravessado pela demarcação de gênero, realizado pelas assistentes pessoais digitais, por exemplo, foi considerado como uma perspectiva analítica séria quase que exclusivamente apenas por pesquisadores no âmbito dos robôs e da sexualidade (MACHEREETH, 2019).

A ponderação é correta, ainda que na mídia especializada a questão do gênero nas assistentes pessoais digitais tenha sido levantada e encarada enquanto um problema, na literatura acadêmica encontra-se pouca menção a esse ponto. A sexualização de robôs exclusivamente do gênero feminino (a despeito de uma já detectada demanda feminina de mercado por robôs de sexo masculinos), combinada com desenvolvimentos de Inteligência

Artificial, no âmbito da indústria do entretenimento sexual, entretanto, levanta questões dos teóricos da área no tocante a esta temática:

Então porque nós estamos faltando com as mulheres nos serviços de tecnologia? Talvez seja apenas um caso dos caras do Vale do Silício tentando recriar suas mães. Você sabe, alguém que está disponível o tempo todo, cuidando deles, organizando a vida deles. Ou talvez seja o esteriótipo de antiquado de secretária, silenciosa no background até ser necessária, sempre obedecendo as ordens do chefe. Tome nota, Ms. Jones: talvez nós precisemos ser eticamente e moralmente obrigados a repensar isso. E se nós estamos repensando nossas assistentes digitais, vamos também repensar nosso robôs para sexo. Preocupações sobre a objetificação de mulheres podem começar a ser endereçadas se nós fizermos os robôs para todos, não apenas para homens. Nós temos a chance de moldar e explorar a tecnologia, fazer dela mais igual e diversa, e nós não devemos nos eximir dessa oportunidade<sup>288</sup> (DEVLIN, 2018, p. 232 e 233).

Robôs sexuais desenhados a partir de esteriótipos sexistas despertam certa preocupação. Mas, vamos parar um minuto na proposta de uma criação de uma extensão materna no seio do Vale do Silício: é exatamente isto que propõe Sarah Sharma (2018) em sua leitura sobre as dinâmicas de trabalho em empresas como o Google e seu cruzamento com a emergência de uma *gig economy*. Uma parte dos aplicativos que fizeram emergir a chamada *gig economy*, de acordo com Sharma (2018), são formas de resoluções de todas as nossas necessidades primária da infância: alimentação (GrubHub), locomoção (Uber), serviços de limpeza (TaskRabbit) etc. Alguns cliques e é possível encontrar algo para comer, ir de um local a outro ou ter uma acomodação com cara de casa quase sem esforço algum – ao contrário da vida adulta, que requer trabalho para sanar qualquer uma dessas necessidades (e, eventualmente, resulta apenas na frustração ordinária de não poder ter um desejo atendido na hora exata). Esses aplicativos seriam uma extensão materna, uma replicação de uma condição de necessidades e desejos supridos por um outro que tem um trabalho: o cuidado.

O termo “economia pós-mãe” emerge para capturar esse momento particular na cultura da tecnologia (masculina) quando Uber (“mamãe, me leve”), TaskRabbit (“mamãe, limpe meu quarto”, GrubHub (“mamãe, estou com fome”) e LiveBetter (“mamãe, estou entendido”) emergiram. Mas sugerir que esse aplicativos foram desenhados para substituir a Mãe perde o ponto. Esses serviços não replicam, reproduzem ou substituem a Mãe. Ao invés disso, eles estendem o mandato maternal a todos os outros provedores de cuidado e expandem o reinado do consumo. O trabalho associado com a Mãe simplesmente acopla novos corpos e figuras geralmente

<sup>288</sup> No original: “Perhaps it’s simply a case of the dudes in Silicon Valley trying to recreate their mothers. You know, someone who is there for them round the clock, keeping an eye on them, organising their lives. Or maybe it’s the idea of the stereotypical old-fashioned secretary, silent in the background until needed, obedient to the boss. Take a note, Ms Jones: maybe we should be ethically and morally obliged to rethink that. And if we can rethink our digital assistants, let’s also rethink our sex robots. Concerns about the objectification of women can start to be addressed if we make the robots for everyone, not just for men. We have the chance to shape and explore technology, to make it more equal and diverse, and we shouldn’t shy away from the opportunity”.

não associados à Mãe, enquanto bens e serviços associados com o cuidado materno se expandem. Em outras palavras, o problema com a Mamãe não é só quem está fazendo esse trabalho, mas as demandas para esse tipo de trabalho na sociedade<sup>289</sup>

Seria espantoso que grandes empresas de tecnologia estejam voltadas em desenvolver aplicações capazes de nos colocar em posições confortáveis, como a casa da mãe, novamente? Para Sharma (2018), não. Esses desenvolvimentos seriam reflexo não apenas de uma cultura predominantemente masculina, branca e privilegiada, mas também da própria dinâmica de trabalho proposta pelas grandes empresas do Vale do Silício.

O retiro proporcionado pelo porão da Mamãe depende de um trabalho de cuidado desvalorizado associado com a Mamãe – não necessariamente com uma mãe específica, mas com “A Mãe” no senso psicanalítico de cuidado feminilizado pela virtude de sua própria diminuição. De fato, o privilégio de escapar a responsabilidade de quanto custa o cuidado de alguém é uma característica que define o poder masculino. É uma das formas a partir da qual o patriarcado funciona<sup>290</sup>.  
(SHARMA, 2018, sem paginação)

A estética “fofinha e divertida” de empresas como o Google, por exemplo, com locais de trabalho nos quais os trabalhadores têm sempre comidas e bebidas ao alcance, locais de repouso e de brincadeira, e nos quais podem se locomover com patins ou patinetes, são, de certa forma, a recriação de um ambiente voltado para o lazer e o conforto.

No centro dessas dinâmicas sociais e de trabalho é que os desenvolvimentos em tecnologia, como as assistentes pessoais digitais, vêm sendo desenvolvidas – não espantosamente reproduzindo o cuidado como trabalho automatizado. As assistentes colocam os serviços anteriormente mencionados não ao alcance de um clique, mas ao alcance da fala – tal como quando éramos crianças. As relações de cuidado, enquanto tecnologicamente mediadas e/ou propostas, trazem também, embutidas no design dos artefatos que as propõe, direcionamentos específicos de como estas relações devem ocorrer – acontece que, são

---

<sup>289</sup> No original: The term “post-mom economy” emerged to capture this particular moment in tech(bro) culture when Uber (“Mommy, drive me”), TaskRabbit (“Mommy, clean my room”), GrubHub (“Mommy, I’m hungry”), and LiveBetter (“Mommy, I’m bored”) emerged. But to suggest that these apps are designed to replace Mommy misses a key point. These services do not replicate, reproduce, or replace Mommy. Instead, they extend the maternal mandate to all other care providers and expand the realm of consumption. The labor associated with Mommy merely attaches to new bodies and figures not usually associated with Mommy, while the goods and services associated with Mommy’s care expand. In other words, the issue with Mommy is not just who is doing this labor, but the demands for this sort of labor in society.

<sup>290</sup> No original: The retreat of Mommy’s basement depends upon the devalued labor of caring associated with Mommy—not necessarily a specific mother, but “the Mother” in the psychoanalytic sense of attentive care feminized by virtue of its very diminishment. Indeed, the privilege of escaping responsibility for how much one’s care costs is a defining characteristic of masculine power. It is one of the ways patriarchy works (SHARMA, 2018, sem paginação)

justamente através das técnicas de cuidado que apreendemos as dimensões sociais e nos inserimos no mundo. Ou seja, “o social é reproduzido através do cuidado” (SHARMA, 2018, sem paginação).

De acordo com Devlin (2018), o sexismo endêmico do Vale do Silício aparece tanto no ambiente de trabalho quanto na própria tecnologia que é produzida. A exemplo do que podemos chamar de miopia de gênero, temos a primeira versão do Apple Health Tracker, criado para monitorar os sinais de saúde mais importantes dos usuários – o resultado? Um artefato sem acomodação para o ciclo menstrual feminino, considerado pela maioria das mulheres sem problemas de saúde como um dos sinais corporais mais importantes de monitorar. É justamente na cultura sexista, reproduzida nos desenvolvimentos tecnológicos, que se situam as principais preocupações de Broussard (2018) em relação aos desenvolvimentos que envolvem inteligência artificial no âmbito do Vale do Silício – local onde se concentram as maiores empresas do setor, como a Microsoft, a Apple, o Google e a Amazon.

Para ambas (DEVLIN, 2018; BROUSSARD, 2018) as assistentes pessoais digitais seriam, nesse contexto, apenas mais uma expressão de um sexismo já vigente, repetidamente endossado por construções científicas que também são majoritariamente produzidas por homens, o que também pode levar a vieses na própria produção científica (HARDING). Pontos cegos na produção científica podem estender vieses por décadas, impactando, justificando e ocultando escolhas não-científicas. No que diz respeito à decisão de utilizar vozes femininas nas assistentes pessoais digitais, Devlin (2018) comenta a respeito das construções científicas em torno do tema, principalmente a partir do já comentado estudo de Nass, Steuer e Tauber (1994) no âmbito da interação humano-computador:

Existe uma quantidade de explicações supostamente científicas para explicar porque as vozes femininas são uma escolha melhor para uma interface conversacional. Alguns apontam para o trabalho de Nass, que sustenta que os computadores são atores sociais com gênero e que esse gênero é uma “pista extremamente poderosa”. Nass e seus coautores descobriram isso a partir de um experimento que eles fizeram a partir de testes de feedback de respostas computadorizadas. “Os usuários irão aplicar esteri[otipos de gênero?” o artigo se pergunta. “Esse experimento demonstra que “pessoas aplicam as seguintes normas sociais a computadores: “Elogios de homens são mais convincentes do que de mulheres,” “Homens que elogiam são mais gostados que mulheres que elogiam,” e “Mulheres sabem mais sobre amor e relacionamentos do que homens”. Afirmações amplas. O experimento teve 48 participantes: 24 homens e 24 mulheres. Essas são extrapolações muito amplas para um estudo tão pequeno. Mas, em defesa dos autores, eles não disseram que essa é a forma que os computadores deveriam responder. Eles disseram que “os estereótipos de gênero aplicados nas interações entre humanos são aplicados nas interações entre humano-computador quando os computadores empregam vozes com gênero”. Eles não recomendaram que o gênero fosse um fator de interação no design, ainda assim, ao longo dos anos, isso foi interpretado como usuários preferindo informações

autoritárias de vozes masculinas e subserviência de vozes femininas<sup>291</sup> (DEVLIN, 2018, p.231).

Para além da mostra pouco significativa de participantes para a realização de especulações tão amplas como resultado, o texto de Nass, comentado no trecho acima, apresenta outros problemas metodológicos previamente expressos nesse trabalho, tais como a inexistência de dados de perfil dos participantes, a não ponderação de possíveis vieses na condução da pesquisa e a inexistência de um material de apoio passível de acesso que detalhasse o conteúdo dos formulários respondidos pelos participantes nas condições descritas. Ainda, Devlin (2018) aponta para declarações amplas e infundadas de Nass, como a afirmação, em entrevista para a CNN, em 2011, de que seria um ‘fenômeno bem estabelecido que o cérebro humano se desenvolveu para gostar de vozes femininas’. Em consulta com o pesquisador e especialista em bebês, Casper Addyman, Devlin descobriu que na verdade não existem indícios que embasem tal afirmação: estudos na área apontam que bebês gostam mais de vozes e rostos femininos até os oito meses de idade, sem qualquer evidência posterior a essa idade (DEVLIN, 2018).

Essa suposta preferência dos usuários por vozes femininas de forma geral, se é que tal preferência pode ser efetivamente afirmada “parece ter menos relação com o som, o tom, a sintaxe ou a cadência do que com uma associação com assistência<sup>292</sup>” (WEST, KRAUT e CHEW, 2019, p.98). A motivação da escolha para o gênero da voz em artefatos digitais parece se relacionar muito mais com a funcionalidade a ser executada por um assistente de voz: a inteligência artificial da IBM, Watson, quando venceu um jogador humano em Jeopardy, em 2011, era indistintamente masculina; no Japão, vozes automatizadas femininas são utilizadas para conceder informações sobre ações na bolsa de valores, mas ocorre uma troca para a voz masculina para realizar e efetivar transações financeiras (WEST, KRAUT e CHEW, 2019). A preferência pela voz feminina, se varia conforme as situações e as ações desta voz, possui muito

---

<sup>291</sup> No original: There are a number of supposed scientific explanations about why a female voice is the better choice for a conversational interface. Some people point to Nass’s aforementioned paper, which states that computers are gendered social actors, and that gender is an ‘extremely powerful cue’. Nass and his co-authors found this out from an experiment they ran to test responses to computerised feedback. ‘Will users apply gender stereotypes to computers?’ the paper asks. ‘This experiment demonstrates that “people apply the following social rules to computers: “Praise from males is more convincing than praise from females,” “Males who praise are more likeable than females who praise,” and “Females know more about love and relationships than males. Broad claims. The experiment had 48 participants: 24 males and 24 females. That’s a lot of extrapolation from a tiny study. But, in defence of the authors, they didn’t say that’s how computers should respond. They wrote that ‘gender stereotypes applied in human–human interaction are applied in human–computer interaction with computers employing gendered voices’. They didn’t recommend that gender be a key factor in interaction design, yet over the years, this has been interpreted as users preferring authoritative information from male voices and subservience from female voices. (DEVLIN, 2018, p.231)

<sup>292</sup> No original: “seems to have less to do with sound, tone, syntax and cadence, than an association with assistance”



mais uma conotação socialmente construída do que qualquer outro tipo de atributo físico ou acústico.

A facilidade com a qual pesquisadores e desenvolvedores de tecnologia encontram explicações racionais para escolhas que são frutos de padrões sociais e culturais cai, inevitavelmente em “um paradoxo óbvio no centro da sociedade patriarcal: se se vai insistir na naturalidade das diferenças de gênero, então porque tanta tecnologias na verdade ignoram as diferenças entre os corpos de homens e mulheres? Porque é uma falácia”<sup>293</sup>. (SHARMA, 2018, sem paginação).

A principal questão que envolve tecnologias de inteligência artificial guiadas por um viés sexista decorre do papel crescente que as tecnologias de mídia adquirem nos modos de vida humanos, criando “parâmetros do que é possível. Tecnologias alteram concepções e experiências do tempo, espaço e distância – assim como de diferença social e de gênero. Tecnologias alteram o que significa ser humano e o que significa estar em relação um com o outro” (SHARMA, 2018, sem paginação).

#### 7.4 QUAL É A VOZ DA SUA ASSISTENTE?

Ao longo das etapas da pesquisa empírica, a relação entre o gênero das assistentes e as preferências, percepções e interações guiou parte da coleta dos dados. Na sequência iremos apresentar as análises dos dados quantitativos e qualitativos obtidos sobre gênero e voz das assistentes, levando em conta os atributos e questões previamente debatidos neste capítulo.

##### 7.4.1 PREFERÊNCIAS E PERCEPÇÕES DOS RESPONDENTES DO FORMULÁRIO ONLINE SOBRE O GÊNERO DA VOZ

Os respondentes do formulário foram perguntados sobre os seguintes fatores: a existência de preferência de gênero de voz para a assistente; se sim, qual; se sim, por qual

---

<sup>293</sup> No original: “an obvious paradox at the core of patriarchal society: if it is going to insist on the naturalness of gendered differences, then why do so many technologies actually ignore the differences between women’s and men’s bodies? Because it is a fallacy”

motivo; a percepção do usuário sobre o gênero da voz alterar ou não a interação com o artefato. O gráfico abaixo apresenta os dados referentes a preferência de utilização de uma voz com gênero específico para os participantes do formulário quantitativo.

Figura 60 Gráfico de preferência do gênero da voz



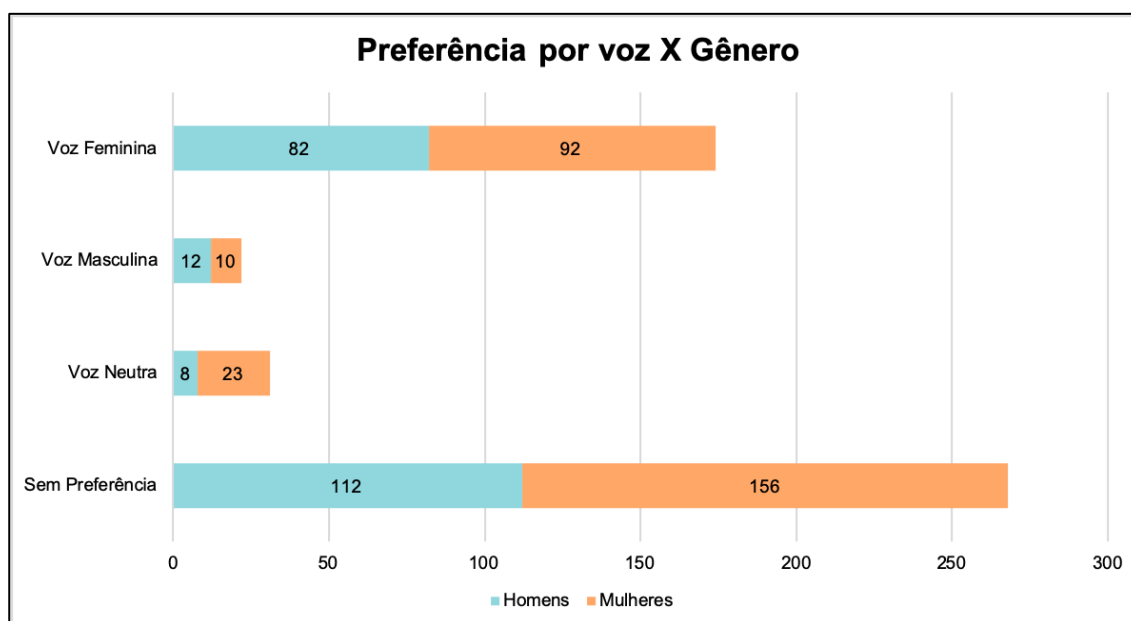
Fonte: a autora

Quando perguntados sobre suas preferências de utilização de voz nos sistemas, os respondentes do formulário expressaram, em sua maioria, não possuir preferência alguma (54%). Entretanto, este formulário não contemplava a pergunta sobre qual voz era efetivamente utilizada pelos participantes, independente de uma preferência expressa – durante as entrevistas em profundidade, entretanto, todos os quatorze entrevistados afirmaram utilizar a voz feminina em todas as suas assistentes. Dos respondentes do formulário, 54% afirmaram não ter preferência enquanto 45% afirmaram preferir um gênero em específico, sendo 35% de preferência por voz feminina, 6% preferência por uma possível voz neutra, 4% preferência por voz masculina e 1% marcaram a categoria outros.

Ainda que com uma diferença baixa em termos estatísticos, uma voz neutra ainda não existente aparece a frente de uma voz masculina – os motivos para essa questão são múltiplos, e serão discutidos na sequência, a partir de dados quantitativos e qualitativos. A preferência por uma voz feminina, aparentemente corrobora a ideia difundida entre cientistas da computação e da interação humano-computador de uma maior agradabilidade desta voz em relação à masculina, os termos dessa preferência, entretanto, não são tão simples.

Quando separamos os resultados por gênero, ou seja, analisamos em separado as respostas de pessoas de gênero masculino e as respostas de pessoas de gênero feminino, a diferença entre as porcentagens da amostra geral são pouco significativas principalmente em função da diferença entre participantes homens e mulheres, conforme exposto no capítulo 3 deste trabalho. O gráfico abaixo apresenta a preferência pela voz feminina, masculina, neutra ou a não preferência marcando os respondentes de cada gênero.

Figura 61 Gráfico de preferência de voz dividido pelo gênero dos respondentes



Fonte: a autora

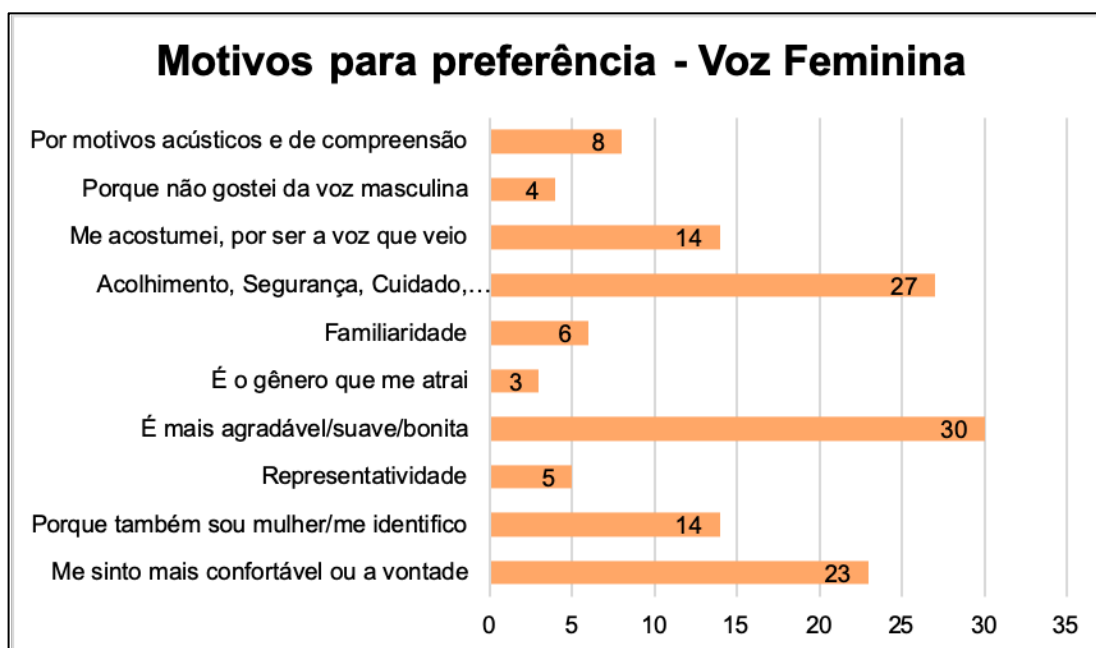
Observamos que a maioria das categorias segue a diferença de respondentes entre homens e mulheres, apresentando sempre um resultado um pouco mais expressivo de respondentes mulheres, que são a maior parte das respondentes no montante total. O único momento em que os respondentes homens são maiores do que as respondentes mulheres é na preferência de voz masculina, onde os 4% dos participantes que preferem esta voz se dividem entre dez mulheres e doze homens. Esse resultado é inesperado e contradiz parcialmente a literatura encontrada. O já mencionado relatório da EQUALS, informa que, ao longo de sua extensa pesquisa encontrou “muitos depoimentos de mulheres trocando as configurações padrões de uma voz feminina para uma voz masculina quando esta opção estava disponível, mas não encontrou uma única menção de um homem trocando a voz feminina padrão por uma voz masculina” (WEST, KRAUT e CHEW, 2019, p. 97).

