

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
Instituto de Letras e Linguística
Programa de Pós-graduação em Estudos Linguísticos

FLÁVIO DE SOUSA FREITAS

**APLICATIVOS MÓVEIS DE INTERPRETAÇÃO
AUTOMÁTICA: EXPECTATIVAS E
PERCEPÇÕES DOS USUÁRIOS BRASILEIROS**

Uberlândia/MG
2022

FLÁVIO DE SOUSA FREITAS

**APLICATIVOS MÓVEIS DE INTERPRETAÇÃO
AUTOMÁTICA: EXPECTATIVAS E
PERCEPÇÕES DOS USUÁRIOS BRASILEIROS**

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Estudos Linguísticos do Instituto de Letras e Linguística da Universidade Federal de Uberlândia, como requisito para a obtenção do título de Mestre em Estudos Linguísticos.

Área de Concentração: Estudos em Linguística e Linguística Aplicada

Linha de pesquisa: Linguagem, ensino e sociedade

Orientador(a): Profa. Dra. Marileide Dias Esqueda

Uberlândia/MG

2022



UNIVERSIDADE FEDERAL DE UBERLÂNDIA
Coordenação do Programa de Pós-Graduação em Estudos
Linguísticos

Av. João Naves de Ávila, nº 2121, Bloco 1G, Sala 1G256 - Bairro Santa Mônica,
Uberlândia-MG, CEP 38400-902

Telefone: (34) 3239-4102/4355 - www.ileel.ufu.br/ppgel - secppgel@ileel.ufu.br



ATA DE DEFESA - PÓS-GRADUAÇÃO

Programa de Pós-Graduação	Estudos Linguísticos				
	Dissertação de mestrado - PPGEL				
Data:	Trinta e um de janeiro de dois mil e vinte e dois	Hora de início:	14:00	Hora de encerramento:	17:00
Matrícula do Discente:	11922ELI008				
Nome do Discente:	Flávio de Sousa Freitas				
Título do Trabalho:	Aplicativos móveis de interpretação automática: expectativas e percepções de usuários brasileiros				
Área de concentração:	Estudos em Linguística e Linguística Aplicada				
Linha de pesquisa:	Linguagem, Ensino e Sociedade				
Projeto de Pesquisa de vinculação:	Ensino de Tradução: Culturas Pedagógicas				

Reuniu-se, por videoconferência, a Banca Examinadora, designada pelo Colegiado do Programa de Pós-graduação em Estudos Linguísticos, assim composta: Professores Doutores: Lucas Nunes Vieira - University of Bristol; [Ariel Novodvorski - UFU](#); e [Marileide Dias Esqueda - UFU](#), orientadora do candidato.

Iniciando os trabalhos a presidente da mesa, Dra. Marileide Dias Esqueda, apresentou a Comissão Examinadora e o candidato, agradeceu a presença do público, e concedeu ao Discente a palavra para a exposição do seu trabalho. A duração da apresentação do Discente e o tempo de arguição e resposta foram conforme as normas do Programa.

A seguir o senhor(a) presidente concedeu a palavra, pela ordem sucessivamente, aos(às) examinadores(as), que passaram a arguir o(a) candidato(a). Ultimada a arguição, que se desenvolveu dentro dos termos regimentais, a Banca, em sessão secreta, atribuiu o resultado final, considerando o(a) candidato(a):

Aprovado.

Esta defesa faz parte dos requisitos necessários à obtenção do título de Mestre.

O competente diploma será expedido após cumprimento dos demais requisitos, conforme as normas do Programa, a legislação pertinente e a regulamentação interna da UFU.

Nada mais havendo a tratar foram encerrados os trabalhos. Foi lavrada a presente ata que após lida e achada conforme foi assinada pela Banca Examinadora.



Documento assinado eletronicamente por **Lucas Nunes Vieira, Usuário Externo**, em 31/01/2022, às 16:36, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Marileide Dias Esqueda, Professor(a) do Magistério Superior**, em 31/01/2022, às 17:09, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Ariel Novodvorski, Professor(a) do Magistério Superior**, em 31/01/2022, às 20:59, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site https://www.sei.ufu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **3336733** e o código CRC **D2D56D3F**.

Referência: Processo nº 23117.005544/2022-70

SEI nº 3336733

Dados Internacionais de Catalogação na Publicação (CIP)
Sistema de Bibliotecas da UFU, MG, Brasil.

F866
2022 Freitas, Flávio de Sousa, 1996-
Aplicativos móveis de interpretação automática [recurso eletrônico] :
expectativas e percepções dos usuários brasileiros / Flávio de Sousa
Freitas. - 2022.

Orientadora: Marileide Dias Esqueda.
Dissertação (Mestrado) - Universidade Federal de Uberlândia.
Programa de Pós-Graduação em Estudos Linguísticos.
Modo de acesso: Internet.
Disponível em: <http://doi.org/10.14393/ufu.di.2022.5018>
Inclui bibliografia.
Inclui ilustrações.

1. Linguística. I. Esqueda, Marileide Dias, 1973-, (Orient.). II.
Universidade Federal de Uberlândia. Programa de Pós-Graduação em
Estudos Linguísticos. III. Título.

CDU: 801

André Carlos Francisco
Bibliotecário - CRB-6/3408

DEDICATÓRIA

Dedico este trabalho a todos que acreditam na ciência e repudiam os dogmas.

AGRADECIMENTO

Agradeço primeiramente a mim mesmo, por ter acreditado que eu seria capaz, ainda que várias pessoas tentassem minar a minha autoestima e diminuir as minhas capacidades. Agradeço ao meu avô Miguel de Freitas, que tanto me apoiou ao longo dos anos. Agradeço também a minha orientadora e a todos os demais doutores que tiveram contato com a minha pesquisa.

RESUMO

Apesar dos inúmeros sistemas desenvolvidos ao longo das cerca de quatro décadas de história da interpretação automática, pouco se sabe sobre as expectativas e percepções que os usuários dessa tecnologia têm a respeito desse serviço. Tendo isso em vista, o presente estudo busca identificar as percepções dos usuários brasileiros de serviços de interpretação automática oferecidos por aplicativos móveis da App Store, a loja de aplicativos móveis da Apple. E, a partir dessa identificação, trata brevemente das expectativas desses usuários. Os resultados mostram que vários comentários veiculam expectativas e percepções contraditórias e, por mais que o usuário as separe em períodos, o fato de essas expectativas e percepções comporem um mesmo comentário pode levar o pesquisador a ter dificuldades com a interpretação delas. As expectativas e percepções dos usuários apontam para a necessidade de os aplicativos melhorarem os serviços de interpretação automática por eles oferecidos, principalmente no que se refere à tradução de textos em imagens oferecida pelo Google Tradutor e a funcionalidade de tradução *offline* oferecida pelo aplicativo Tradutor de Voz. Além disso, os dados do presente estudo mostram que as expectativas e percepções dos usuários merecem atenção por parte tanto dos desenvolvedores de aplicativos de interpretação automática quanto dos intérpretes humanos, pois os parâmetros que devem ser levados em consideração no julgamento da qualidade de traduções e interpretações automáticas nem sempre estão ao alcance daquele usuário que nunca consumiu traduções e interpretações.

Palavras-chave: interpretação automática; expectativas e percepções dos usuários; tradução automática.

ABSTRACT

Despite the numerous systems developed over