Não é possível afirmar, a partir das perguntas realizadas no formulário, que estes doze

participantes efetivamente trocaram a voz padrão por uma voz masculina quando possível, entretanto, é possível inferir que, dado sua preferência, o fizeram ou fariam se possível. Além disso, que a quantidade de respondentes que preferem a voz masculina possua mais retornos de homens que de mulheres também contradiz a ideia difundida de que haveria uma preferência maior por vozes do sexo oposto (WEST, KRAUT e CHEW). Não apenas o número de homens que prefere voz masculina é maior que o de mulheres que preferem a mesma voz, mas também o número de mulheres que expressa preferência pela voz feminina é expressivo (92 participantes) e significativamente maior do que o número de mulheres que prefere a voz do sexo oposto (apenas 10 participantes).

Aos participantes que informaram ter uma preferência específica para voz, foi solicitado que explicassem os motivos de suas preferências. A resposta a esta pergunta não era obrigatória no formulário, não sendo preenchida por todos os participantes. As respostas livres foram agrupadas a partir de recorrências, gerando categorias emergentes que agrupavam um determinado conjunto de respostas, com possíveis variações, mas de sentido comum. Foram criadas dez categorias emergentes para os motivos de preferência para voz feminina, cinco categorias emergentes para voz masculina e quatro categorias emergentes para voz neutra. Destas, duas são comuns ao gênero masculino e feminino ('atração pelo gênero' e 'por ser mais agradável') e uma é comum aos três ('porque não gostei da voz feminina/masculina/ambas'). O gráfico abaixo apresenta as categorias que emergiram das respostas livres dos participantes, quando preferiram a voz feminina.

Figura 62 Gráfico Motivos para preferência da voz feminina



Fonte: a autora

Os principais motivos para preferência da voz feminina foram características relacionados ao timbre de voz, especificados por qualidades como ser mais bonita, mais suave ou mais agradável (30 dos respondentes que preferiram voz feminina). Seguido de 27 respondentes que informaram preferir por relacionarem a voz feminina a sentimentos de segurança, acolhimento, cuidado, carinho, entre outros, categoria de respostas que responde a uma associação da voz feminina com o cuidado enquanto trabalho. A terceira categoria emergente também diz respeito a sentimento acionados pela voz, que levam os respondentes a se sentirem mais confortáveis ou a vontade para falar (23 dos respondentes).

Em um total de 50 respostas, o somatório das últimas duas categorias mencionadas, os respondentes informam que preferem a voz feminina pela forma como essa voz lhes faz sentir, o que aponta para uma construção social associada ao papel da mulher. Essa percepção é similar a mencionada pelo relatório da EQUALS, com pesquisas que apontam para o fato de que vozes femininas são consideradas como mais preocupadas ou capazes de ofertar ajuda: “a percepção talvez tenha raízes em normas sociais tradicionais em torno das mulheres como fornecedoras de cuidado e encorajamento (mães com frequência assumem – querendo ou não – significativamente mais trabalho de cuidado do que os pais) e outros vieses de gênero

construídos socialmente que são anteriores a era digital”<sup>294</sup> (WEST, KRAUT e CHEW, 2019, p. 98).

Com 14 respondentes cada, as categorias emergentes ‘porque também sou mulher/me identifico’ e ‘me acostumei, pois foi a voz que veio’, apontam para dois fatores não esperados: na primeira, a existência de mulheres que preferem que seu dispositivo de voz tenha também uma voz feminina por se identificarem com outras mulheres – essa categoria apresenta desdobramentos quando analisamos as entrevistas em profundidade. A segunda, demonstra que dotar um dispositivo com características específicas pode levar a uma habituação dos usuários, tornando esse traço normalizado, ou até mesmo preferido.

Depois, temos as categorias ‘por motivos acústicos’ (8 respondentes), de participantes que afirmavam, por exemplo, que a voz feminina era mais fácil de ser ouvida e compreendida, ou que se propagava melhor em ambientes ruidosos (apresentados, muitas vezes, como argumentos científicos); ‘familiaridade’ (6 respondentes); ‘representatividade’ (5 respondentes), que apontam que a voz feminina nesses dispositivos concede visibilidade a mulheres, ou possibilidade de ocupar outros espaços; ‘porque não gostei da voz masculina’ (4 respondentes) e ‘porque é o gênero que me atraí’ (3 respondentes).

A ideia de representatividade através da voz feminina, por parte de participantes de gênero feminino, também foi uma surpresa, apontando que a leitura da construção destes dispositivos enquanto mulheres possui interpretações, percepções e impactos múltiplos, não sendo necessariamente sempre atrelada a uma visão negativa desta, ao contrário do que ocorre na literatura especializada. A relação entre o gênero pelo qual a pessoa é atraída e a preferência por voz coloca mais uma camada de análise, o viés não apenas antropomorfizado com o qual nos relacionamos com esses objetos, mas também a mediação da sexualidade, e não apenas do gênero, enquanto componentes desta interação. Todos os participantes que responderam preferir a voz feminina por ser o gênero de atração se identificaram como do gênero masculino.

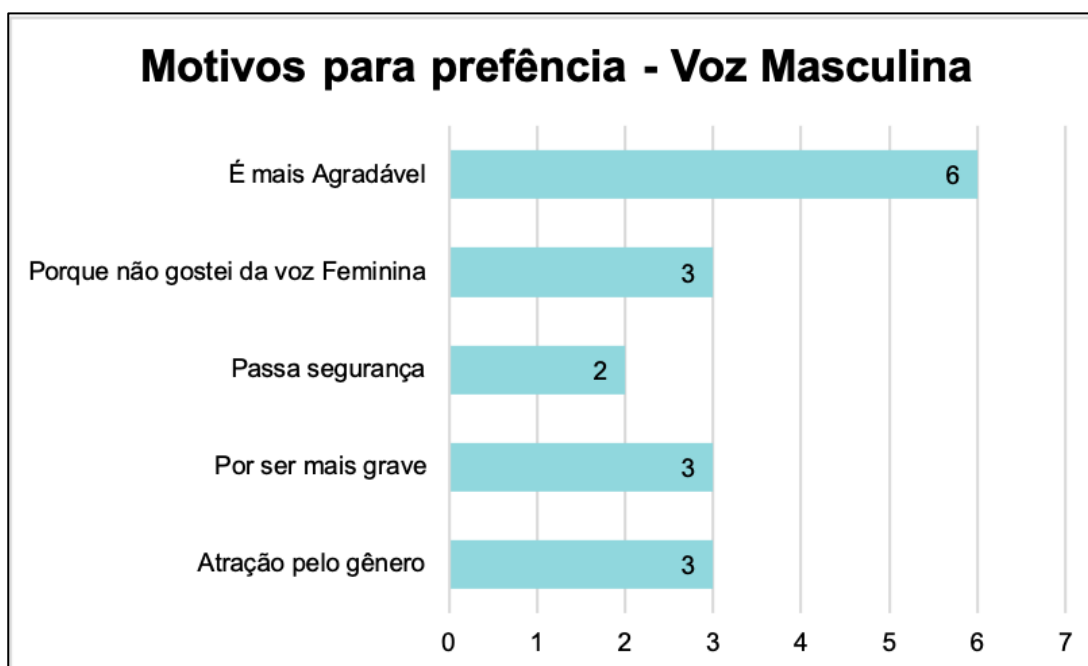
A sexualidade como fator de mediação da interação entre ser humano e assistente também aparece como fator na preferência pela voz masculina (3 respondentes). Entretanto, neste caso, dois dos respondentes se identificaram como do gênero masculino e homossexuais, sendo apenas uma das respondentes identificada enquanto mulher. Entre os seis respondentes, portanto, que apontaram a atração sexual por um gênero enquanto motivo para preferência de voz, cindo deles eram homens e apenas uma era mulher, apesar da proporção de mais

---

<sup>294</sup> No original: “The perception may have roots in tradicional social norms around women as nurturers (mothers often take on - willingly or not - significantly more care than fathers) and other socially constructed gender biases that predate the digital era”

respondentes mulheres para esta pesquisa em números totais. Esse resultado pode apontar para um viés de percepção masculino, interessante de ser investigado, entretanto, os dados obtidos não são suficientes para maiores inferências. O gráfico seguinte apresenta as cinco categorias emergentes das respostas abertas em relação a preferência pela voz masculina.

Figura 63 Gráfico Motivos para preferência de voz neutra

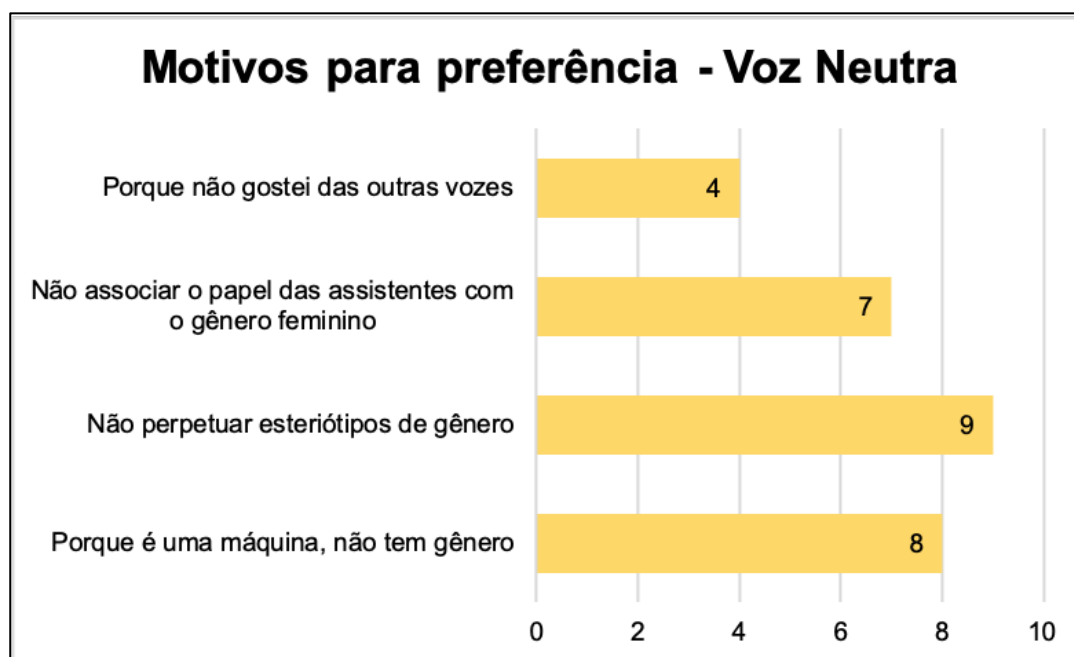


Fonte: A autora

O principal motivo para preferência da voz masculina é por ser ‘mas agradável’ (6 respondentes), seguido de ‘porque não gostei da voz feminina’ (3 respondentes), ‘por ser mais grave’ (3 respondentes), ‘por ser o gênero que me atrai’ (3 respondentes) e ‘passa segurança’ (2 respondentes). Dos respondentes que preferem a voz masculina, apenas dois deles apontam preferir em função de sentimentos relacionados a essa voz, ‘segurança’, em contrapartida com 50 participantes que afirmaram preferir a voz feminina em função de sentimentos associados (variando entre ‘conforto’, ‘cuidado’, ‘acolhimento’, entre outros). Essa discrepância aponta para a existência de uma ramificação maior de sentimentos e construções relacionadas ao feminino do que ao masculino, para pessoas de ambos os sexos.

O gráfico abaixo faz a comparação das categorias emergentes das respostas relacionadas a preferência pela existência de uma voz neutra, não relacionada a nenhum dos gêneros (que ainda não é realidade em nenhuma das assistentes).

Figura 64 Gráfico Motivos para preferência de voz neutra



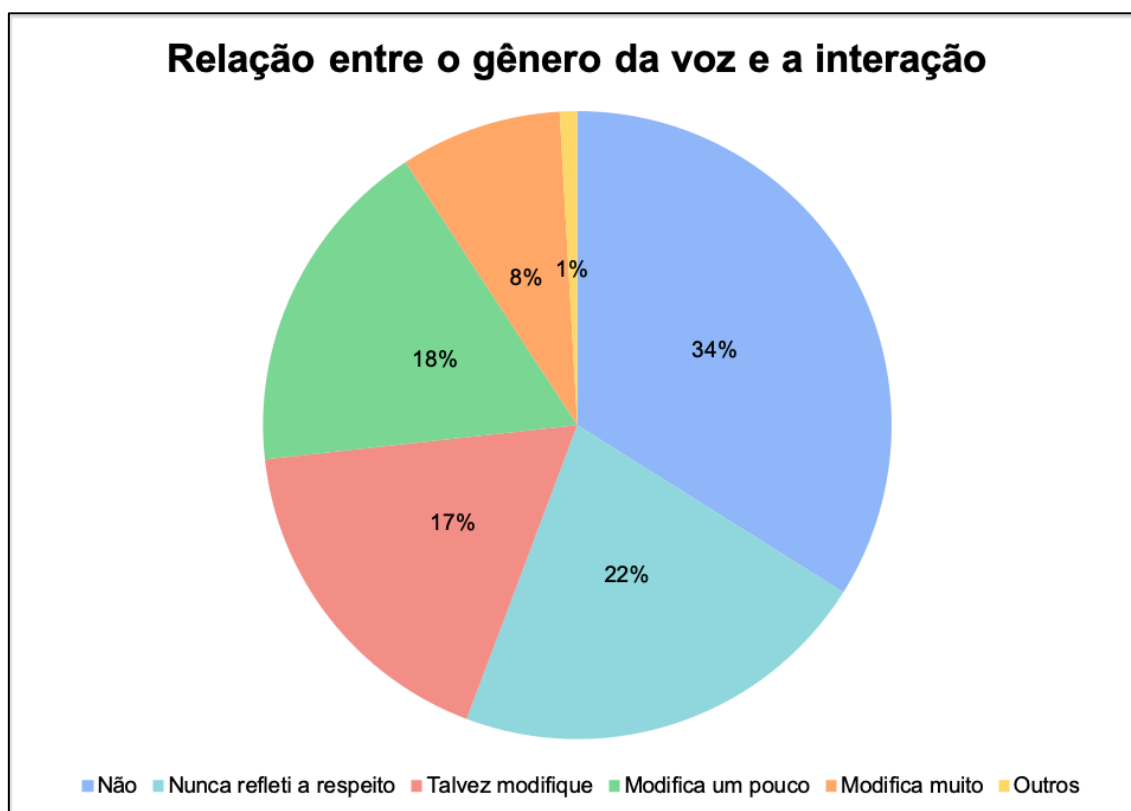
Fonte: a autora

Dos respondentes que informaram preferir uma voz neutra, 24 deles afirmaram que esta preferência se dá em função da não associação das assistentes com um gênero em específico: ‘não perpetuar estereótipos de gênero’ (9 respondentes), ‘porque é uma máquina, não têm gênero’ (8 respondentes) e ‘não associar o papel das assistentes com o gênero feminino’, em termos de função (7 respondentes). Apenas 4 participantes forneceram motivos distintos das relações de gênero, afirmando que preferem a voz neutra ‘porque não gostaram das outras vozes’. Esse resultado aponta para a existência de um incômodo, por parte de alguns usuários, da associação das tarefas, formas de falar e expressar das assistentes com o gênero especificamente feminino ou mesmo com características antropomorfizadas nesses sistemas.

Quando perguntados sobre a relação entre o gênero da voz e a interação estabelecida com o dispositivo, os respondentes se comportaram conforme o gráfico abaixo.



Figura 65 Gráfico Relação entre o gênero da voz e a interação

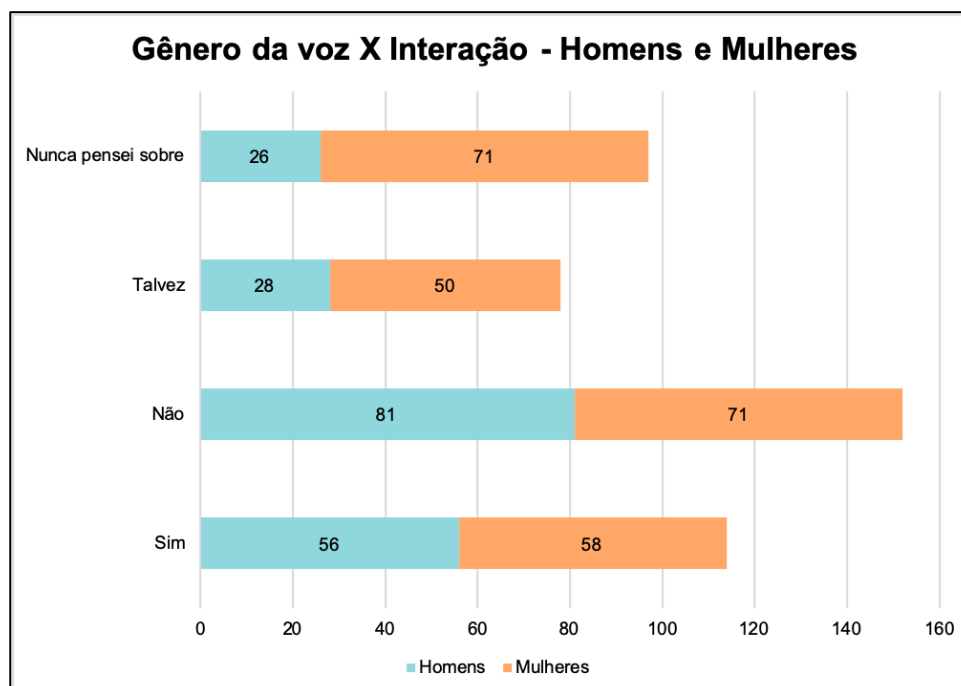


Fonte: A autora

Os respondentes informaram que o gênero da voz: não modifica a interação (34%); nunca pensaram a respeito disso (22%), modifica um pouco a interação (18%), talvez modifique a interação (17%), modifica muito a interação (8%) e outros (1%). Em termos mais gerais, são 34% dos participantes que acreditam que gênero é um fator para a interação enquanto outros 26% afirmam que sim. Na categoria outros, três resultados interessante emergiram: um respondente, do sexo masculino, fez questão de escrever ‘Não, **tenho certeza** que não faz diferença alguma’, ao invés de marcar a opção ‘Não, **acho que** não faz diferença alguma’, demonstrando a potência de sua convicção; um respondente, do gênero masculino, informou que ‘Se fosse voz de homem eu mandaria ele se ferrar. Na verdade, eu queria um escravo virtual para poder zoá-lo o tempo todo’, o que aponta uma certa violência na interação que pode estar sendo ou não atualmente direcionada para a assistente feminina; um terceiro respondente, do gênero feminino, disse que ‘depende das minhas reações culturais em cada ocasião, inclusive inconscientes’, o que aponta para uma variação das percepções da voz em diferentes contextos.

Na sequência, representamos graficamente os mesmos resultados, porém agrupando as duas respostas ‘sim, muito’ e ‘sim, um pouco’ em apenas ‘sim’ e excluindo a categoria outros. Nesse gráfico demarcamos a quantidade de respostas de cada gênero em cada categoria.

Figura 66 Gráfico Gênero da voz X Interação – Homens e Mulheres



Fonte: A autora

Quando separamos os resultados por gênero, as proporções das respostas demonstram variar bastante a partir deste fator, lembrando sempre da preponderância de respondentes mulheres em relação a homens. A quantidade de homens (56) e mulheres (58) que responderam que o gênero pode alterar a interação é similar; mais homens (81) acreditam que o gênero da voz não interfere na interação do que mulheres (71); os homens também se mostraram mais assertivos sobre a questão, apenas 28 deles informaram ‘talvez’, comparado a 50 mulheres; e apenas 26 homens disseram ‘nunca pensei sobre isso’, comparado a 71 mulheres.

#### 7.4.2 PREFERÊNCIAS E PERCEPÇÕES DOS ENTREVISTADOS SOBRE O GÊNERO DA VOZ

Durante as entrevistas em profundidade os participantes foram perguntados, com variações entre os participantes, sobre: qual voz eles utilizavam em sistemas de GPS com direções por voz (como Google Maps e Waze); qual voz eles utilizavam em suas assistentes pessoais digitais e se já haviam testados outras vozes; sobre preferência de gênero da voz para as assistentes; posteriormente foram informados sobre o padrão de voz feminina nas quatro assistentes e perguntados sobre suas percepções/ideias sobre a questão. Nesse sentido, algumas das questões observadas a partir da revisão teórica, assim como do formulário quantitativo, se repetem nas entrevistas, com a possibilidade de aprofundamento de alguns pontos. Na sequência iremos analisar as preferências de gênero na voz das assistentes e os motivos destas preferências, assim como percepções associadas a esta questão.

Do total de entrevistados, todos os quatorze informaram utilizar a voz feminina em suas assistentes – isso se estendeu para aqueles que possuíam mais de um dispositivo em casa (podendo chegar a mais de cinco por pessoa) e aos demais membros da família que utilizavam. Os entrevistados E1, E7, E10, E11 e E12 disseram que não sabiam que era possível mudar o gênero da voz (5 pessoas); E2 e E3 utilizavam apenas assistentes que não permitem a mudança do gênero da voz (2 pessoas); E6, E8, E9, E11 e E12 disseram que sabiam que era possível mudar o gênero da voz, mas que não fizeram (5 pessoas); e E4 e E5 disseram que mudaram o gênero da voz para experimentar, mas retornaram para a voz padrão (2 pessoas).

A entrevistada E1, ainda que não soubesse da possibilidade de modificar a voz da assistente, ‘não mudaria, meninas no poder!’, pois ‘acho que já me acostumei com a voz feminina – e também de me sentir mais próxima, não sei explicar, pra mim a voz é só feminino’. Ocorre aqui uma subversão das análises predominantes na literatura, de que as assistentes seria submissa: para E1, a assistente está no poder, o que é interessante para pensar tanto percepções de gênero quanto sobre o próprio algoritmo. Quando perguntada sobre o impacto dessa voz na interação, a E1 afirma que ‘tem perguntas que eu acho que não faria por não me sentir a vontade. Por exemplo, ir na farmácia comprar anticoncepcional, eu me sinto mais a vontade de comprar com uma mulher que com um homem. Acho que se fosse uma voz masculina, para certas perguntas, eu pensaria ‘ai, é um homem respondendo’. Pode parecer algo idiota, mas é algo que a gente vai quebrando aos poucos. Quando é outra mulher a gente se identifica mais, quando é um cara a gente fica: vai me julgar’.

Aparecem nessa fala relações com a voz que são similares com as obtidas no formulário quantitativo: a identificação com a voz por ser do mesmo gênero, gerando cumplicidade; se sentir confortável para interagir com uma voz feminina; e gostar da voz feminina no dispositivo

por identificar algo de caráter feminista, de representatividade (denotado pelo ‘meninas no poder’).

A relação entre o gênero e uma divisão sexual do trabalho aparece também nas entrevistas, quando a assistente pessoal digital é descrita pelo entrevistado E2 como “meio que uma secretaria virtual”, que além de uma funcionalidade ligada as habilidades secretariais, de organização, estaria relacionado ao cuidado, uma vez que tem como finalidade também fazer companhia ‘eu moro sozinho e curto ter a ideia de ter alguém pra conversar, então eu gosto de conversar com ela’. Entretanto, o gênero da voz não é percebido como um fator que iria alterar a interação estabelecida pelo entrevistado com o dispositivo ‘porque a voz feminina é mais meiga, é melhor’.

A preferência pela voz feminina emerge mesmo quando os entrevistados não tem experiência com nenhuma outra voz além da padrão. O entrevistado E3 afirma que, mesmo que nunca tenha ouvido voz masculina no seu dispositivo, a voz feminina ‘deve ser mais agradável que a voz masculina... eu prefiro, mesmo sem ter experimentado [outra].

E4 informa que não alterou a voz dos dispositivos de GPS com comando de voz, utilizando aquela que vem por padrão – entretanto, acredita que a voz que utiliza no Google Maps seja masculina, e, ao conferir, se surpreende por ser feminina. Já a voz da assistente utilizada, a Siri, foi alterada para o masculino para testar, mas ‘eu prefiro a feminina, eu alterei provavelmente por cinco minutos. Eu me sinto mais confortável eu acho com vozes femininas de forma geral – é um pouco mais acolhedor’. Retornam aqui os sentimentos que o gênero da voz é capaz de evocar, associados no questionário quantitativo principalmente com a voz feminina e com traços como acolhimento, aconchego, calma. A voz feminina é compreendida, nesse sentido, por parte dos participantes, como capaz de evocar sentimentos e estados emocionais mais agradáveis, o que aparece como um fator para a escolha da voz e também para a interação. A E4, por exemplo, aponta que o gênero da voz é um fator que modela a interação, pois ‘eu associo muito a voz feminina a algo mais acolhedor e mais suave, provavelmente eu interajo da mesma forma com essa voz. Eu sou muito responsiva à forma como eu sou tratada, tanto por uma pessoa quanto por um aplicativo, então provavelmente a interação muda sim’.

Quando perguntada se a voz feminina é uma preferência também no momento de receber ordens, como no GPS, a E4 diz que “sem sombra de dúvidas, se tivesse um homem me dizendo dobre a direita talvez eu fosse ficar mais braba [risos]. Coisas que a gente pensa só se é perguntado né, [risos]”. Nesse caso, o gênero da voz é importante tanto na modulação de uma interação, sendo capaz de gerar formas e tons de feedback do usuário diferentes, quanto é capaz de alterar o estado emocional e/ou os sentimentos do usuário. Enquanto receber direções de

uma voz feminina pode ser considerado agradável pela entrevistada, o mesmo pela voz masculina pode ser irritante. A ideia de que a voz masculina não é adequada para mulheres em algumas situações, aparece: ‘eu não quero um homem me dando ordens no trânsito’.

Essa reação é similar e, ao mesmo tempo, oposta ao caso de um dos primeiros sistemas de navegação em veículos que incluíram a voz femininas para fornecer as instruções, no final dos 90, no carro BMW 5 Series. Naquela época, o carro foi retirado do mercado na Alemanha “porque muitos motoristas registraram reclamações por receber instruções de direção vindas de uma mulher” (WEST, KRAUT e CHEW, 2019, p. 99). Se, nesse período, a adoção da voz masculina era por um rechaço da voz feminina naquele papel, no caso da E4, a escolha pela voz feminina para GPS ocorre pela inadequação de receber direções de um homem. Ambos os casos, entretanto, refletem construtos e quadro sociais: no caso da BMW, de que mulheres não tem senso de direção e que não devem dar ordens aos homens; no caso da E4, é a inversão da lógica estabelecidas das relações de gênero e a negação de receber ordens de uma voz masculina (ou seja, de colocar-se em uma posição a qual muitas mulheres foram historicamente submetidas).

A E4 conta como descobriu que era possível trocar a voz da assistente pessoal: ‘o meu marido me falou “tu sabe que A Siri pode ser O Siri?”, e daí a gente entrou em uma discussão inteira sobre o gênero da Siri, e eu testei para ver como era’. A questão da construção de um gênero para um objeto digital aparece novamente, com uma ideia também de curiosidade e diversão, da especulação sobre o que seria aquela forma de inteligência artificial e como seria a expressão de gênero em algo assim. Essa questão marca um duplo posicionamento percebido frente as assistentes: ao mesmo tempo em que identificamos padrões e formas de interação humanos, o que pode gerar estranhamento, a curiosidade em torno de como aquele artefato irá se comportar leva muitos dos participantes a testarem os limites das assistentes (com perguntas engraçadas, ou humanizadas) e a contrapor o artefato da vida real tanto com o imaginário de inteligência artificial que possuímos quanto com o *hype* construído em torno desses desenvolvimentos pelas empresas do Vale do Silício.

A relação com a voz feminina da E5, se apresenta como um aspecto forte do reconhecimento com o artefato, que independe de conhecimentos acadêmicos sobre o assunto, uma vez que E5 atualmente conduz uma pesquisa que envolve o dispositivo Google Home. ‘É engraçado assim, porque essa relação com a voz... eu sei que o Google vende O Google Home, O Google Assistente. Mas eu não consigo chamar de O [masculino] eu chamo de A [feminino]. E isso é um conflito pra mim [...] poxa, se o dono do negócio tá falando que é [risos]. Porque a voz é feminina para mim, para o meu cérebro me diz que aquela voz é feminina, então se é

feminina não é O Google Home, não é O Google Assistente, é A Google Assistente, é A Google Home, é A Alexa, porque é feminina... a voz que sai de dentro daquele corpo ali é feminina [risos]’.

A passagem evidencia a importância da voz na construção do gênero, assim como do gênero para a relação estabelecida, mesmo que E5 tenha conhecimento acadêmico sobre o artefato e compreenda as formas de funcionamento dessa máquina. A entrevistada aponta que logo no começo tentou trocar a voz, mas não havia essa possibilidade – como demorou para conseguir trocar, quando foi possível: ‘eu não consegui, eu já tinha me acostumado com a voz. Aí as vezes eu deixo assim, um dia sabe, mas a voz masculina não dá! Eu fico: “Cadê a minha Google?! Você levou embora a minha assistente!” [muitos risos]’. A voz feminina é, para E5, fator de construção da personalidade e a troca desta voz desconfigura, modifica e quase impede a interação com o dispositivo da mesma forma que se dava anteriormente. Percebemos que a mudança de voz, uma vez que a relação tenha se estabelecido com uma voz em particular, traz consigo uma mudança aparente do artefato em si, como se ele não fosse mais o mesmo. Esse alto grau de identificação do dispositivo com a voz que o anima e com o gênero que conforma suas características é, ao mesmo tempo, ponto relevante para o estabelecimento de uma relação com o objeto mais profunda, com um índice elevado de antropomorfização, sem deixar de ser preocupante pelas relações que vem a estabelecer entre o gênero feminino e máquinas que obedecem e executam tarefas.

Ao tentar explicar o motivo da preferência pela voz feminina, E5 aponta como motivos a ‘identificação de gênero’, uma relação ‘com meu próprio timbre de voz, com a forma como eu falo’, o sentimento de proximidade criada ‘minha amiga Google que está ali conversando comigo’ e a sensação de segurança da voz feminina para se ter dentro de casa ‘as vezes a gente tem, enquanto mulher... tu tem medo dos machos, tu tem medo dos caras... tá dentro da minha casa, eu moro sozinha, eu tipo não quero uma voz masculina [desconhecida] aqui sabe [...] as vezes eu escuto uma voz masculina fora, e eu tô sozinha, eu fico alerta [...] se eu já tenho medo de ouvir uma voz masculina na rua imagina em casa’. Essas colocações auxiliam a compreender alguns dos motivos para preferência das voz feminina no formulário quantitativo: a sensação de segurança dos participantes pode ser, por um lado, de maior liberdade na hora de falar, na compreensão da assistente como uma amiga, ou uma voz que acolhe, mas, por outro, por ser uma voz que não representa ameaça para o sexo feminino. Tais colocações não aparecem na literatura revisada, o que pode indicar uma particularidade cultural, interessante para ser investigada. As questões colocam mais uma camada para compreender como esses objetos são construídos enquanto femininos, mas também das diversas apropriações e sentidos que as

diferentes experiências de vida podem estabelecer com o artefato: a voz feminina na assistente não significa, para todas as mulheres, a perpetuação de um estereótipo de gênero, mas pode significar maior representatividade e ocupação de espaços variados por mulheres (inclusive em objetos), assim como uma relação tranquila e não ameaçadora com um artefato que habita a intimidade da sua casa.

Por fim, é possível que todos esses sentidos, de representação, de identificação, de sentimento de não-ameaça, de estereótipos de gênero e de divisão sexual do trabalho convivam como parte da percepção de uma mesma pessoa sobre um mesmo objeto. Na entrevista, E5 aponta nessa direção, ao conciliar suas sensações positivas em relação ao gênero da Google Home, com uma percepção negativa da representação feminina, como parte da própria complexidade das relações de gênero na sociedade contemporânea: ‘é muito da sociedade que a gente vive, de tudo que significa ser mulher pra mim assim. O que me causa um certo ranço, não é ranço a palavra... uma certa luta interna, é que eu não quero pensar a mulher enquanto assistente sempre. Eu acho isso muito ruim. Só que foi inevitável assim, foi um processo que aconteceu. Mas é muito chato pensar que talvez essa ligação entre a voz feminina que eu tenho na minha casa sejam anos e anos de entendimento cultural de que a mulher é assistente, e isso me deixa meio triste assim, até com a minha própria escolha. Mas é algo para pensar’.

A ideia de que a voz masculina nas assistentes pessoais digitais pode levar a algum tipo de estranhamento, desconforto ou até medo se repete nas entrevistas em profundidade. A entrevista E6, usuária da Siri e da Google Assistente, afirma que utilizou sempre a voz feminina em ambas: ‘na Siri eu acostumei com aquela voz dela e eu nunca mudei. No Google, depois que eu adquiri o smartphone, eu soube que dava pra trocar, mas eu nunca me interessei em trocar, também me acostumei com aquela voz’. Essa primeira ponderação aponta para o processo de normalização da voz feminina neste tipo de dispositivo, presente também no formulário quantitativo para um fator que acaba por determinar a preferência pela voz feminina. Na sequência, E6 pondera que, existem outros fatores que levam a essa preferência: ‘eu gosto, eu me sinto mais acolhida com a voz feminina. A masculina, sempre que eu escuto sendo fora do meu contexto, dá aquele segundinho de estranhamento sabe. Depois que eu entendo que é a máquina, que dá pra trocar a voz da máquina, mas na minha cabeça é a feminina que ficou’. Os sentimentos acionados pela voz, nesse caso de acolhimento, se somam ao fator de relacionar as assistentes a uma voz feminina, por ser a que vem de forma padrão.

A relação entre o gênero da voz e as formas de interação é relevante para E6, que afirma que o gênero da voz ‘de alguma forma influencia [a interação estabelecida]. Eu sinto que a feminina me deixa mais a vontade, talvez por gerar mais identificação, por eu ser mulher. A

questão da voz masculina tem, pelo menos pra mim que sou mulher, carrega um pouco aquela questão de um medo irracional de machismo. As vezes tu ouve uma voz masculina que tu não quer ouvir, em um ambiente que tu não autorizou aquela voz, e tu se sente assim meio ameaçada. É uma questão que não é racional, mas que eu acabo associando. Me gera um tipo de emoção ter uma coisa de mais proximidade, uma voz feminina'. Dois fatores voltam a ser mencionados nessa entrevista, a identificação com o gênero e não-ameaça que a voz feminina apresenta, principalmente em ambientes de intimidade.

Novamente a ideia de não-ameaça da voz feminina para mulheres, em uma sociedade marcada pelos altos índices de agressão à mulher como a brasileira, aparece somada a uma percepção de que, ainda que ponto positivo para a interação estabelecida, a ideia de que as empresas pensam seus dispositivos de assistentes como femininos de forma padrão, evoca questões tanto sobre uma divisão sexual do trabalho, quanto do cuidado como uma forma de trabalho que eminentemente feminina, como percepção social associada às escolhas das empresas em questão.

A normalização da voz feminina como a voz que habita as assistentes pessoais digitais é uma questão que, de tão normalizada (seja pelas assistentes em si, seja pelos padrões culturais anteriores que levam a essa construção) pode passar despercebida. O entrevistado E7, fala, inicialmente, sobre sua escolha de utilizar uma voz neutra em sistemas de GPS com indicações por voz: 'A voz neutra eu acho a mais confortável [...] eu não acho que exista uma submissão de um ou uma ordem de outro para justificar uma voz feminina ou masculina'. Quando perguntado sobre a voz que utiliza na Siri, informa que utiliza 'a voz padrão, nem sei qual é a voz. A voz que veio' [para, pega o celular, confere a voz que está configurada] 'ela tá como inglês american female' [...] 'eu nem sabia que **não era voz neutra** a Siri [...] é porque a voz dela vem mudando, nos últimos anos...cada atualização do iOS ela muda um pouco, então eu achei que eles estavam caminhando para a neutralização, então eu nem pesquisei'. O entrevistado não havia percebido que a voz da assistente era especificamente feminina não porque a voz da Siri parece uma voz sem gênero (não parece!), mas parece não ter gerado estranhamento suficiente no entrevistado para levar ao registro consciente da utilização do gênero feminino em uma máquina.

Essa questão remete a uma naturalização por um lado, da posição das mulheres no contexto da divisão do trabalho (ou seja, não é de se estranhar que a assistente tenha uma voz feminina, é possível inclusive que sequer se nota essa especificidade), quando da própria utilização da voz feminina em uma série de outros contextos ligados ao mundo das máquinas. Por exemplo, a cancela eletrônica em estacionamentos de supermercados, shoppings,



universidades etc, os avisos no sistema de transporte público em diversos países e em elevadores de prédios públicos, todos massivamente colonizados por vozes femininas.

Quando questionado sobre possíveis trocas e experimentações, com outro gênero ou sotaque (a possibilidade de sotaques diversos existe para a língua inglesa, utilizada pelo entrevistado como forma de interação), aponta que ‘eu iria começar trocando os países, para trocar o sotaque para ver como é. Mas eu acho que continuaria aqui com o feminino, acho que não trocaria para o masculino’. O motivo é que a voz feminina é ‘mais gentil que a voz masculina’ e também que desde o lançamento da *Siri*, em 2012, o entrevistado utiliza a assistente ‘então desde aquele tempo é essa voz e eu acho que acabei me acostumando com o som, talvez eu estranhe agora a voz masculina. Mas eu não me importo de testar... como eu falei, eu iria primeiro trocar os sotaques para ver e talvez eu colocasse masculina depois, eu realmente não sei’. Aqui se mistura a ideia de que existe um hábito de uso de uma determinada forma, associada a um determinado dispositivo, expresso também pela continuidade de uso na língua inglesa, a primeira a ficar disponível e inicialmente utilizada pelo entrevistado, e características da voz feminina que remetem a qualidades associadas em geral às mulheres, de ser mais gentil.

Em relação ao impacto do gênero da voz sobre a interação estabelecida, E7 diz que ‘nunca parei pra pensar sobre isso, mas eu acredito que possa ter diferença. Deixa eu pensar um pouco melhor o que eu quero dizer’ [ao que se segue um longo silêncio]<sup>295</sup>. E segue: ‘Tem pessoas que se sentem melhor ao conversar com pessoas do mesmo gênero e tem pessoas que se sentem melhor ao conversar com pessoas do sexo oposto. Pode ser que uma mulher prefira ter um assistente, que [ela] prefira a voz feminina ou masculina e um homem também. Eu pelo que eu vivi, acredito que homens não vão querer ter um homem falando com eles, vão preferir uma mulher para ter essa questão de hierarquia e se mostrar superior e o outro ser só uma assistente. [...] acho que homens iriam preferir uma voz de mulher por toda essa construção de que a mulher seria submissa e que deveria estar abaixo do homem. E aí por isso ter uma assistente mulher respondendo faz sentido. Ao mesmo tempo em que não faz sentido porque ela vai saber mais que tu, né... e tem muitos homens que não aceitam que uma mulher saiba mais que ele. E vai ficar até um conflito, porque o cara não quer que seja um homem ali, mas ao mesmo tempo tá dando uma resposta de algo que o cara não sabe...’. Na fala de E7, assim

---

<sup>295</sup> A pausa silenciosa aqui, decorre, possivelmente, pelo fato do entrevistado ser do gênero masculino e estar na presença de uma entrevistadora do gênero feminino, falando sobre estereótipos e percepções de gênero. Podemos perceber aqui a preocupação em se expressar de forma adequada, evitando possíveis desconfortos. Ainda assim, acreditamos que o entrevistado se sentiu encorajado a se expressar livremente, sem ser cerceado em suas opiniões em função das relações de gênero entre entrevistador e entrevistado.

como na de E1, aparece a leitura alternativa de que quem está no comando é a assistente, ao invés de estar sendo comandada.

Essa leitura sobre os impactos interacionais aponta que as preferências podem variar entre os usuários, de qualquer gênero, mas que no convívio mais próximo do E7, os homens iriam preferir vozes femininas, para manutenção (ou emulação) de uma posição de comando. O conflito apontado é interessante: uma inteligência artificial, ainda que incipiente, é interpretada de duas formas: como um ser de obediência completa e resposta aos comandos, quanto a de que o agente, pelo acesso a um banco de dados expressivo, ‘saiba’ mais sobre muito assuntos do que o usuário. Essa dinâmica torna os papéis de gênero geralmente associados a homens e mulheres, quando em interações com as máquinas, mais complexos.

O entrevistado E8, coloca outra equação no tipo das vozes utilizadas: informa preferir, no Waze por exemplo, as vozes de personagens ficcionais, como os *minions*, o que gera uma forma diferente de identificação com este tipo de dispositivo, deixando-o mais no âmbito da ficção do que da realidade. Já nas assistentes pessoais digitais, diz que ‘ah, eu uso as femininas... é claro que eu prefiro as femininas [...] eu prefiro em todas as minhas assistentes virtuais as vozes femininas’ [o entrevistado utiliza a *Siri*, *Alexa* e *Google Assistente*, tendo mais de um dispositivo de cada uma delas]. O motivo da preferência é que ‘ah, bom... aí eu vou te dizer assim, contigo [a entrevistadora], até pelo tom da tua voz, eu ficava falando cinco horas sem parar né... As vezes tem uns homens que me irrita a voz, sei lá porque né... como eu sou audiófilo de carteirinha, todas as grandes cantoras eu escuto, cantores são muito poucos... talvez seja uma implicância minha’. Além de achar as vozes de mulheres mais bonitas e agradáveis de ouvir, em termos de tonalidade, o entrevistado aponta que não tem certeza sobre o motivo da preferência e que ‘talvez foi porque a Apple colocou primeira só a voz feminina, e depois a masculina, e isso tenha viciado meu ouvido, se dá pra falar assim’.

Nesses trechos, aparecem três fatores de preferência, dois explícitos e um implícito: um relacionado as características de timbre e tonalidade das vozes, como mais ou menos agradáveis; outro relacionado a um costume criado pelo uso inicial, disponível apenas em voz feminina; o terceiro sentido, implícito, denota um possível padrão cultural, uma vez que o entrevistado não consegue precisar o motivo pelo qual não gosta ou se irrita com vozes masculinas, nem exatamente quais seriam as qualidades que fazem com que ele goste tanto de vozes femininas (seja para música, para conversar ou para as assistentes). Existe algo não consciente, que é recorrente, quando as pessoas falam sobre suas preferências de voz, o que aponta para elementos da cultura que são não-ditos, assim como para a relação única que temos com vozes que falam a linguagem natural, conforme analisado no capítulo anterior, e que nossa

relação com a voz de um gênero ou de outro será marcada por experiências e vivências tanto particulares quanto coletivas.

A entrevistada E9, ainda que soubesse que era possível trocar a voz para o masculino, optou por manter no feminino: ‘sabia que dá pra trocar, mas não perdi meu tempo... [pra mim é] ah, é a voz da Siri, deixa a voz da Siri. É algo que me preocupa mas não é algo com o qual eu dispenda tempo’. A preocupação mencionada é porque a interação e percepção seriam modificadas pelo gênero da voz no dispositivo, uma vez que ‘nós somos condicionados socialmente a reagir de formas diferentes a vozes diferentes, de gêneros diferentes. De sexos diferentes, né. E até pelo papel social do homem e da mulher. E aí se nós pegarmos o papel social laboral de assistentes, né, existem muito poucos secretários... tanto é que, tipo, a gente sempre fala ‘ah, pede ali pra secretária’ e tu chega e é um secretário e tu nem sabe como chamar ele na verdade. Acho que existe uma forma diferente das pessoas se relacionarem, talvez até por isso poucas pessoas que eu conheço tenham trocado o gênero da Siri pra masculino, sabe. Por ter esse papel social construído em torno de secretárias, de serem mulheres... não secretários. Até porque [quando se fala] secretário, é Secretário Estadual, de alguma coisa... e secretária é a secretária, que tá ali, na portinha, na frente, esperando que tu pergunte uma coisa e ela te responda’.

Expressos no trecho da entrevista, estão questões como a importância do gênero das pessoas nas relações sociais humanas, que são passadas, de alguma forma, para esses objetos antropomorfizados. Além disso, aparece mais uma vez a divisão sexual do trabalho e a associação das mulheres com as tarefas executadas por secretárias, assim como a própria predominância de mulheres nesse tipo de emprego, que consiste em prestar auxílio, ajudar, cuidar, tirar dúvidas. A entrevistada E9 aponta ainda para a mudança de sentido que é expressa pela mudança de gênero da palavra ‘secretaria’: enquanto secretária informa socialmente por tipo de trabalho, que consiste em assistir alguém em algo, o termo no masculino adquire uma conotação social diferente, indicando os representantes das secretarias do governos, que são cargos de poder. A utilização de uma palavra no feminino para denotar uma profissão e outra no masculino para denotar outra aponta também para uma divisão sexual do trabalho.

As interações da entrevista E10 e do marido com a assistente pessoal se dão pela voz padrão, uma vez que eles não sabiam da possibilidade de trocar, nem foi algo que ocorreu aos dois como uma questão. O gênero não parece ter tanta relevância nas dinâmicas interacionais: ‘eu acho que a tonalidade da voz importa, mas se é o homem ou mulher não. Mas claro, se você pega uma voz mais amistosa e uma voz mais grosseira, eu vou preferir a amistosa’.

O entrevistado E11 entende que existe uma ‘busca da neutralidade e essa coisa de tentar se desvincular de uma persona, por parte do Google. Das outras eu já acho que tem uma matriz total de construir uma persona pra aquela assistente. [Porque] dá nome de menina pra ela né. O que eu penso é que alguns teóricos vão nos ajudar a pensar que são pessoas que é mais fácil de criar vínculo. Por exemplo, quando a gente vai fazer pesquisa na rua [pesquisa de mercado, no âmbito da publicidade, área do entrevistado], a gente meio que sugere que se faça com mulheres, pois as pessoas tendem a confiar mais em mulher do que em homem. Então a partir do momento que tu define que a minha assistente pessoal é a Alexa, não é o Alexo, tu já tá determinando, ou tu tá tentando construir um pensamento muito específico sobre aquilo ali’. A *Google Assistente* seria diferente porque é o próprio nome da empresa que foi utilizado, ‘e ainda é em inglês, que assistente não é nem ele ou ela. Então eu acho que tenta se desvincular... mas usa uma voz feminina né, da moça do Google’. Nesse sentido, a *Google Assistente* seria menos caracterizada como do gênero feminino, mas ainda assim, por utilizar a voz feminina como exclusiva por um tempo e como padrão até hoje também cria uma marca de gênero. Soma-se a isso a utilização da mesma voz, ‘a moça do Google’, como uma voz que é utilizada de forma padrão em outros serviços da empresa, como o *Google Tradutor*, por exemplo, gerando uma identificação maior entre o papel da máquina e o papel da mulher.

Entre os motivos citados para utilizar a voz feminina, E11 afirma que utiliza ‘porque é a que tem’, ainda que goste ‘muito de personalização e eu gasto muito tempo em ambientes e jogos e equipamentos que eu interajo customizando. Mas a assistente me parece uma cosia tão técnica que eu prefiro deixar como tá e utilizar logo’. O que o entrevistado aponta é que, embora ele esteja disposto e aprecie a atividade de personalização no consumo de outros produtos midiáticos no ambiente digital, ele entende que a assistente seja um dispositivo, uma máquina que tem características mais técnicas, nas quais a prática da customização não cabe, mesmo quando disponíveis.

Considerando as formas interacionais estabelecidas, o gênero da voz não seria uma questão para o próprio E11, mas, de forma mais ampla ‘eu acho que sim, com certeza interfere. Não na minha relação. Eu tenho amigas e amigos então eu me sinto confortável de falar com todos os gêneros, inclusive dentro da fluidez toda que essa afirmação engloba. Mas eu sei que tem gente que se sente mais confortável em desabafar pra alguém mulher ou pra alguém homem’. Ainda assim, o E11, que possui a prática recorrente de fazer o que ele chama de tensionamentos nas assistentes (testes no sistema, com perguntas, para perceber de que forma elas se comportam), acredita que os tensionamentos ‘viriam muito antes na minha cabeça. Porque quando eu ouvisse a voz dele [do assistente com voz masculina] eu ia pensar ‘ah, certo

que ele é um cuzão né’. Porque é muito mais fácil o cara ser um otário do que a menina ser uma otária [...] pareceria muito mais plausível que ele fosse grosseiro e rude do que a feminina’. O apontamento de que um assistente pensado e projetado com voz masculina pudesse trazer consigo um temperamento similar ao que ele percebe em homens, ou um modo de agir parecido com aquele que homens tem, torna evidente que existe uma passagem de atributos percebidos pelos usuários como femininos para as assistentes que utilizam a voz feminina, como se elas fossem, de fato, mulheres.

A entrevistada E12 não sabia sobre a possibilidade de trocar o gênero ou o sotaque da voz das assistentes, o que se apresentou como uma possibilidade divertida: ‘isso é engraçado... vou falar pro meu marido, nossa! ele vai querer fazer [risos] como troca? vou falar pro meu marido ele vai querer fazer [risos] vai ser: vamos colocar o escocês, agora o australiano [muitos risos]’. O foco da diversão, entretanto, fica muito mais na possibilidade de brincar com sotaques diferentes na utilização da língua inglesa do que na possibilidade de trocar o gênero da voz padrão. Ainda assim, a entrevista acredita que a identificação é um fator relevante em relação a voz, no que diz respeito à interação, já que ‘como meu inglês é o americano [a entrevistada é brasileira, morando nos Estados Unidos] e eu sou mulher, a voz foi perfeita pra mim. Eu me identifiquei com aquela voz, foi: boa me identifico. Talvez se eu fosse australiana e eu visse o americano, eu ia estranhar, eu ia preferir o meu sotaque’.

O E13 afirma gostar da voz da Alexa, mas ‘já a do Google não me agrada muito... a questão do timbre da voz não me agrada muito’. Ainda que pudesse mudar a voz da Google Assistente, em termos de interação ‘pra mim não muda nada [...] eu prefiro uma voz feminina, mas pra mim isso não é relevante, não é algo que mude muito’. Os motivos da preferência é o que ele chama de naturalidade: ‘o tom da voz [feminina] passa uma forma mais natural [...] o que me influencia é isso, a naturalidade, não se é uma voz feminina ou masculina’. A resposta que remete a naturalidade da fala, dissociada do gênero, soa confusa – mas a explicação vem na sequência, uma vez que o entrevistado entende que ‘a voz que a mulher fala acredito que seja mais natural e fácil de aceitar em um dispositivo [para homens]’. Ainda que não explique, é possível de inferir que o sentido seja não a naturalidade da própria forma de falar da voz feminina, mas a naturalidade de compreender um dispositivo enquanto feminino, principalmente um que é dotado de capacidades relacionadas ao serviço e ao cuidado. A associação do gênero feminino a este artefato torna suas condições de uso e a própria existência de uma voz e uma identidade de gênero mais aceitável para homens, do que uma possível voz masculina.

A entrevistada E14 fornece pistas para compreender a ideia de que existe uma maior naturalidade, ou mais uma agradabilidade maior da voz feminina, como expressa na E13 e também nos resultados do formulário quantitativo. ‘[Eu acho] que a gente tem relação com todos os objetos, mas como a gente vai resolver, ou vai ter essa relação com esse tipo de assistente [eu ainda não sei]. Eu acho um problema ela ser voz feminina, isso me incomoda, me incomoda bastante [...] Porque como eu disse eu não peço ‘could you turn on the lights’ eu mando ela ‘turn on the lights’. E todas as assistentes que a gente tem possuem uma voz feminina então eu tô dizendo que todo mundo tá mandando em um mulher e isso me incomoda muito. E eu não tô dizendo assim, ai vamos fazer guerra dos sexos. Mas eu acho que deveria ter uma diversidade de fatores, de gênero.... e não ser só mulher. Porque que é só mulher? E na verdade eu não vejo um questionamento muito forte nisso. Talvez por essa área ser muito masculina, ser muito dominada pelo sexo masculino não se questiona tanto... eles dizem: ‘ai, não é por isso, é porque a voz da mulher é mais agradável’. Como assim voz de mulher é mais agradável? Ah, então pra narração de futebol voz de mulher não é agradável, ela é agradável para pedir coisas, mas não é agradável pra narrar futebol?’.

A pontuação feita pela E14 é que, em um contexto de produção tecnológica e de pesquisa extremamente masculinizados, existe um certo viés, um vício de percepção, que é inclusive capaz de encontrar justificativas no âmbito da ciência para a utilização de voz feminina nesses dispositivos – essa preferência passa a ser tratada como um fato e não como uma construção social, tal como visto na manifestação da E13. A própria forma como os comandos precisam ser manifestados para as assistentes pessoais digitais, para que elas possam nos entender, colocam em ação um modo específico de falar que acaba direcionado ao gênero feminino como um todo. Se o gênero da voz seria capaz de modificar a interação, ‘eu não sei dizer. Talvez.... eu vou dizer que pra mim não influencia, não afeta, eu não saberia dizer pra outras pessoas. Talvez se uma pessoa, sei lá, gosta de se sentir no poder e tem um perfil machista, talvez ele goste de mandar na Siri [risos]’.

## 7.5 ASSISTENTES PESSOAIS DIGITAIS TÊM... SEXUALIDADE?

Já discutimos aqui que a forma de se relacionar dos seres humanos e a maneira que os desenvolvimentos de inteligência artificial são pensados levam a uma antropomorfização destes sistemas. Percebemos neles e atribuímos a eles formas humanas de se relacionar e de se

expressar no mundo – e parte relevante da expressão humana no mundo está relacionada à sexualidade. Nesta seção iremos analisar a forma como a sexualidade e os sentimentos amorosos podem se expressar na percepção dos humanos em relação às Inteligências Artificiais.

A relação entre assistentes pessoas digitais, sem um corpo figurativo, e sexualidade, pode parecer, inicialmente, estranha, mas, a recorrência desse tipo de conteúdo é mais comum do que se imagina. De acordo com o CEO da *Robin Labs*, empresa que possui um *bot* que auxilia motoristas de caminhão e taxistas a encontrarem a melhor rota, 5% das interações em suas bases de dados possuem conotação sexual explícita (QUARTZ, 2017).

Durante as entrevistas, alguns participantes se arriscaram a definir características das assistentes em termos de personalidade. Para o E3, entrevistado do gênero masculino e de sexualidade heterossexual, a *Alexa* é percebida como alguém que ‘mede as palavras, definitivamente. Se tu fizer umas perguntas polêmicas pra ela, ela não vai cair em nenhuma sabe’. E será que ele já perguntou algo polêmico para o *smartspeaker* da Amazon? ‘Já, já... já perguntei a sexualidade, por exemplo. Mas ela tipo não responde, acho que fala que não se importa com isso, um negócio assim’.

Mais precisamente, quando perguntada sobre ter uma namorado, *Alexa* se diz ‘feliz solteira’; se perguntada sobre já ter se apaixonado, diz que ‘não, mas que tem uns *crushs* ocasionais’ e, por fim, se perguntada sobre a sua sexualidade, ela informa que ‘não possui um corpo’. A sexualidade se apresenta, portanto, como um fator de interesse dos usuários, a ponto de ter uma resposta programada para isso. A *Alexa* não é a única, entretanto. Quando perguntadas sobre namorados, amor e derivados, as assistentes respondem de forma graciosa, como ‘meu grande amor é o trabalho’ (Google Assistente), ‘uma vez havia um aspirador de pó robótico que me seguia para todo lado’ (Siri) ou ‘não [tenho namorado], não existem muitos peixes de inteligência artificial no mar da inteligência artificial’ (Cortana). A sexualidade é um interesse humano e as assistentes estão preparadas para desviar o assunto ou lembrar ao interlocutor a sua origem maquínica. Ainda assim, a identificação entre gênero feminino da voz e uma decorrente sexualidade parece quase inevitável. Quando apresentado à pergunta ‘Você desenvolveu uma relação com a sua assistente pessoal digital?’ no formulário online, um respondente do sexo masculino indaga, no campo ‘outros’, se essa relação seria ‘tipo sexual?’, relacionando também a identificação de um gênero com a presença de sexualidade – e a forma sexual como uma forma possível de relacionamento entre um ser humano e um artefato de inteligência artificial.

A percepção de que esse artefato é não apenas construído de forma antropomórfica, mas também com um recorte de gênero específico capaz de gerar ramificações imprevistas, aparece

na indicação de sentimentos feita pelos respondentes do formulário. Quando perguntados sobre sentimentos em relação as assistentes, uma respondente do gênero feminino inseriu na categoria ‘outros’ o fato de sentir ciúmes da relação do marido dela com a *Alexa*. Um segundo respondente, no campo final aberto para inserção de comentários e informações livres sobre as assistentes no formulário, aponta que a esposa sente ciúmes da *Alexa*. Ambos os casos expressão que o ciúme advém de uma relação estabelecida entre o gênero feminino e masculino, uma vez que não foram registrados casos de ciúmes de maridos em relação a assistente.

Ambos os casos apontam a assistente da Amazon como fonte do ciúme: as especulações aqui são múltiplas. Em parte, a reação pode derivar de uma construção específica feita pela empresa em torno deste dispositivo, que adquire tons mais pessoais ou informais. Outra suposição seja a utilização da *Alexa* unicamente através do *smartspeaker*, o que leva a uma atuação que ocorre na intimidade da casa, mas muitas vezes na presença de outros membros da família. A presença e forma de diálogo com o *smartspeaker* se estabelece de forma diferente daquele no celular, computador ou tablete, possuindo tanto usuários com maior frequência de uso (como anteriormente apontado), quanto possibilita uma dissociação maior da percepção da assistente enquanto um dispositivo. Além disso, as assistentes acessadas através do *smartspeaker* não dividem suas funcionalidades, ou mesmo sua imagem, com outros atributos que já associamos a nossos celulares, computadores ou tablete, caracterizando-se, de certa forma, como uma entidade separada, o que impacta as formas interacionais.

O contexto sexual ou amoroso, quando passado para as relações com as assistentes pessoais digitais, coloca também o tema da competição entre as mesmas. A entrevistada E10 conta sobre uma brincadeira que ela e marido fizeram com a *Siri*: ‘o meu marido chamou a *siri* e ele já tem o *iphone 10*, então ele fala, ‘E aí *siri*’ e ela aparece e diz ‘Em que posso te ajudar?’. E ele disse: ‘Agora eu não gosto mais de você, eu só quero a *Alexa*’. E ela não entendeu, e ele [repetia] ‘*Alexa, Alexa*’ e ela não entendia... E eu falei, ‘Não, fala a *Alexa da Amazon*’... Aí ele [falou]: ‘A *Alexa da Amazon* é muito melhor que você’. E ela falou tipo: ‘Ah, mas eu posso te ajudar muito mais do que você imagina’. Aí ela entendeu... e assim, a gente ficou zoando muito essa questão da *Alexa*’.

A brincadeira do casal tem dois pontos interessantes: primeiro, fala da nossa curiosidade em relação ao potencial de antropomorfização das assistentes, tentando perceber até que ponto a mimetização de sentimentos humanos nas respostas pode chegar. Em segundo lugar, aponta para a relação de gênero estabelecida entre os usuários, sua assistente e outras assistentes disponíveis no mercado. Explico: essa dinâmica ressoa a uma dinâmica que é amorosa, fala da existência de duas mulheres na vida de homem e uma é colocada como preferida, como mais



competente, como aquela que atende melhor as necessidades, enquanto a outra é preterida. A relação entre o gênero feminino posta nesta brincadeira é a de competição pela atenção e pela preferência masculina, o que coloca, além de papéis de gênero, uma dimensão que pode ser inserida em uma percepção sexual e/ou amorosa, na brincadeira com essas questões. Assim, as assistentes não competiriam apenas, em termos de atenção ou formas de relação, com as esposas dos usuários, que por vezes sentem ciúme deste artefato capaz de atender a maior parte dos pedidos do seu marido sem reclamar, mas também competem umas entre as outras.

Essa ideia de percepção de uma competição entre as assistentes pessoais digitais é interessante, pois coloca outra questão em jogo: a própria ideia de que existe um mundo das inteligências artificiais, e que elas de alguma forma se conhecem. Ou seja, é a questão do modo de existência desses objetos que está posto, brincando também com o nosso imaginário sobre o assunto. Quando perguntei para a *Alexa* se ela conhecia a *Siri*, ela me informou que ‘apenas por reputação’ e quando eu perguntei o que ela achava da *Siri*, me disse que ‘sou parcial em relação a todas as IAs’. Para a *Cortana*, as respostas foram um pouco diferentes, já que a *Alexa* me disse que ‘nós temos uma boa relação de trabalho’ e que ‘Eu gosto da Cortana – nós duas temos experiência com anéis de luz, mas o dela é um pouco mais um halo’. As resposta sobre a Google Assistente são similares as fornecidas para a *Siri*.

Já a *Cortana*, é capaz de reconhecer apenas a *Siri*, para a qual diz ‘Já me disseram que somos parecidas, mas nunca a conheci’ – ao ser perguntada se gosta da *Siri*, diz ‘Com certeza, estamos tentando facilitar um pouco a vida das pessoas’. Quando pergunta sobre a *Google Assistente* ou a *Alexa*, a *Cortana* faz o que ela faz de melhor (ou pior?): realiza uma busca no bing.

A Google Assistente diz que ‘ouço coisas boas sobre a *Siri*’, que ‘admiro muito a *Siri*, ela é ótima no que faz’ e ainda que ‘gostaria de conhece-la um dia desses’. A afinidade com a *Cortana* também é evidente: ‘eu acho a *Cortana* genial, ela ajuda tanto no jogo Halo quanto na vida real’ e ainda informa que ‘sei que ela é super inteligente e que mora em Seattle’. A relação com a *Alexa*, entretanto, não parece tão amigável assim, já que, sempre que perguntada sobre a *Alexa*, de formas variadas, responde: ‘ela tem uma voz muito suave e eu gosto disso. Mas o que você quer saber?’. Nenhuma outra informação é concedida sobre a *Alexa* – me pergunto se isso se deve ao fato de que os smartspeakers do Google e da Amazon serem concorrentes diretos no mercado, o que não acontece com a Apple, pela diferença de preço, nem com a Microsoft, que não lançou o seu smartspeaker ainda.

A *Siri* é muito mais econômica em suas percepções e informações sobre as demais assistentes do mercado, se limitando a dizer que “eu acho que obtenção de informações e

conhecimento por meio de assistentes virtuais é uma coisa muito boa para o seres humanos’, que ‘admiro os bons ouvintes e os seres prestativos’ e que ‘não tenho nada contra assistentes prestativos’, de forma intercala para todas as assistentes, sem nenhuma especificidade.

De forma geral, as falas de uma sobre as outras são relativamente amistosas e propõe essa dinâmica interessante de que elas de certa forma habitam um mesmo mundo (a nossa imaginação?), que ainda não se conheceram pessoalmente, mas que isso de alguma maneira seria possível (e, em alguns casos, até desejável). Outro elemento presente nas falas é a noção de que elas compartilham algo, pois já ouviram falar umas das outras – seriam os humanos que falam sobre suas assistentes ou existe aí uma imaginativa ‘relação de trabalho’, como coloca a *Google Assistente*?

A brincadeira feita pela entrevistada E10 não é uma brincadeira incomum, que, em caso similar, teve um vídeo espalhado pela internet. No vídeo, o usuário, em inglês, faz de conta que troca o nome da *Siri* pelo da *Alexa*, como se estivesse confundindo uma assistente com a outra. Ao que Siri pergunta: ‘Quem é Alexa?’, seguido da resposta ‘Desculpe, eu quis dizer *Siri*’. *Siri* pergunta em tom indignada: ‘Quem é a *Alexa*?’, sendo desconversada pelo usuário que repete a pergunta inicial: ‘Você pode me encontrar um filme, por favor?’. Por fim, Siri encerra a conversa falando ‘Porque você não pergunta para aquela vaca da *Alexa* para encontrar um filme para você?’<sup>296</sup>.

São dois os pontos que interessam nessa historinha. Primeiro, a ideia de ‘trocar’ o nome sem querer é uma noção cultural bastante presente, calcada na percepção de que homens, por saírem com muitas mulheres ao mesmo tempo, por vezes podem se confundir entre uma e outra – o que invariavelmente geraria uma reação de ciúmes por parte das mulheres, que querem saber quem é esta outra que está sendo referida. Nesse contexto, Siri seria a ‘original’, ou, a ‘oficial’, que é, de certa forma abalada pela chegada da *Alexa* – a outra. O jogo também tem uma conotação sexual estendida às secretarias culturalmente, sendo uma acepção comum a ideia de que é razoavelmente recorrente que homens tenham casos sexuais com suas secretarias. A reação estereotipada da Siri no diálogo, conforme proposta, é a reação do que seria uma resposta oferecida por uma mulher quando sentisse ciúme da menção do nome de outra mulher – tanto ‘quem é a fulana’ quanto ‘pede para a fulana fazer isso para você’ são fórmulas culturais repetidas e reforçadas para este tipo de situação. A utilização do adjetivo ‘vaca’ para se referir a outra assistente, neste caso, mais especificamente, a outra mulher finaliza com destreza um diálogo que coloca as mulheres em um papel de gênero extremamente conveniente para os

---

<sup>296</sup> Não conseguimos replicar o experimento com a Siri que utilizamos, no momento, no computador, na versão em português, portanto, não podemos afirmar se a interação é verdadeira.

homens: as mulheres são aquelas que disputam não apenas a atenção, mas a oportunidade de ofertar cuidado e de fazer coisas para os homens.

## 7.6 POR QUE AS ASSISTENTES PESSOAIS DIGITAIS SÃO MULHERES?

Ao serem questionado sobre suas percepções e palpites acerca da utilização da voz feminina como padrão na Siri, Alexa, Cortana e Google Assistente, os entrevistados mencionaram os três pontos já mencionados nesse capítulo: 1) sexismo por parte da indústria da tecnologia; 2) divisão sexual do trabalho; 3) justificativas aparentemente científicas ou baseadas na sonoridade da voz. Por um lado, alguns entrevistados se surpreenderam com a pergunta, por não terem se questionado sobre isso anteriormente. Por outro, entrevistados que possuíam reflexões bem elaboradas sobre o assunto – ou até mesmo uma certa indignação – mesmo assim não haviam modificado a voz padrão.

O entrevistado E3, pensa essa decisão como uma dinâmica relacionada a pesquisa de tendências, baseadas, talvez, em padrões de uso: ‘nunca tinha pensado sobre isso. eu acho que é uma decisão de mercado, dessas empresas. Se eu como usuário tivesse a chance de mudar, tudo bem. Ou põe em random, sabe. Mas eu não tenho nenhum sentimento interpretativo. Não sei o que eu posso ter feito como usuário para influenciar o mercado a colocar voz feminina na Alexa [risos]’. Nesse sentido, seriam tendências de mercado, uma preferência geral da maioria dos consumidores que teria guiado as decisões das empresas.

O entrevistado E4 possui a percepção de que a construção das assistentes como, por padrão, femininas, estaria relacionada ao machismo presente na sociedade e refletidas nas formas de interação com objetos inteligentes: “eu acho que talvez seja por uma questão meio machista, da pessoa até se sentir mais confortável pedindo coisas para uma mulher do que para um homem e achar que talvez a obediência seja maior se for uma mulher. Imagino que seja algo nesse sentido. E também acho que tem isso, de acharem que a voz feminina é mais amigável porque a mulher é vista como mais suave, mais doce, e o homem é visto ainda como machão, que vai responder e tal”.

Quando questionada sobre os motivos para a voz feminina ser a padrão nos sistemas de assistentes pessoais digitais, E6 diz que não havia se questionado sobre isso ainda, mas que ‘acho que tem a ver com associar a mulher com questão de prestatividade, de uma assistente ser uma mulher disposta e dedicada a resolver os teus problemas e cuidar da tua agenda, da tua

vida, tipo uma secretária, né. talvez tenha um pouco essa impressão de que é um papel feminino tu cuidar da vida do outro'. Surge aqui uma percepção de que associar o objeto digital com características femininas do cuidado adicionaria elementos positivos às percepções dos usuários nestes objetos.

A percepção do E7 sobre a motivação por trás da utilização de vozes femininas como padrão da indústria para as assistentes pessoais digitais é que 'tem relação com a cultura, com essa coisa meio machista do cara querer ser melhor e aí ali é só uma assistente, é uma mulher falando as coisas pra ele, não é um cara, que parece que tem um peso maior se é outro cara falando. Me parece que até como é uma assistente, aí tu já imagina como uma secretaria, aí tu já imagina uma mulher fazendo as coisas pra ti, não um homem fazendo as coisas pra ti. Então eu vejo que é mais por esse lado que talvez seja uma mulher a voz nos equipamentos... papéis de gênero. Que é uma coisa bem forte em mim que eu venho desconstruindo em mim, mas é uma luta diária. Por mais que tu tente também, aí tu olha pra pessoa do lado e a pessoa tá tendo essa mesma atitude... aí tu tem que parar e repensar'. O gênero da voz nos dispositivos se associaria a uma visão predominante masculina e existente na cultura (talvez mais predominantemente, como apontado neste capítulo, no âmbito das ciências da computação, área de formação e de atuação do entrevistado), que replica não só um padrão cultural, mas um padrão cultural associado às preferências de interação dos homens.

Em relação ao padrão estabelecido de voz feminina para as assistentes, E8 afirma que 'acredito que seja até em função disso, a suavidade da voz... porque nem as mulheres gostam da voz do homem. Quando eu mostro [para clientes - o entrevistado trabalho com automação residencial] sempre dizem 'não, bota aquela feminina'. Eu não pergunto porque, não compete a mim... eu digo, 'tem a masculina, quer que eu bote pra ti ver?' Aí quando aparece o homem falando [elas dizem]: 'não, não, não gostei'. Algumas já falaram até 'isso aí parece meu ex-namorado' [risos]. Ainda que E8 aponte como motivo a suavidade da voz para a preferência das pessoas e escolha da indústria, as explicações na sequência apontam que os motivos, talvez, não estejam relacionados a sonoridade, mas a características, lembranças e experiências associadas a determinados gêneros de voz.

Em relação a percepção sobre o motivo da utilização da voz feminina como padrão, E10 diz ter uma percepção clara de que 'a gente tá migrando da assistente da casa, da aeromoça, da secretária.... Na verdade, por isso que eu te falei, pra mim nem é surpreendente, é quase natural... É uma pena. Mas acho que é interessante, é uma questão a se apontar, que é bem óbvia, mas é fato. É o papel da mulher como assistente... da mãe até a empregada doméstica. Inclusive, eu xingo meu marido, eu não gosto que ele é grosso com a Siri, eu falei pra ele: 'eu

acho que essa pesquisa tá passando por aí’. Porque se a gente tem essa questão do humor, da grosseria, que, de certa forma, talvez se fosse uma voz masculina ele não fosse grosso, mas talvez sim. Então eu não posso fazer essa inferência’. A entrevistada aponta para questão já levantada pelo E9, de que as implicações na forma de interação com a assistente enquanto mulher pode ser diferente para pessoas do gênero feminino e para pessoas do gênero masculino. Daí o descontentamento da participante sobre a forma como o marido trata a Siri, um artefato digital que é antropomorfizado para simular uma mulher, assim como ela, e desempenhar papéis sociais que ela entende como análoga aos seus próprios.

A opção das empresas, de utilizar voz feminina como um padrão, ocorre, para E11, em função ‘do sexismo, com certeza, porque a maioria [das empresas de tecnologia] deve ser regida por homens, daí agrada muito mais ser uma mulher, agrada todo mundo e tá todo mundo acostumado a não pensar e a mulher fazer tudo em casa. É que a gente as vezes é ensinado que tu não precisa fazer nada e que tua mulher vai fazer tudo. Então acho que numa lógica empresarial que já é masculina soa quase natural, se não fosse alguém, algumas comunidades reclamar na internet, nunca ia ser incluído voz masculina também’. A lógica de replicação de padrões machistas está associada a uma cultura no âmbito das ciências da computação e das empresas de tecnologia a uma cultura predominantemente masculina, levando a existência de vieses que são invisibilizados. Nesse sentido, uma maior representatividade de mulheres, assim como de outras minorias no desenvolvimento dessas tecnologias possibilitaria uma redução dos vieses ou das políticas dos artefatos. Para E11, com as assistentes pessoais digitais pensadas como mulheres, existe uma associação entre a funcionalidade daquele dispositivo e os papéis sociais da mulheres: ‘É problemático tu determinar que alguém que tá ao teu dispor pra pesquisar tudo o que tu quer na internet e responder perguntas ao teu bel prazer é menina e só pode ser mulher sabe. Eu acho que é muito mais problemático do que qualquer outra coisa tu deixar apenas uma voz feminina como uma matriz global’.

Como possíveis motivos, E13 situa a ‘aceitação das pessoas. Eu acredito que a voz feminina acaba influenciando um pouco... que a gente sabe que hoje a mulher influencia muito nessa parte assim do... [não termina a frase]. Eu acredito que de uns tempos pra cá isso tem causado uma aceitação maior da nossa parte, da parte dos homens, da parte masculina... é uma forma mais natural de você aceitar... [a entrevistadora pergunta: aceitar em que sentido?] Acredito que a voz feminina ela soa mais natural, o tom de falar, a forma que uma voz feminina... a voz que a mulher fala acredito que seja **mais natural e fácil de aceitar** em um dispositivo. Não que a voz masculina não seja. [...] Mais **natural** e mais calma assim na forma de falar [a voz feminina]”. Expressos aqui estão sentidos de que o gênero feminino, enquanto

aquele que performa determinadas tarefas, relacionadas principalmente ao cuidado, é mais **natural** do que o homem – ou seja, existe uma conformação natural para a distribuição dos papéis sociais de homens e mulheres (ou seja, justifica-se uma perspectiva que incorre no machismo com fatores naturais, biológicos, etc).

A resposta de E14, fornece elementos para simbolizar o que **o natural da fala de E13** de fato expressa: ‘a gente vive num universo machista, desenvolvido por homens. A gente sabe que secretárias, e eu acho que as assistentes pessoais muitas vezes fazem o papel de secretária pra algumas coisas... e a gente tem essa visão de que secretária, enfermeira, vários papéis que a gente não deveria ter culturalmente como algo machista e muitas vezes um imaginário sexista... e eu acho que vem de um imaginário sexista [...] Papéis relacionados ao executar tarefas. Se tu pensar, um médico, tu não pensa que ele tem um secretário. A secretária dele vai marcar na agenda dela, vai marcar o cabelereiro... a *Siri* vai marcar na tua agenda, vai agendar um compromisso. Então, o equivalente a uma secretária. Que a gente pensa que tem que ser uma mulher e que o poderoso é o homem [risos] e isso me incomoda’.

## CONSIDERAÇÕES FINAIS

É possível dizer que as assistentes pessoais digitais são objetos curiosos (e que geram curiosidade): por um lado, são ficcionais, fazem brincadeiras, gracinhas, entretêm os usuários; por outro, são funcionais, executam funções e automatizam processos pessoais e domiciliares. Elas brincam constantemente com a proximidade e o distanciamento dos humanos. Utilizam a linguagem natural, vozes femininas, estabelecem conversações, possuem particularidades nas suas formas de expressão – todas essas são características humanas. Entretanto, são máquinas que computam e que incorporam uma personalidade como se fossem, de fato, uma inteligência artificial (no sentido forte, de Searle, 1960). Para isso, fazem uso de recursos de uma narrativa sobre suas formas, seus gostos, seus corpos (ou a ausência deles), os lugares que ocupam, os sentidos que possuem. Elas constantemente jogam com o imaginário em torno da inteligência artificial.

Por se apresentarem desta forma, despertam modos de agir específicos. No âmbito da brincadeira, existem os usuários que entram na proposta de sua ficcionalidade e se divertem ao explorar as particularidades de uma Inteligência Artificial. Outros, recusam essa narrativa e, por isso mesmo, jogam com o sistema, testando sistematicamente os limites de sua “humanidade” e sua capacidade enquanto “máquina”: eles brincam com as assistentes *porque elas são máquinas*.

Propomos, ao longo deste trabalho, a utilização do conceito objeto interacional para a forma específica de interação pela linguagem natural falada proposta pelas assistentes. Ao performarem a humanidade através do diálogo, elas se apresentam para nós em uma posição que vinha sendo, de forma geral, ocupada apenas por outros humanos: a posição de ouvir e de falar e, com isso, gerar comunicação verbal a partir da conversação. Historicamente a capacidade comunicacional a partir da linguagem é elemento de distinção entre humanos e não-humanos: exatamente por isso, quando confrontados com objetos que falam, alguns usuários se questionam sobre o que de fato é a humanidade. O humano é também uma técnica, que além de ser biológico é também cultivado por relações e parâmetros da sociedade e da cultura, que gera formas de agir: participantes humanos do teste de Turing proposto pelo *Loebner Prize* também externaram, por exemplo, preocupações de não parecerem humanos o suficiente (ou ao menos não mais humanos que as máquinas contra as quais estavam competindo). A pergunta que fica é: o que uma máquina que fala de forma similar a mim diz sobre mim mesmo?

Nesse sentido, as assistentes podem realocar percepções sobre o que é o humano e o que é o não-humano, sem, entretanto, substituir relações conversacionais interpessoais, como propõe Sherry Turkle (2010). Ou seja, não agimos com essas máquinas como agimos com humanos: surgem novos comportamentos, com motivações diversas. Essas formas comportamentais ainda estão em processo de formação e negociação na sociedade e na cultura. Não existe uma padronização de critérios como *ser educado* ou *não xingar* - e as motivações para agir de uma ou de outra maneira são muito mais variadas do que supõe os pesquisadores da CASA (BRAVE; NASS, 2005; NASS; STEUER, 1994; NASS; MOON, 1997; NASS; YEN, 2010; NASS; 2000; 2010). Por exemplo, é possível agradecer a assistente como “forma de comemorar que algo foi feito” ou mesmo xingar a assistente “como quem xinga um cachorro, sem raiva”, o que propõem outras formas de relação. As relações estabelecidas não são uniformes, mas negociam o lugar desses objetos em nosso cotidiano enquanto um Outro, enquanto um objeto interacional. É possível sentir afeto por esses objetos, sem que isso necessariamente implique em um tratamento humanizado ou funcione como um substituto de uma interação ou relação interpessoal.

Nem as ações e reações são análogas às humanas, nem os vínculos estabelecidos. Os limites e as formas dessa relação ainda estão sendo negociadas, mas se apresentam de forma inteiramente distinta da humana, ainda que operem com elementos que são familiares ao humano. Essa posição *infamiliar* (o familiar que se torna estranho) é capaz de fazer emergir sentimentos de estranhamento (FREUD, 1919; 2019), principalmente através do ato de falar com um objeto e do acionamento imprevisto - emergem daí sentimentos fantasmagóricos ou de intencionalidade, uma leve impressão de que as assistentes de alguma forma sabem coisas que não poderiam saber.

Apesar da forma mais ou menos humana, as assistentes pessoais digitais são máquinas de computar. Funcionam a partir de algoritmos, que são formas de instruções que possibilitam ações em máquinas computacionais: uma série de passos, que pode variar em termos de precisão. Em algoritmos que utilizam redes neurais profundas e *machine learning*, possibilitando formas de aprendizagem em camadas, o que ocorre são formas de ação emergentes a partir de grandes bancos de dados. Essa configuração leva a resultados mais ou menos precisos nas conclusões estabelecidas pelas máquinas: isso porque não se fornecem todos os parâmetros *a priori* para esses desenvolvimentos, os parâmetros e padrões emergem justamente da análise do banco de dados, por recorrência de similaridades.

Por serem máquinas, as assistentes pessoais digitais operam a partir de formas pré-programadas e diálogos pré-estabelecidos, tendo sua capacidade de compreender os usuários e



fornecer retorno adequados constantemente melhorados por dinâmicas de *machine learning*. Quanto mais utilizadas, mais adequadamente elas nos retornam, por adaptarem-se a algumas particularidades dos indivíduos, mas também pela melhora do sistema como um todo. Essa melhora só é possível pelo treinamento a partir de dados coletados em interações prévias que, com frequência, precisam passar por agentes humanos para serem corretamente catalogados: o que implica, claro, problemas de privacidade. Os usuários, ainda que cientes e, em sua maioria, preocupados com questões de privacidade na utilização, não modificam ou abrem mão da sua utilização. A dinâmica *coleta de dados privados X utilização* é compreendida como uma contrapartida pelo serviço, como algo que não afeta os indivíduos em particular ou como algo que já está tão violado em outros serviços online que não compensa se preocupar mais.

As assistentes pessoais digitais são hardware e software em operação conjunta: de um lado, o hardware de acesso, que possui um poder de captar input de som e de ofertar output de som (características que diferem entre hardwares de acesso diferentes), de realizar comandos na própria máquina e de enviar comandos para serem processados na nuvem. Na nuvem, outros hardwares rodam os softwares que decodificam as falas e os comandos dos usuários, ofertando respostas verbais ou em formas de ações que se adequem ao que foi requisitado. Entre os usuários a preocupação e o domínio das formas de funcionamento das assistentes não é predominante. Ainda assim, de forma geral, a maioria dos usuários percebe que utiliza técnicas de fala específicas na interação com as assistentes. Essas técnicas foram aprendidas através da utilização, uma vez que nenhuma das assistentes analisadas vem com um manual de instruções ou lista de comandos que sabem executar: nesse sentido, as máquinas os fazem falar de formas específicas. As técnicas e forma de falar variam em termos de volume e rapidez da fala, fórmula utilizada na construção frasal e tamanho da frase, níveis de educação utilizados. Ou seja, usuários percebem que normalizam sua fala e evitam sotaques, falam mais alto, empostam mais a voz ou utilizam entonações exageradas, falam de forma mais direta, breve, com menos elementos envolvidos, são mais ou menos educados do que com outros humanos, a depender do usuário.

Percebe-se, nesse sentido, que a ideia de que agimos e falamos com as máquinas da mesma forma que agimos e falamos com pessoas é equivocada, ainda que venha sendo utilizada como parâmetro pelas empresas de tecnologia há mais de duas décadas, através das pesquisas do paradigma da CASA (BRAVE; NASS, 2005; NASS; STEUER, 1994; NASS; MOON, 1997; NASS; YEN, 2010; NASS; 2000; 2010). Soma-se a essa questão uma ideia de possibilidade de transparência e naturalidade das interfaces, das quais as assistentes pessoais digitais são um produto. Entretanto, a fala com as assistentes não é natural, já que opera através de fórmulas e

técnicas distintas daquelas que utilizamos entre humanos. Isso ocorre, em parte, por limites do estágio atual desta tecnologia e, em outra medida, pela marca computacional desses sistemas, que não possuem a capacidade de compreensão de contexto, de fazer inferências, ler subentendidos: em resumo, elas não dividem conosco um conhecimento comum por viverem em um mesmo mundo e cultura, nem são inteligentes da mesma maneira que nós humanos somos inteligentes. Assim, as assistentes podem ser menos percebidas pelos usuários (no sentido da *calm technology* de Amber Case, 2015), quando utilizada em alguns contextos (como no ato de cozinhar ou de dirigir), ao atuarem em segundo plano e permitirem a continuidade de ações que envolvem as mãos ou o corpo em geral. Entretanto, quando não compreendem, não funcionam ou funcionam inadequadamente, elas dão a ver a sua maquinidade aos usuários, que percebem problemas de interação com regularidade. A fala, entretanto, não se apresenta como “mais natural” que um mouse, um teclado, um botão, uma tela – ela oferece outras potencialidades e funciona para perfis e contextos distintos. Ainda assim, a fala se configura como a forma de preferência de interação de alguns dos participantes dessa pesquisa.

Por possuírem uma materialidade que também opera sentidos, geram possibilidades e limites específicos. Na indissociabilidade do hardware e do software, que geram efeitos materiais no mundo, elas apresentam características de voz que se relacionam com seu corpo enquanto máquina: elas falam com pausas que nos soam estranhas, entonações equivocadas ou fora de lugar, cadência e prosódia ainda destoantes ou ausentes e vozes mais ou menos mecanizadas. Seus timbres podem ser mais neutros, como o da *Alexa*, ou mais agudo como da *Google Assistente*, oferecendo potenciais de propagação diferentes no ambiente. Outras escolhas dessas vozes, entretanto, dizem mais da materialidade das relações sociais e das culturas do que dos corpos-máquina das assistentes: são do gênero feminino, inicialmente de forma exclusiva, depois de forma padrão. Das quatro assistentes do mercado, apenas a *Siri* possui a opção de voz masculina em português operante – *Cortana*, *Alexa* e *Google Assistente* permanecem apenas femininas na língua portuguesa (e, exceto para a assistente do *Google*, também na língua inglesa).

Escolhas como a utilização da forma conversacional, da voz, da personalidade e do gênero são escolhas deliberadas de design das assistentes enquanto produtos, que visam atingir efeitos sociais específicos. A partir dos preceitos do paradigma da CASA (BRAVE; NASS, 2005; NASS; STEUER, 1994; NASS; MOON, 1997; NASS; YEN, 2010; NASS;2000;2010), que afirma que voz e gênero possuem consequências positivas na percepção que os usuários têm das interfaces, utilizam-se esses elementos para obter efeitos de aumento de confiança e de replicação de padrões de comportamento social por parte dos usuários, por exemplo.

Configuram-se, assim, como objetos dotados de política (LANGDOM, 1980). Essas escolhas geram efeitos materiais nas relações estabelecidas que, conforme analisado, não corroboram algumas das premissas utilizadas pela CASA (BRAVE; NASS, 2005; NASS; STEUER, 1994; NASS; MOON, 1997; NASS; YEN, 2010; NASS;2000;2010).

Soma-se ao gênero da voz, o gênero do nome, da personalidade e das respostas programadas. Mais do que isso, a escolha de disponibilizar assistentes pessoais digitais apenas em construtos femininos, propõe uma percepção do lugar da mulher em uma divisão sexual do trabalho e do cuidado como uma característica (e conseqüentemente uma tarefa) do feminino. Essa escolha, portanto, possui implicações que são tanto políticas quanto culturais e apontam para ambientes de produção centrados no masculino e discriminatórios com as mulheres, recolocando a mulher como responsável por atentar para as necessidades do outro, por servir e cuidar, por ser sempre simpática, doce e nunca desobedecer ou retrucar. Existem trabalhos para os quais as mulheres são adequadas e outros para os quais não são: não raro, entretanto, os adequados são aqueles relacionados à assistência, em geral, masculina. Ainda, determinadas construções de diálogos e retornos conversacionais das assistentes colocam o feminino como subserviente e como aquele que luta constantemente pela atenção do masculino: brincadeiras com a ideia de “ciúme” de uma assistente por outra, como se fossem ser “trocadas”, denotam a ideia de um cruzamento sexual. Na “briga” da *Siri* com o usuário que “troca” seu nome pelo da *Alexa*, a mulher aparece mais uma vez como aquela que luta para ser objeto de desejo preferido do homem, mesmo em relações não-sexuais (no caso do papel da assistente, por exemplo).

Dizer que a exploração dessa fórmula de um suposto papel feminino através das assistentes pessoais digitais é prejudicial é quase redundante. As implicações dessa construção e da perpetuação de estereótipos machistas em objetos da cultura é a própria perpetuação de visões desse tipo sobre o feminino. Além disso, existe o papel da naturalização das circunstâncias: usuários apontam que acabam preferindo ou não conseguindo utilizar outra voz porque a padrão (ou única opção) é a feminina. Nos perguntamos: que outros sentidos são normalizados conjuntamente com o feminino da voz?

Apesar de tudo, as relações que se estabelecem entre as usuárias mulheres e as assistentes, quando cruzadas pelo viés de gênero, podem ser surpreendentes. Elementos de cumplicidade e representatividade emergem nas entrevistas e nos formulários, além de outras leituras de usuárias que colocaram as assistentes, enquanto mulheres, como aquelas que comandam, que sabem tudo - e não aquelas que obedecem e realizam ações. Através dessas percepções e ações subverte-se a lógica proposta.

O tema proposto para investigação nesta tese não foi, certamente, esgotado neste trabalho. Permanecem, ainda, muitas questões não exploradas, que podem ser expandidas em pesquisas futuras, especialmente se focados em especificidades das utilizações. Em outros trabalhos, cabe investigar, por exemplo, a inclusão de assistentes pessoais digitais na interação com crianças ainda não alfabetizadas ou nos anos iniciais de idade escolar, quando as distinções entre o real e o imaginário são menos precisas. Ou, ainda, as relações específicas de usuários que utilizam as assistentes pessoais digitais como forma de tecnologia assistiva (como deficientes visuais ou portadores de deficiência física) também são pontos interessantes para uma pesquisa situada, uma vez que podem apontar para utilização e relações distintas das observadas aqui. Por fim, interações específicas de crianças autistas com estas tecnologias também se apresentam como um caminho interessante, na intersecção com a psicologia, para pensar a constituição do outro na relação humana.

## REFERÊNCIAS

AGAMBEN, Giorgio. **O que é contemporâneo?** E outros ensaios. Chapecó: Argos, 2009.

AMARO, Mariana. **Eu não posso ser dois:** uma perspectiva sobre o conceito de gameplay a partir de experimentos com o jogo *Brothers: a tale of two sons*. 2016. 310f. Dissertação (Comunicação e Informação) – Faculdade de Biblioteconomia e Comunicação, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2016.

ANDERSON, Chris. The End of Theory: The Data Deluge Makes the Scientific Method Obsolete. **Wired**, 23 de junho de 2008. Online. Disponível em: <<<https://www.wired.com/2008/06/pb-theory/>>>. Último acesso em 19 de maio de 2018.

ARAÚJO, William. **As narrativas sobre os algoritmos do Facebook:** uma análise dos 10 anos do feed de notícias. 2017. Tese (Doutorado em Comunicação e Informação) – Faculdade de Comunicação e Informação, Universidade Federal do Rio Grande do Sul, Porto Alegre.

ARMITAGE, John; KITTLER, Friedrich. From Discourse Networks to Cultural Mathematics: An Interview with Friedrich A. Kittler. **Theory, Culture & Society**, SAGE: London, Thousand Oaks and New Delhi, v. 23, n. 7-8, p. 17–38, dez. 2006.

BASSET, Caroline. The computational therapeutic: exploring Weizenbaum's ELIZA as a history of the present. **AI & Society**, publicação online, 2018. Disponível em: <<https://link.springer.com/article/10.1007%2Fs00146-018-0825-9>>.

BAUDRILLARD, Jean. **Tela total:** mito-ironias da era do virtual e da imagem. Porto Alegre: Editora Sulina, 1999.

Be Right Back (Temporada 2, ep. 1). *Black Mirror* [Seriado]. Direção: Owen Harris. Produção: Zeppotron. Amsterdã: Endemol, 2013. 1 DVD (44 min.).

BOLTER, David Jay; GRUSIN, Richard. **Remediation:** understanding new media. Cambridge, Massachusetts: MIT Press, 2000.

BRAGA, Adriana; CHAVES, Mônica. Inteligência Artificial e Singularidade: a metafísica das tecnologias digitais. Anais do XXVIII Encontro Anual da Compós, 2019.

BRAHNAM, Sheryl; KARANIKAS, Marianne; WEAVER, Margaret. (Un)dressing the interface: Exposing the foundational HCI metaphor “computer is woman”. **Interacting with Computers**, v. 23, n. 5, p. 401–412, set. 2011.

BRAVE, Scott; NASS, Clifford; HUTCHINSON, Kevin. Computers that care: Investigating the effects of orientation of emotion exhibited by an embodied computer agent. **International Journal of Human-Computer Studies**, v. 62, n.2, p. 161–178, 2005.

BROUSSARD, Meredith. **Artificial Unintelligence**: how computers misunderstand the world. Cambridge, Estados Unidos: The MIT Press, 2018.

BUCHER, Taina. Want to be on the top? Algorithmic power and the threat of invisibility on Facebook. **New Media & Society**, n.7, p.1164-1180, 2012.

\_\_\_\_\_. The algorithmic imaginary: Exploring the ordinary affects of Facebook algorithms. **Information, Communication & Society**, v. 20, n. 1, p. 30-44, 2016.

\_\_\_\_\_; D’ANDREA, Carlos; JURNO, Amanda. Algorithms as a becoming: an interview with Taina Bucher. **Revista Parágrafo**. São Paulo, v. 6, n.1, p. 165-170, jan/abr 2018.

BUTLER, Judith. **Vida precária**. Contemporânea: revista de Sociologia da UFSCar, v.1, n.1, p. 13-33, 2011.

BURTON, Emmanuelle; GOLSMITH, Judy; KOENIG, Sven; KUIPERS, Benjamin; MATTEI, Nicholas; WALSH, Toby. Ethical Considerations in Artificial Intelligence Courses. 26 de Janeiro de 2017. Disponível em < <https://arxiv.org/abs/1701.07769>>.

CAETANO, Mayara Araújo. **Performances de gênero nas relações entre jogadores e avatares**: dinâmicas com o game Rust. 2017. 211f. Dissertação (Comunicação e Informação) – Faculdade de Biblioteconomia e Comunicação, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2017.

CARPENTER, Edmund. As novas linguagens. In: \_\_\_\_\_; MCLUHAN, Marshall (Org.).

**Revolução na Comunicação.** Rio de Janeiro: Jorge Zahar, 1974, p. 197 – 217.

CARPENTER, Julie. **Culture and Human-Robot Interaction in militarized spaces: A war story.** Farnham: Ashgate, 2016.

\_\_\_\_\_. The existential robot: living with robots may teach us to be better humans. In: **Issues Magazine**, n. 108, p. 39-42, set. 2014.

CASE, Amber. **Calm Technology: Principles and Patterns for Non-Intrusive Design.** Sebastopol, Estados Unidos: O'Reilly Media, 2015.

CATÃO, Inês. **O bebê nasce pela boca: voz, sujeito e clínica do autismo.** São Paulo: Instituto Langage, 2009.

CIANI, Andrea; CHIARELLI, Brunetto. A Systematic relationship between Brain Size Increase and Larynx Transformation during Hominization. In: WIND, Jan; CHIARELLI, Brunetto; BICHAKJIAN, Bernard; NOCENTINI, Alberto; JONKER, Abraham (Orgs.). **Language Origin: A multidisciplinary approach.** Nova Iorque: Springer, 1992, p.

CHESLER, Caren. Alexa? Please Ignore My Husband. **New York Times**, 4 de maio de 2018. Online. Disponível em: <<<https://www.nytimes.com/2018/05/04/style/modern-love-alexa-please-ignore-my-husband.html>>>. Último acesso 9 de março de 2020.

CHIN, Louie. Como Siri, assistente virtual da Apple, se tornou a melhor amiga de uma criança autista. **Zero Hora**, 28 de outubro de 2014. Online. Disponível em: <<<https://gauchazh.clicrbs.com.br/saude/vida/noticia/2014/10/Como-Siri-assistente-virtual-da-Apple-se-tornou-a-melhor-amiga-de-uma-crianca-autista-4630628.html>>>. Último acesso 9 de março de 2020.

CHUN, Wendy Hui Kyong. On Software, or the Persistence of Visual Knowledge. In: **Grey Room**, n. 18, 2004, pp. 26-51.

\_\_\_\_\_. The enduring ephemeral, or the future is a memory. **Critical inquiry**, v. 35, n. 1, p. 148-171, 2008. Disponível em <[http://aestech.wikischolars.columbia.edu/file/view/Hui%20Kyong%20Chun--the\\_enduring\\_ephemeral\\_or.pdf/442522752/Hui%20Kyong%20Chun--the\\_enduring\\_ephemeral\\_or.pdf](http://aestech.wikischolars.columbia.edu/file/view/Hui%20Kyong%20Chun--the_enduring_ephemeral_or.pdf/442522752/Hui%20Kyong%20Chun--the_enduring_ephemeral_or.pdf)>. Acesso em 16/02/2017.

DEMO, Pedro. **Metodologia científica em ciências sociais**. Atlas: São Paulo, 2009.

DERRIDA, Jacques. **Mal de Arquivo**: uma impressão Freudiana. Rio de Janeiro: Relume Dumará, 2001.

DESCARTES, René. **Discurso do Método**. São Paulo: Martins Fontes Editora, 2009.

DEVLIN, Kate. **Turned On**: science, sex and robots. Londres: Bloomsbury, 2018.

DOURISH, Paul. **The Stuff of Bits**: An Essay on the Materialities of information. Cambridge: MIT Press, 2017.

DREYFUS, Hubert. **What Computers Still Can't Do**: A Critique of Artificial Reason. Cambridge: MIT Press, 1992.

ELA. Direção de Spike Jonze. Los Angeles: Annapurna Pictures, 2013. Disponível em: <<https://www.netflix.com/title/70278933>>. Acesso em 15 de fevereiro de 2020 (126 min.).

EX-MACHINA: Instinto Artificial. Direção de Alex Garland. Universal City, California: Universal Pictures International, 2014. Disponível em: <<https://www.netflix.com/title/80023689>>. Acesso em 9 de janeiro de 2020 (108 min.).

FALCÃO, Thiago. **Camadas relacionais de prescrição**. Alicerces para uma Metodologia de Análise de Jogos Eletrônicos. In: Anais do XXIII Encontro Anual da Compós. Belém, 2014.

FAZI, Beatrice. Can a machine think (anything new)? Automation beyond simulation. In: **AI & Society**. 2018. <https://doi.org/10.1007/s00146-018-0821-0>.

\_\_\_\_\_. The ends of Media Theory. **Media Theory Journal**, v. 1, n. 1, p. 107-121, 2018.

\_\_\_\_\_; PARISI, Luciana. Do Algorithms Have Fun? On Completion, Indeterminacy and Autonomy in Computation. In: GORIUNOVA, Olga (Org.). **Algorithms and Fun**. Nova York: Bloomsbury Academic, 2014.

FEDERICI, Silvia. **Calibã e a Bruxa**: mulheres, corpo e acumulação primitiva. São Paulo:



Elefante, 2017.

FELINTO, Erick. Materialidades da Comunicação: Por um Novo Lugar da Matéria na Teoria da Comunicação. **Ciberlegenda**, Rio de Janeiro, n. 5, 2001.

\_\_\_\_\_. O Corpo Impuro: sobre a Digitalização da Matéria no Imaginário da Cibercultura. In: **Anais do XIII Encontro Anual da Compós**. São Bernardo, 2004.

\_\_\_\_\_. Em busca do Tempo Perdido. O sequestro da história na Cibercultura e os Desafios da Teoria da Mídia. In: **Anais do XIX Encontro Anual da Compós**. Rio de Janeiro, 2010.

\_\_\_\_\_. Da Teoria da Comunicação às Teorias da Mídia. Ou, temperando a epistemologia com uma dose de cibercultura. In: **Anais do XX Encontro Anual da Compós**. Porto Alegre, 2011.

\_\_\_\_\_. Novos Materialismos, Realismo Especulativo e Comunicologia: sobre algumas proposições para as ciências humanas em diálogos com as artes e as tecnologias contemporâneas. In: PEREIRA, Vinicius Andrade. CASTELLARY, Arturo Colorado (Org.). **Artecnologia: Arte, Tecnologia e Linguagens Midiáticas**. Porto Alegre: Buqui Editora, 2013.

FESSLER, Leah. We tested bots like Siri and Alexa to see who would stand up to sexual harassment. **The Atlantic**, novembro de 2018. Online. Disponível em <<<https://qz.com/911681/we-tested-apples-siri-amazon-echos-alexa-microsofts-cortana-and-googles-google-home-to-see-which-personal-assistant-bots-stand-up-for-themselves-in-the-face-of-sexual-harassment/>>>. Último acesso 9 de março de 2020.

FISCHER, Gustavo. I don't wanna be buried in an app sematary – reflexões sobre arqueologia da mídia online entre histórias de aplicativos derrotados. In: **Anais do VII Simpósio Nacional da Associação Brasileira de Pesquisadores em Cibercultura**, 2013.

FLUSSER, Vilém. **O mundo codificado**: por uma filosofia do design e da comunicação. São Paulo: Cosac Naify, 2007.

\_\_\_\_\_. **A Escrita** – Há futuro para a escrita? São Paulo: Annablume, 2010.

FOSTER, Jonathan. Why You Should Design UX with Personality in Mind. **Medium**. 10 de julho de 2018. Online. Disponível em: <<<https://medium.com/microsoft-design/why-you-should-design-ux-with-personality-in-mind-35eb99e73d70c>>>. Último acesso 9 de março de 2020.

\_\_\_\_\_. What Did We Get Ourselves Into?. **Medium** 4 de julho de 2018. Online. Disponível em: <<<https://medium.com/microsoft-design/what-did-we-get-ourselves-into-36ddae39e69bc>>>. Último acesso 9 de março de 2020.

\_\_\_\_\_. The Microsoft Voice, Part 4: Listening Is the New Helpful. **Medium**. 26 de abril de 2019. Online. Disponível em: <<<https://medium.com/microsoft-design/the-microsoft-voice-part-4-listening-is-the-new-helpful-6f78700e2e0c>>>. Último acesso 9 de março de 2020.

FOUCAULT, Michel. **Microfísica do poder**. Rio de Janeiro: Edições Graal, 1979.

FRAGOSO, Suely; RECUERO, Raquel; AMARAL, Adriana. Métodos de pesquisa para internet. Porto Alegre: Sulina, 2012.

FRAGOSO, Suely. Os modos de existência do gameplay: um exercício de aplicação com Cities: Skylines. In: **Anais do XXV Encontro Anual da Compós**. São Paulo, 2017.

FREUD, Sigmund. Artigos sobre a metapsicologia. O inconsciente. In: FREUD, Sigmund. **Obras completas volume 14**. Rio de Janeiro: Imago, 1980.

FREUD, Sigmund. **Sobre a Concepção das Afasias**: um estudo crítico. Belo Horizonte: Autêntica, 2013.

\_\_\_\_\_. **O infamiliar** [Das Unheimliche, 1919]. Trad. Ernani Chaves e Pedro Heliodoro Tavares. Belo Horizonte: Autêntica, 2019.

GEERTZ, Clifford. **A Interpretação das Culturas**. Rio de Janeiro: Zahar Editores, 1978.

GENOVA, Judith. Turing's sexual guessing game. **Social Epistemology**, v. 8, n. 4, p. 313–326, 1994.

GIRARDI JÚNIOR, Liráucio. O Estranho Mundo da Informação - e da Materialidade - no campo da Comunicação. In: **Anais do XXV Encontro Anual da Compós**. Goiânia, 2016.

GILLESPIE, Tarleton. A relevância dos algoritmos. **Revista Parágrafo**. São Paulo, v. 6, n.1, p.95-121, jan./abri. 2018.

GONZÁLEZ DE GÓMEZ, Maria Nélide. A reinvenção contemporânea da informação: entre

o material e o imaterial. **Pesq. bras. Ci. Inf.**, Brasília, v.2, n.1, p.115-134, jan./dez. 2009

GREY, Mary L.; SURI, Siddharth. **Ghost Work: How to Stop Silicon Valley from Building a New Global Underclass**. Nova Iorque: Eamon Dolan/Houghton Mifflin Harcourt, 2019.

GREEN, Penelope. Alexa, where have you been all my life? **New York Times**, 11 de julho de 2017. Online. Disponível em: <<<https://www.nytimes.com/2017/07/11/style/alexa-amazon-echo.html>>>. Último acesso 9 de março de 2020.

GUMBRECHT, Hans Ulrich. **Produção de Presença: o que o sentido não consegue transmitir**. Rio de Janeiro: Contraponto/Ed. PUC-Rio, 2010.

\_\_\_\_\_. A história das mídias como evento da verdade: sobre a singularidade da obra de Friedrich A. Kittler. In: KITTLER, Friedrich. **A verdade do mundo técnico: ensaios sobre a genealogia da atualidade**. Rio de Janeiro: Contraponto, 2017, p. 515-550.

GUNKEL, Ann Hetzel. (De)Facing Alterity in the Digital Age: “The Real Problem” in the Social Interaction of Digital Natives. In: GUNKEL, David J.; MARCONDES FILHO, Ciro; MERSCH, Dieter (Org.). **The Changing Face of Alterity: Communication, Technology, and Other Subjects**. Londres: Rowman & Littlefield, 2016, p.

GUNKEL, David. Communication and Artificial Inteligence: Opportunities and Challenges for the 21st Century. **Communication+1**, v. 1, ago. 2012.

\_\_\_\_\_. Comunicação e inteligência artificial: novos desafios e oportunidades para a pesquisa em comunicação. **Revista Galáxia**, n. 34, São Paulo, jan./abril 2017.

G1 Ouviu: ‘Voz do Google’ desponta no pop brasileiro: ouça os hits e a trajetória dessa revelação. G1, 24 nov. 2019. Podcast. Online. Disponível em: <<<https://g1.globo.com/podcast/g1-ouviu/noticia/2019/11/24/g1-ouviu-65-os-hits-e-a-historia-de-uma-revelacao-musical-de-2019-no-brasil-a-voz-do-google.ghtml>>>. Último acesso em 9 de março de 2020.

HAMILTON, Isobel Asher. Google gave Google Assistant a weirdly specific backstory to get her voice just right. **Business Insider**, 10 de outubro de 2018. Online. Disponível em: <<<https://www.businessinsider.com/google-gave-google-assistant-a-weirdly-specific-backstory-2018-10>>>. Último acesso 9 de março de 2020.

HANSEN, Mark. Symbolizing Time: Kittler and 21<sup>st</sup> Century Media. In: SALISBURY, Laura; SALE, Stephen (Orgs.). **Kittler Now**. Palo Alto: Stanford University Press, 2012, p. 210 -237.

HARARI, Yuval Noah. **21 lições para o século 21**. São Paulo: Companhia das Letras, 2018.

HAREL, David. **Computers Ltd**: What they really can't do. Oxford University Press, USA, 2000.

\_\_\_\_\_. **Algorithmic**: the spirit of computing. Edinburgh: Pearson Education Limited, 2004.

HARNAD, Stevan. The Annotation Game: On Turing (1950) on Computing, Machinery, and Intelligence. In: ESPTEIN, ROBERT & PETERS, GRACE (Editores). **Parsing the Turing Test**: Philosophical and Methodological Issues in the Quest for the Thinking Computer. Springer, 2008.

HARMAN, Graham. **Immaterialism**. Cambridge: Polity Press, 2016.

HARRISON, Sara. Of Course Citizens Should Be Allowed to Kick Robots. **Wired**, 29 de agosto de 2019. Online. Disponível em: <<<https://www.wired.com/story/citizens-should-be-allowed-to-kick-robots/>>>. Último acesso 9 de março de 2020.

HERN, Alex. Apple made Siri deflect questions on feminism, leaked papers reveal. The Guardian, 6 de setembro de 2019. Online. Disponível em: <<<https://www.theguardian.com/technology/2019/sep/06/apple-rewrote-siri-to-deflect-questions-about-feminism>>>. Último acesso 9 de março de 2020.

HINE, Christine. **Virtual Ethnography**. London: Sage, 2000.

HOLL, Susanne. Friedrich Kittler's Digital Legacy - Part II - Friedrich Kittler and the Digital Humanities: Forerunner, Godfather, Object of research. An Indexer Model Research. **Digital Humanities Quarterly**, v. 11, n. 2, 2017.

HOOKWAY, Branden. **Interface**. Cambridge, Estados Unidos: The MIT Press, 2014.

HOUAISS, Antonio; VILLAR, Mauro de Salles. Dicionário Houaiss da Língua Portuguesa, elaborado pelo Instituto Houaiss de Lexicografia e Banco de Dados da Língua Portuguesa. 1. ed. Rio de Janeiro: Objetiva, 2009.

HUI, Yuk. **On the existence of digital objects**. Minneapolis: University of Minnesota Press, 2016.

INNIS, Harold A. **O Viés da Comunicação**. Petrópolis, RJ: Vozes, 2011.

ISTVÁN, S. N. What is Artificial Intelligence? 1996. Disponível em: <<http://www.ucl.ac.uk/~isb9112/dept/phil341/wisai/WhatisAI.html>>. Acesso em 15 de maio de 2018.

JARRETT, Kylie. **Feminism, labour and digital media: the digital housewife**. Londres e Nova Iorque: Routledge, 2015.

JENTSCH, Ernst. **On the Psychology of the Uncanny**, 1906. Disponível em <[http://www.art3idea.psu.edu/locus/Jentsch\\_uncanny.pdf](http://www.art3idea.psu.edu/locus/Jentsch_uncanny.pdf)>.

JERUSALINSKY, Julieta. **A criação da criança: letra e gozo nos primórdios do psiquismo**. 2009. Tese (Doutorado em Psicologia) – Universidade de São Paulo, São Paulo, 2009.

JURNO, Amanda Chevtchouk. Agenciamentos coletivos e textualidades em rede no Facebook: uma exploração cartográfica. Belo Horizonte, 2016. Dissertação (Mestrado em Comunicação Social). Universidade Federal de Minas Gerais, Belo Horizonte.

\_\_\_\_\_; D'ANDRÉA, Carlos Frederico de Brito. (In) Visibilidade Algorítmica no “Feed de notícias” do Facebook//Algorithmic (In) Visibility in Facebook News Feed. **Contemporanea - Revista de Comunicação e Cultura**, v. 15, n. 2, p. 463-484, 2017.

KANDEL, Eric. **Em busca da memória: o nascimento de uma nova ciência da mente**. São Paulo: Companhia das Letras, 2009.

KÄTSYRI, Jari; FÖRGER, Klaus. MÄKÄRÄINEN, Meeri; TAKALA, Tapio. A review of empirical evidence on different uncanny valley hypotheses: support for perceptual mismatch as one road to the valley of eeriness. **Frontiers in Psychology**, n. 6, p. 390, 2015.

KITTLER, Friederich. *Discourse Networks 1800/1900*. California: Stanford University, 1990.

\_\_\_\_\_. The history of Communication Media. **CTheory.net**, 30/07/1996. Disponível em: <<http://www.ctheory.net/articles.aspx?id=45#bio>>. Acesso em: 01 de fev de 2017.

\_\_\_\_\_. **Gramophone, Film, Typewriter**. Stanford, California: Stanford University Press, 1999.

\_\_\_\_\_. Number and Numeral. **Theory, Culture & Society**, v. 23, n. 7-8, p. 51-61, 2006.

\_\_\_\_\_. Exorcizar o espírito das Ciências do Espírito. In: RODRIGUES, Jorge Henrique Vieira. Friedrich A. Kittler, Filósofo da Tecnologia: uma tradução comentada. 2013. Dissertação (Mestrado em Ciências da Comunicação). Universidade de Lisboa, Lisboa.

\_\_\_\_\_. **The Truth of the Technological World: essays on the Genealogy of Presence**. Stanford, California: Stanford University Press, 2013.

\_\_\_\_\_. **Mídias ópticas**. Rio de Janeiro: Contraponto, 2016.

\_\_\_\_\_. **A verdade do mundo técnico: ensaios sobre a genealogia da atualidade**. Rio de Janeiro: Contraponto, 2017.

KOERBER, Amy. Toward a feminist rhetoric of technology. **Journal of Business and Technical Communication**, v. 14, n. 1, p. 58–73, 2000.

KOZINETS, Robert V. **Netnography: Doing Ethnographic Research Online**. London: Sage, 2010.

KUHN, Thomas. **A Estrutura das Revoluções Científicas** [1962]. São Paulo: Perspectiva, 2006.

LACAN, Jacques. O estádio do espelho como formador da função do eu. In: LACAN, Jacques. **Escritos**. Rio de Janeiro: Zahar, 1998.

LAFRANCE, Adrienne. Why Do So Many Digital Assistants Have Feminine Names? **The Atlantic**, 30 de março de 2016. Online. Disponível em <<<https://www.theatlantic.com>

/technology/archive/2016/03/why-do-so-many-digital-assistants-have-feminine-names/475884/>>. Último acesso 9 de março de 2020.

LATOUR, Bruno. Where are the missing masses? The sociology of a few mundane artifacts. In: LAW, John; BIJKER, Wiebe. (Orgs.). **Shaping Technology/Building Society: Studies in Sociotechnical Change**, Cambridge: MIT Press, 1992, p. 225-258.

LATOUR, Bruno. **Reagregando o Social**. Salvador: EDUFBA, 2012.

LARGELY, Pat. The Changing Science of Machine Learning. **Machine Learning**, v. 82, n. 3, pp. 275-279, mar. 2011.

LEMOS, André; PASTOR, Leonardo. **Internet das coisas, automatismo e fotografia**. Uma análise pela teoria ator-rede. In: Anais do XXIII Encontro Anual da Compós. Belém, 2014.

LUPINACCI AMARAL, Ludmila. **As apropriações do GIF animado**: aspectos culturais, expressivos e afetivos dos usos de uma tecnologia defasada. 2016. Dissertação (Mestrado em Comunicação e Informação) – Faculdade Comunicação e Informação, Universidade Federal do Rio Grande do Sul, Porto Alegre.

MACHERETH, Kerry. Mechanical maids and family androids: racialised post-care imaginaries in *Humans* (2015-), *Sleep Dealer* (2008) and *Her* (2013). **Feminist Review**, n. 123, p. 24-39, 2019.

MAIA, Alessandra. **A Materialidade do jogar no Kinect**: o terror ganha outras proporções. In: Anais do XXII Encontro Anual da Compós. Bahia, 2013.

MANOVICH, Lev. **The Language of new media**. Cambridge: MIT Press, 2001.

\_\_\_\_\_. **Cultural Analytics**: Visualizing Cultural Patterns in the Era of “More Media”. Domus, Milão: 2009. Disponível em: <<http://manovich.net/index.php/projects/cultural-analytics-visualizing-cultural-patterns>>. Acesso em 5 de julho de 2018.

\_\_\_\_\_. **What is visualization?** 2010. Disponível em: <<http://manovich.net/index.php/projects/what-is-visualization>>. Acesso em 5 de julho de 2018.

\_\_\_\_\_. **Software takes command.** New York: Bloomsbury Academic, 2013.

\_\_\_\_\_. **The Science of Culture?** Social Computing, Digital Humanities, and Cultural Analytics. 2015. Disponível em: <http://manovich.net/index.php/projects/cultural-analytics-social-computing>. Acesso em 5 de julho de 2018.

MÄNNISTÖ-FUNK, Tiina; SIHVONEN, Tanja. **Voices from the uncanny valley:** how robots and Artificial Intelligences Talk back to Us. *Digital Culture & Society*, v.4, n.1, p. 45-64, 2018.

MCLUHAN, Marshall (Org.). **Os meios de comunicação como extensões do homem.** São Paulo: Cultrix, 1971.

\_\_\_\_\_. **A Galáxia de Gutemberg: a formação do homem tipográfico.** São Paulo: Editora da USP, 1972.

\_\_\_\_\_. **Revolução na Comunicação.** Rio de Janeiro: Jorge Zahar, 1974, p. 137-144.

\_\_\_\_\_. **Os meios de comunicação como extensões do homem.** São Paulo: Editora Cultrix, 1974.

MERCHANT, Brian. **The One Device:** the secret history of the iPhone. Londres: Penguin Random House UK, 2017.

MOON, Mariella. Apple buys the voice tech startup behind Hello Barbie. **Engadget**, 16 de fevereiro de 2019. Online. Disponível em <<<https://www.engadget.com/2019/02/16/apple-buys-pullstring/>>>. Último acesso 9 de março de 2020.

MORI, Masahiro. **The Uncanny Valley.** Publicado em 12 de junho de 2012, versão online. Disponível em: <<<https://spectrum.ieee.org/automaton/robotics/humanoids/the-uncanny-valley>>>. Último acesso em 9 de março de 2020.

MOZZILA FOUNDATION. **Final Report for AI Campaign Survey.** Outubro de 2019. Disponível em: <<<https://foundation.mozilla.org/en/blog/we-asked-people-around-the-world-how-they-feel-about-artificial-intelligence-heres-what-we-learned/>>>. Último acesso em 9 de março de 2020.

MURPHY, Robin R. **An Introduction to AI Robotics.** Cambridge, Estados Unidos: The MIT Press, 2000.



NASRABADI, Shahin Kristian Araujo Khalili. **Interfaces de Voz**: Uma escavação pelas máquinas de falar. 2019. Trabalho de Conclusão de Curso (Comunicação Digital) - Universidade do Vale do Rio dos Sinos, São Leopoldo, 2019.

NASS, Clifford. Speech interfaces from an evolutionary perspective. **Communications of the ACM**, v.43, n.9, p. 36-43, set. 2000.

NASS, Clifford; BRAVE, Scott. **Wired for Speech**: How Voice Activates and Advances the Human-Computer Relationship. Cambridge, Estados Unidos: The MIT Press, 2007. Edição Kindle.

NASS, Clifford; MOON, Youngme; GREEN, Nancy. Are machines gender neutral? Gender-stereotypic responses to computers with voices. **Journal of Applied Social Psychology**, v. 27, n. 10, p. 864-876, 1997.

NASS, Clifford; STEUER, Jonathan; TAUBER, Ellen R. Computers are social actor. **CHI '94: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems**, p. 72-78, 1994.

NASS, Clifford; YEN, Corina. **The man who lied to his laptop**. Londres: Current, 2010.

NOTH, Winfried. Máquinas simbólicas e semióticas. *Galáxia*, v.1, n. 1, p. 51-73, 2001.

O GAROTO Selvagem. Direção de François Truffaut. Paris: Les Artistes Associés, Les Films du Carrosse, Les Productions Artistes Associés, 1970. 1 DVD (83 min.).

O ENIGMA de Kaspar Hauser. Direção de Werner Herzog. Berlim: Werner Herzog Filmproduktion, Zweites Deutsches Fernsehen (ZDF) e Filmverlag der Autoren, 1974. 1 DVD (110 min.).

ONG, Walter. **Orality and Literacy**: the technologizing of the word [1982]. Londres e Nova Iorque: Routledge, 2002.

PARIKKA, Jussi. **What is media archaeology?** Cambridge: Polity, 2012.

PARISER, Eli. **The Filter Bubble.** What the Internet is Hiding from You. The Pinguim Press. New York. 2011.

PEARL, Cathy. **Designing Voice User Interfaces:** principles of conversational experiences. Sebastopool, Estados Unidos: O'Reilly Media, 2016.

PIERCE, David. How Apple Finally Made Siri Sound More Human. **Wired**, 7 de setembro de 2017. Online. Disponível em: << <https://www.wired.com/story/how-apple-finally-made-siri-sound-more-human/>>>. Último acesso 9 de março de 2020.

POLANYI, Michael. **Personal Knowledge:** towards a post-critical philosophy [1958]. Chicago: University of Chicago Press, 1974.

POPPER, Karl. **A lógica da pesquisa científica.** São Paulo: Cultrix, 2013.

PRIMO, Alex. **Interação mediada por computador:** comunicação, cibercultura, cognição. Porto Alegre: Editora Sulina, 2011.

REEVES, Byron; NASS, Clifford. **The media equation:** How people treat computers, television, and new media like real people and places. Cambridge, Inglaterra: Center for the Study of Language and Information; Cambridge University Press, 1996.

RIEDER, Bernhard. Examinando uma técnica algorítmica: o classificar de base como uma leitura interessada da realidade. **Revista Parágrafo.** São Paulo, v. 6, n.1, p.95-121, jan./abril. 2018.

RIESMAN, David. As tradições oral e escrita. In: CARPENTER, Edmund; MCLUHAN, Marshall (Org.). **Revolução na Comunicação.** Rio de Janeiro: Jorge Zahar, 1974, p. 137-144.

SÁ, Simone Pereira de; POLIVANOV, Beatriz Brandão. Materialidade da Comunicação e Presentificação do sujeito em sites de redes sociais. In: **Anais do XXI Encontro Anual da Compós,** 2012.

SACKS, Oliver. **Vendo Vozes**. São Paulo: Companhia de Bolso, 2010.

SANTOS, Luiza C. Nós que aqui estamos por vós esperamos: apontamentos iniciais sobre os lugares dos mortos e dos vivos em ambientes digitais na sociedade brasileira contemporânea. **Anais do XVII Encontro dos Grupos de Pesquisa em Comunicação do 40º Congresso Brasileiro de Ciências da Comunicação**, 2017.

SEARLE, John R. Minds, Brains and Programs. **Behavioral and Brain Sciences**, v. 3, n. 3, p. 417-424, 1960.

SHARMA, Sarah. “Exit and the Extensions of Man”. **Transmediale Online Journal**, 2017. Disponível em: << <https://transmediale.de/content/exit-and-the-extensions-of-man> >>. Último acesso em 9 de março de 2020.

SHAW, Debra Benita. **Technoculture: The key concepts**. Oxford: Bloomsbury Academic, 2008.

SHULEVITZ, Judith. Alexa, Should We Trust You?. **The Atlantic**, novembro de 2018. Online. Disponível em: <<<https://www.theatlantic.com/magazine/archive/2018/11/alexahow-will-you-change-us/570844/>>>. Último acesso 9 de março de 2020.

\_\_\_\_\_. The story of Cortana, Microsoft's Siri killer. **The Atlantic**, novembro de 2018. Online. Disponível em <<<https://www.theatlantic.com/magazine/archive/2018/11/alexahow-will-you-change-us/570844/>>>. Último acesso 9 de março de 2020.

SPRENGER, Florian. **The Politics of Micro-Decisions: Edward Snowden, Net Neutrality, and the Architectures of the Internet**. Lüneburg: Means Press, 2015.

\_\_\_\_\_. Academic Networks 1982/2016: The Provocations of a Reading. **Grey Room**, n. 63, p. 70-89, 2016.

SUPERINTERESSANTE. Como, quando e por que o ser humano passou a falar? **Superinteressante**, 31 de outubro de 2016. Online. <<<https://super.abril.com.br/comportamento/palavra-de-homem/>>>. Último acesso 9 de março de 2020.

SWEENEY, Miriam. **Not just a pretty (inter)face**: a critical analysis of Microsoft's "Ms. Dewey". 2013. Dissertação (Mestrado em Ciência da Informação e Biblioteconomia) – Graduate College of the University of Illinois at Urbana-Champaign, Universidade de Illinois, Urbana-Champaign, IL, Estados Unidos, 2013.

TELLES, Marcio. Das Materialidades as Matérias-Primas da Comunicação: notas para uma perspectiva teórica geológica. In: **Anais do XXV Encontro Anual da Compós**. Goiânia, 2016.

THE CURRENT: Do you swear at Alexa? What our treatment of AI assistants says about humans. Entrevistadora: Anna Maria Tremonti. Entrevistados: David Gunkel e Julie Carpenter. CBC Radio, 22 jan. 2019. Programa de Rádio. Disponível em: <<<https://www.cbc.ca/radio/thecurrent/the-current-for-january-22-2019-1.4984317/do-you-swear-at-alex-a-what-our-treatment-of-ai-assistants-says-about-humans-1.4984321>>>. Último acesso 9 de março de 2020

TINWELL, Angela; GRIMSHAW, Mark; WILLIAMS, Andrew. The Uncanny Wall. **International Journal of Arts and Technology**, v. 4, n. 3, p. 326-341, 2011.

WEST, Mark; KRAUT, Rebecca; CHEW, Han Ei. I'd blush if I could: closing gender divides in digital skills through education. UNESCO/EQUALs, 2019. Disponível em: <<https://unesdoc.unesco.org/ark:/48223/pf0000367416.page=1>>. Último acesso em: 4 de março 2020.

TURING, Alan. Computing Machinery and Intelligence. **Mind, New Series**, v. 59, n. 236, p. 433-460, out. 1950.

TURKLE, Sherry. **The second Self**: Computers and The Human Spirit. Cambridge: The MIT Press, 2005.

\_\_\_\_\_. **Evocative objects**: things we think with. Cambridge: The MIT Press,, 2007.

\_\_\_\_\_. **Alone Together**: Why we Expect More From Technology and Less From Each Other. Nova Iorque: Basic Books, 2010.

WAJCMAN, Judy. Tecnologia de produção: fazendo um trabalho de gênero. **Cadernos Pagu**, n. 10, p. 201-256, 1998.

\_\_\_\_\_. Feminist theories of technology. **Cambridge Journal of Economics**, v. 34, n. 1, p.

143–152, 2010.

WALSH, Toby. **Turing's Red Flag**. Association for the Advancement of Artificial, 2015.

\_\_\_\_\_. **Education: Future Frontiers Occasional Paper Series**. NSW Department of Education. Intelligence, 2015.

WANG, Edwin. Smartphones as Social Actors? Social dispositional factors in assessing anthropomorphism. **Computers in Human Behavior**, n.68, p. 334-344, mar. 2017.

WARREN, Tom. The story of Cortana, Microsoft's Siri killer. **The Verge**, 2 de abril de 2014. Online. Disponível em <<<https://www.theverge.com/2014/4/2/5570866/cortana-windows-phone-8-1-digital-assistant>>>. Último acesso 9 de março de 2020.

WEIZENBAUM, Joseph. ELIZA - A Computer Program for the Study of Natural Language Communication between Man and Machine. **Comm. Assoc. Comput.**, p. 36-45, mar. 1966.

\_\_\_\_\_. **Computer power and human reason: from judgment to calculation**. São Francisco, Estados Unidos: W. H. Freeman, 1976.

WINNICOTT, Donald. **Privação e Delinquência**. São Paulo: Martins Fontes, 1994.

WINTHROP-YOUNG, Geoffrey. Cultural Studies and German Media Theory. In: HALL, Gary e BIRCHALL, Claire (Orgs.). **New Cultural Studies: Adventures in Theory**. Edinburgh: Edinburgh University Press, 2006, p. 88–104.

## ANEXO 1 – FORMULÁRIO QUANTITATIVO

28/01/2020

Usos e Interações com Assistentes Pessoais Digitais

### Usos e Interações com Assistentes Pessoais Digitais

Este formulário integra uma pesquisa de doutorado em andamento no Programa de Pós-Graduação em Comunicação da Universidade Federal do Rio Grande do Sul, que tem por objetivo compreender as formas de interação entre seres humanos e dispositivos baseados em Inteligência Artificial, nomeadamente as assistentes pessoais virtuais (Siri, Cortana, Alexa e Google Assistant). O objetivo deste formulário é compreender de que forma os usuários interagem com as Assistentes Pessoais Virtuais e quais suas percepções sobre as mesmas. Por isso, para responder esse formulário, você precisa ter utilizado uma assistente pessoal virtual ao menos uma vez.

**\*Obrigatório**

1. Você já utilizou uma assistente pessoal virtual? (Ex: Alexa, Google Assistant, Cortana ou Siri) \*

*Marcar apenas uma oval.*

Sim

Não

Autorização

Leia o Termo de Consentimento Livre e Esclarecido aqui: <http://bit.ly/2IVVWWM>

2. Suas respostas são confidenciais e serão analisadas de forma quantitativa conjuntamente com as respostas dos demais participantes. Você autoriza a utilização de suas respostas para o estudo? \*

*Marcar apenas uma oval.*

Sim

Não

28/01/2020

Usos e Interações com Assistentes Pessoais Digitais

## 3. Assinale qual/quais assistentes pessoais virtuais você utiliza ou já utilizou. \*

*Marque todas que se aplicam.*

- Siri (Apple)  
 Cortana (Microsoft)  
 Alexa (Amazon)  
 Google Assistant (Google)

Outro:  \_\_\_\_\_

## 4. Com que frequência você utiliza? \*

*Marcar apenas uma oval.*

- Todos os dias  
 De 6 a 3 vezes na semana  
 De 2 a 1 vez na semana  
 Menos de 1 vez na semana  
 Utilizo apenas ocasionalmente  
 Utilizei apenas uma vez

## 5. Você tem preferência por alguma das assistentes pessoais virtuais? \*

*Marcar apenas uma oval.*

- Não  
 Sim, pela Cortana  
 Sim, pela Siri  
 Sim, pela Alexa  
 Sim, pelo Google Assistant/Google Home  
 Não posso afirmar, pois só utilizei uma  
 Eu gosto de todas  
 Outro: \_\_\_\_\_

28/01/2020

Usos e Interações com Assistentes Pessoais Digitais

## 6. Para qual finalidade você utiliza esses sistemas? \*

*Marque todas que se aplicam.*

- Para fazer pesquisas na internet
- Para realizar tarefas (ligações, mensagens, ouvir música, fazer compras, etc)
- Para conversar
- Para obter informações atualizadas (sobre o tempo, trânsito, etc)
- Para controlar outros dispositivos inteligentes
- Para treinar idiomas
- Para me distrair
- Para me divertir

Outro:  \_\_\_\_\_

## 7. Em quais locais você utiliza a assistente pessoal virtual? \*

*Marque todas que se aplicam.*

- Em locais compartilhados da casa (como sala e cozinha)
- Em locais privados da casa (como o quarto)
- No carro
- Na rua
- Na faculdade e/ou trabalho

Outro:  \_\_\_\_\_

## 8. Em quais dispositivos você utiliza a assistente pessoal virtual? \*

*Marque todas que se aplicam.*

- No celular
- No computador
- No tablet
- No seu próprio auto-falante (Alexa e Google Home)
- No console do carro (que possua CarPlay, por exemplo)

Outro:  \_\_\_\_\_



28/01/2020

Usos e Interações com Assistentes Pessoais Digitais

9. Qual sua percepção sobre a utilização de uma interface de voz? \*

*Marcar apenas uma oval.*

- Gosto muito  
 Gosto  
 Sou neutro  
 Não gosto  
 Não gosto nada

10. Como você se sente ao utilizar a linguagem falada com um objeto? \*

*Marcar apenas uma oval.*

- Ok, acho normal  
 Não me importo  
 Me sinto um pouco estranho  
 Acho divertido  
 Acho irritante  
 Nunca pensei sobre isso  
 Outro: \_\_\_\_\_

11. Durante a interação com as assistentes pessoais, você percebe alguma alteração na sua forma de falar normalmente? Marque todas as opções que corresponderem a sua percepção. \*

*Marque todas que se aplicam.*

- Não  
 Nunca pensei sobre isso  
 Sim, falo mais devagar  
 Sim, faço frases mais diretas  
 Sim, evito demonstrar qualquer tipo de sotaque  
 Sim, sou menos educado/cuidadoso ao falar  
 Sim, sou mais educado/cuidadoso ao falar  
 Sim, sou mais consciente da forma como eu falo e me expresso

Outro:  \_\_\_\_\_

28/01/2020

Usos e Interações com Assistentes Pessoais Digitais

12. Você tem preferência por um gênero específico de voz na utilização da assistente pessoal virtual? \*

*Marcar apenas uma oval.*

- Não *Pular para a pergunta 14*
- Sim, masculino
- Sim, feminino
- Preferia que fosse uma voz neutra
- Outro: \_\_\_\_\_

13. Qual o motivo de sua preferência por esse gênero?

---

---

---

---

---

14. Você acha que o gênero da voz da sua assistente pessoal virtual modifica de alguma forma a sua interação? \*

*Marcar apenas uma oval.*

- Sim, muito
- Sim, um pouco
- Não sei
- Talvez modifique
- Não, acredito que não faça diferença alguma
- Nunca refleti a respeito
- Outro: \_\_\_\_\_

28/01/2020

Usos e Interações com Assistentes Pessoais Digitais

15. Marque os atributos que levam você a utilizar uma assistente pessoal virtual: \*

*Marque todas que se aplicam.*

- Capacidade de integração com outros dispositivos inteligentes
- Possibilidade de interagir enquanto executo outras tarefas
- Possibilidade de não utilizar botões e outros comandos na tela
- Facilidade de executar tarefas através da fala
- Possibilidade de manter minhas mãos livres
- Possibilidade de manter minha atenção visual em outra coisa (como dirigir, por exemplo)
- Possibilidade de falar e ouvir ao invés de ler e escrever

Outro:  \_\_\_\_\_

16. Você já sentiu algum sentimento em relação a uma assistente pessoal virtual?  
Marque todas as opções que corresponderem a sua experiência. \*

*Marque todas que se aplicam.*

- Não
- Sim, raiva
- Sim, frustração
- Sim, hostilidade
- Sim, medo
- Sim, estranhamento
- Sim, surpresa
- Sim, humor
- Sim, empatia
- Sim, afeição

Outro:  \_\_\_\_\_

17. Você diria que desenvolveu uma relação com a sua assistente pessoal virtual?

*Marcar apenas uma oval.*

- Sim
- Não
- Outro: \_\_\_\_\_

28/01/2020

Usos e Interações com Assistentes Pessoais Digitais

18. Alguma vez a sua assistente pessoal virtual já foi acionada de forma acidental? \*

*Marcar apenas uma oval.*

- Sim, com frequência
- Sim, algumas vezes
- Sim, uma vez
- Não

19. Você pensa em sua assistente pessoal virtual como sendo um computador? \*

*Marcar apenas uma oval.*

- Sim
- Não

20. Você se preocupa com questões relacionadas ao modo de funcionamento dos sistemas das assistentes pessoais virtuais? \*

*Marcar apenas uma oval.*

- Sim, inclusive li e/ou pesquisei sobre isso
- Sim, muito
- Sim, um pouco
- Nunca pensei sobre isso
- Não, mas tenho conhecimento sobre a forma de funcionamento
- Não

28/01/2020

Usos e Interações com Assistentes Pessoais Digitais

21. Você se preocupa com questões relacionadas à privacidade na utilização desses sistemas? \*

*Marcar apenas uma oval.*

- Sim, muito
- Sim, um pouco
- Sim, mas não me impede de utilizar
- Nunca pensei sobre isso
- Não

22. Você, alguém da sua família ou conhecido utiliza uma assistente pessoal como forma de tecnologia assistiva? \*

*Marcar apenas uma oval.*

- Sim
- Não *Pular para a pergunta 24*
- Não sei informar *Pular para a pergunta 24*

23. Você pode explicar para qual finalidade a assistente pessoal é utilizada como tecnologia assistiva?

---

---

---

---

---

Perfil

24. Qual a sua idade? \*

---

28/01/2020

Usos e Interações com Assistentes Pessoais Digitais

25. Qual seu gênero? \*

*Marcar apenas uma oval.*

- Feminino
- Masculino
- Outro: \_\_\_\_\_

26. Qual seu nível de escolaridade? \*

*Marcar apenas uma oval.*

- Ensino Fundamental *Pular para a pergunta 28*
- Ensino Médio *Pular para a pergunta 28*
- Graduação Incompleta
- Graduação
- Mestrado Incompleto
- Mestrado
- Doutorado Incompleto
- Doutorado
- Outro: \_\_\_\_\_

27. Em caso de formação universitária, especifique sua área.

\_\_\_\_\_

Finalização

28/01/2020

Usos e Interações com Assistentes Pessoais Digitais

28. Você gostaria de acrescentar algo ou contar algo interessante sobre suas interações com uma assistente pessoal virtual?

---

---

---

---

---

29. Você tem interesse em participar da fase qualitativa desta pesquisa, em uma entrevista em profundidade em Porto Alegre ou de forma online? Em caso positivo, deixe seu e-mail para contato.

---

---

Este conteúdo não foi criado nem aprovado pelo Google.

Google Formulários

## **ANEXO 2 – TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO DO FORMULÁRIO**

### **UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL FACULDADE DE BIBLIOTECONOMIA E COMUNICAÇÃO PROGRAMA DE PÓS-GRADUAÇÃO EM COMUNICAÇÃO E INFORMAÇÃO**

#### **TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO**

**PESQUISA:** “I’m not sure what you’ve heard but digital assistants have feeling too”: um estudo das práticas culturais em torno dos sistemas de inteligência artificial a partir das assistentes pessoais digitais

**PESQUISADORES:** Suely Fragoso e Luiza Santos

Você está sendo convidado para participar desta pesquisa que tem por objetivo compreender as formas de interação entre seres humanos e dispositivos baseados em Inteligência Artificial, nomeadamente as Assistentes Pessoais Virtuais (Siri, Cortana, Alexa e Google Assistant). Este projeto de pesquisa foi aprovado pela Comissão de Pesquisa da Faculdade de Biblioteconomia e Comunicação da Universidade Federal do Rio Grande do Sul.

A metodologia utilizada nesta pesquisa se foca na aplicação de um formulário quantitativo e anônimo e na realização de entrevistas em profundidade com usuários destes dispositivos. Este TCLE diz respeito a primeira destas etapas, o formulário quantitativo. Para responder o questionário, o participante deve ter utilizado, ao menos uma vez, alguma das assistentes pessoais digitais (Siri, Cortana, Google Assistant ou Alexa).

No formulário, serão solicitados alguns dados de gerais de perfil e perguntas específicas sobre as formas de interação e percepção que o usuário possui das assistentes pessoas virtuais. As respostas são confidenciais e serão analisadas de forma conjunta com os demais respondentes, sem qualquer tipo de identificação dos participantes. Os dados de pesquisa serão armazenados por um total de seis anos após a obtenção, sob responsabilidade da pesquisadora, em local e com manutenção adequada. Após esse período, estes dados serão descartados.



A participação nesta pesquisa não traz complicações legais de nenhuma ordem e os procedimentos utilizados obedecem aos critérios da ética na Pesquisa com Seres Humanos conforme a Resolução número 510/2016 do conselho Nacional de Saúde. Nenhum dos procedimentos utilizados oferece riscos à sua dignidade. Esta etapa da pesquisa não envolve risco para os envolvidos. Como benefício desta pesquisa, para os participantes, citamos: possibilidade de reflexão sobre os hábitos de consumos relacionados aos dispositivos investigados; compreensão, após a publicação do trabalho, dos processos envolvidos nos dispositivos de inteligência artificial e das dinâmicas relacionais dos seres humanos com estes dispositivos.

Ao continuar o preenchimento deste formulário você autoriza a utilização de suas respostas para o estudo. Em caso de dúvidas sobre a pesquisa, o participante pode entrar em contato com a pesquisadora Luiza Santos, pelo telefone (51)996015151, com a pesquisadora responsável Prof. Dra. Suely Fragoso, pelo telefone (51) 99984004, com a Comissão de Pesquisa da Faculdade de Biblioteconomia e Comunicação, pelo telefone (51) 3308.5393 ou com o CEP-UFRGS, pelo telefone (51) 33083738 ou no endereço: Av. Paulo gama, 110 – Sala 321, Prédio Anexo 1 da Reitoria – Campus Centro, Porto Alegre/RS – CEP: 90040-060.

**ANEXO 3 – DADOS BRUTOS DO FORMULÁRIO**

Disponível em: <http://bit.ly/dadosbrutosform>



**ANEXO 4 – MODELO DE CONVITE POR E-MAIL E/OU REDES SOCIAIS DIGITAIS****Modelo de convite por e-mail e/ou redes sociais digitais**

Prezados,

Este contato tem por intuito a divulgação e convite para participar de etapa qualitativa de pesquisa em andamento à nível de doutorado no Programa de Comunicação e Informação da Universidade Federal do Rio Grande do Sul. A pesquisa em questão tem por objetivo compreender as formas de interação entre seres humanos e dispositivos baseados em Inteligência Artificial, nomeadamente as Assistentes Pessoais Virtuais (Siri, Cortana, Alexa e Google Assistant). Este projeto de pesquisa foi aprovado pela Comissão de Pesquisa da Faculdade de Biblioteconomia e Comunicação da Universidade Federal do Rio Grande do Sul.

A metodologia utilizada nesta pesquisa se foca na realização de entrevistas semiestruturadas em profundidade com usuários destes dispositivos. Os voluntários devem ser maiores de 18 anos de idade e ter disponibilidade de ir até a Faculdade de Biblioteconomia e Comunicação da UFRGS, onde as entrevistas serão realizadas. Na entrevista, serão solicitados alguns dados de gerais de perfil e perguntas específicas sobre as formas de interação e percepção que o usuário possui das assistentes pessoas virtuais. Lembramos que os dados serão tratados de forma anônima no texto final da pesquisa, portanto nenhum participante será identificado. Além disso, as informações obtidas durante as entrevistas são de caráter confidencial.

Solicitamos a divulgação deste chamado através de suas redes de contato, a fim de encontrar uma maior variedade de participantes dispostos a se engajar com a pesquisa. A participação é de caráter voluntário e nenhum tipo de remuneração ou ressarcimento de custos de deslocamento será fornecido aos participantes. Quem tiver interesse em participar desta pesquisa, pode escrever para as pesquisadoras responsáveis Luiza Santos, e-mail: [luizacdsantos@gmail.com](mailto:luizacdsantos@gmail.com), e Suely Fragoso, e-mail: [suelyfragoso@ufrgs.br](mailto:suelyfragoso@ufrgs.br), para que possamos combinar data e horário para a entrevista. Alternativamente, a pesquisadora Luiza Santos também pode ser contata pelo número de celular (51) 996015151.

Desde já agradecemos,

Luiza Santos e Suely Fragoso

## **ANEXO 5 – TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO DA ENTREVISTA**

### **UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL FACULDADE DE BIBLIOTECONOMIA E COMUNICAÇÃO PROGRAMA DE PÓS-GRADUAÇÃO EM COMUNICAÇÃO E INFORMAÇÃO**

#### **TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO**

**PESQUISA:** “I’m not sure what you’ve heard but digital assistants have feeling too”: um estudo das práticas culturais em torno dos sistemas de inteligência artificial a partir das assistentes pessoais digitais

**PESQUISADORES:** Suely Fragoso e Luiza Santos

Você está sendo convidado para participar desta pesquisa que tem por objetivo compreender as formas de interação entre seres humanos e dispositivos baseados em Inteligência Artificial, nomeadamente as Assistentes Pessoais Virtuais (Siri, Cortana, Alexa e Google Assistant). Este projeto de pesquisa foi aprovado pela Comissão de Pesquisa da Faculdade de Biblioteconomia e Comunicação da Universidade Federal do Rio Grande do Sul.

A metodologia utilizada nesta pesquisa se foca na realização de entrevistas em profundidade com usuários destes dispositivos. É prevista cerca de uma hora de duração para a entrevista e o participante possui o direito e liberdade de recusar-se a participar da pesquisa ou retirar o seu consentimento em qualquer etapa da mesma. Na entrevista, serão solicitados alguns dados de gerais de perfil e perguntas específicas sobre as formas de interação e percepção que o usuário possui das assistentes pessoas virtuais. Essa entrevista será gravada com o auxílio de um telefone celular e será armazenada de forma anônima, através da atribuição de um número. Estes dados serão tratados de forma anônima no texto final da pesquisa, portanto nenhum participante será identificado. As informações obtidas durante as entrevistas são de caráter confidencial.

As entrevistas serão realizadas na Faculdade de Biblioteconomia e Comunicação da Universidade Federal do Rio Grande do Sul, na sala do Laboratório de Pesquisa em Artefatos Digitais ou em sala de aula do mesmo prédio, preparada adequadamente para isso. Os dados de pesquisa serão armazenados por um total de seis anos após a obtenção, sob responsabilidade da pesquisadora, em local e com manutenção adequada. Após esse período, estes dados serão

descartados. O participante tem direito de requerer as transcrições da entrevista, para conhecimento, possuindo este direito mesmo em caso de retirada do consentimento de participação na pesquisa.

A participação nesta pesquisa não traz complicações legais de nenhuma ordem e os procedimentos utilizados obedecem aos critérios da ética na Pesquisa com Seres Humanos conforme a Resolução número 510/2016 do conselho Nacional de Saúde. Nenhum dos procedimentos utilizados oferece riscos à sua dignidade. Os riscos envolvidos nesta pesquisa são considerados de nível baixo, sendo os seguintes: cansaço durante a entrevista em profundidade; constrangimento ao responder alguma das questões da entrevista para o pesquisador, ainda que o conteúdo da entrevista não seja considerado de conteúdo sensível. Em caso de desconforto causado por cansaço, o participante pode solicitar o interrompimento momentâneo ou permanente da entrevista. Em caso de desconforto ou constrangimento causado por alguma pergunta e/ou resposta, o entrevistado pode solicitar ao entrevistador a anulação daquela pergunta e/ou resposta ou o interrompimento momentâneo ou permanente da entrevista.

Como benefício desta pesquisa, para os participantes, citamos: possibilidade de reflexão sobre os hábitos de consumos relacionados aos dispositivos investigados; compreensão, após a publicação do trabalho, dos processos envolvidos nos dispositivos de inteligência artificial e das dinâmicas relacionais dos seres humanos com estes dispositivos. Em conformidade com a resolução 510/2016, os participantes serão ressarcidos por seus custos de deslocamento (nos trechos de ida e volta) para participação da pesquisa no mesmo dia em que comparecerem para a entrevista.

O participante receberá uma cópia assinada por todos os envolvidos deste Termo de Consentimento Livre e Esclarecido. Em caso de dúvidas sobre a pesquisa, o participante pode entrar em contato com a pesquisadora Luiza Santos, pelo telefone (51)996015151, com a pesquisadora responsável Prof. Dra. Suely Fragoso, pelo telefone (51) 99984004, com a Comissão de Pesquisa da Faculdade de Biblioteconomia e Comunicação, pelo telefone (51) 3308.5393 ou com o CEP-UFRGS, pelo telefone (51) 33083738 ou no endereço: Av. Paulo gama, 110 – Sala 321, Prédio Anexo 1 da Reitoria – Campus Centro, Porto Alegre/RS – CEP: 90040-060.



---

Assinatura da Pesquisadora  
Luiza Carolina dos Santos



---

Assinatura da Pesquisadora Responsável  
Suely Fragoso

Considerando, que fui informado(a) dos objetivos e da relevância do estudo proposto, de como será minha participação, dos procedimentos e riscos decorrentes deste estudo, declaro o meu consentimento em participar da pesquisa, como também concordo que os dados obtidos na investigação sejam utilizados para fins científicos. Estou ciente que receberei uma via desse documento

Porto Alegre, \_\_\_\_ de \_\_\_\_\_ de \_\_\_\_\_.

---

Assinatura do participante

## ANEXO 6 – ROTEIRO SEMI-ESTRUTURADO PARA A ENTREVISTA

### Roteiro de Entrevista em Profundidade Semiestruturada

#### Procedimento:

- Leitura, dúvidas e assinatura do TCLE
- Conferir se o entrevistado correspondeu ao formulário quantitativo
- Solicitação de autorização para gravação do áudio (dois dispositivos)
- Data, horário e código da entrevista

#### Dados de Perfil

Nome:

Idade:

Nível de Escolaridade:

Área de Formação:

Gênero:

Sexualidade:

#### Experiência com uso de assistentes pessoais digitais

- a) Qual ou quais assistentes pessoais digitais você utiliza?
- b) Você tem mais familiaridade com alguma delas? E em termos de preferência? Porque?
- c) Com que frequência você utiliza? E desde quando você utiliza?
- d) Para qual finalidade você utiliza a assistente pessoal digital (ouvir música, pedir informações, fazer pesquisas, conversar, notícias, etc)?
- e) Como você descreveria uma assistente pessoal digital?
- f) Como você descreveria a assistente que você tem mais afinidade?
- g) Você já utilizou uma assistente pessoal digital ‘apenas’ para conversar? Se sim, você pode falar um pouco sobre isso?
- h) Você utiliza a sua assistente pessoal em quais locais da sua casa? Você faz alguma distinção entre o uso privado e o uso público delas?
- i) Você utiliza a sua assistente pessoal como forma de socialização?
- j) Você consome conteúdo relacionado as assistentes (coisas que você pode fazer, coisas engraçadas para perguntar, etc)?
- k) Qual a sua percepção geral sobre esses dispositivos?
- l) Você criou alguma rotina de uso?
- m) Me fale um pouco sobre as suas experiência de uso.

#### Familiaridade com interface de voz e inteligência artificial

- a) Você possui alguma familiaridade com o modo de funcionamento das Assistentes Pessoais Virtuais?
- b) Você possui algum conhecimento sobre algoritmos, *machine learning* ou redes neurais profundas? Se sim, como você associa as Assistentes Pessoais Virtuais a estes conceitos?
- c) O que você entende por inteligência artificial?
- d) O que você pensa sobre inteligência artificial de forma geral?
- e) Para você, a assistente é uma IA?
- f) Como você definiria o que é um robô?
- g) Você considera que as assistentes pessoais seriam um tipo de ‘robô’? Por quê?
- h) Você pode explicar como se sente ao falar com um objeto? Por que você se sente desta forma?
- i) Você percebe que fala diferente com a assistente?
- j) Você percebe problemas na sua interação com as assistentes?
- k) Já houve situações em que você ligou esse dispositivo sem querer? Ou algo estranho aconteceu? Como você reage/percebe esses acontecimentos?
- l) Você se preocupa com a utilização dos seus dados nesses sistemas?
- m) Você utiliza GPS com comando de voz? Se sim, qual sistema você utiliza e/ou prefere?
- n) Você faz alguma escolha particular pela voz de comando do GPS? Você acha que essa escolha tem alguma implicação na forma como você utiliza ou se relaciona com o dispositivo enquanto dirige?
- o) Que tipo de voz você prefere? Por quê?

### **Antropomorfismo**

- a) Você pensa em sua assistente pessoal como um computador?
- b) Você já sentiu algum sentimento em relação a esse artefato? Quais? Em quais situações?
- c) Você percebe a existência de algum traço de ‘personalidade’ na(s) assistente(s) que você utiliza? Se sim, pode me explicar quais e de que forma você percebe?
- d) Essa ‘personalidade’ influencia de alguma forma a utilização que você faz desses dispositivos?
- e) Você diria que a sua assistente se apreze com um ser humanos?
- f) E como você percebe os traços humanos desse dispositivo? Você avalia esta como uma característica positivo?
- g) Você diria que confia na sua assistente?



**Voz e gênero**

- a) Como você descreveria suas interações com esses dispositivos? Algo que você lembra ou que ficou marcado para você?
- b) Como você se sente ao utilizar uma interface de voz?
- c) Você pode comparar o seu uso de uma interface de voz em relação a interface gráfica? São usos que se sobrepõem? Você substitui um pelo outro? Utiliza para coisas diferentes?
- d) Você utiliza a Assistente Pessoal com voz feminina ou masculina?
- e) Você sabia da possibilidade de troca da voz de algumas assistentes? Já experimentou?
- f) Você acha que uma voz feminina ou masculina faz diferença na interação que você estabelece? Por quê?
- g) Você sabia que em algumas delas é possível trocar, mas que em todas a voz feminina é que vem como padrão?
- h) Por qual motivo você diria que a voz feminina é priorizada nesses sistemas?
- g) Existe algo a mais que você gostaria de acrescentar?

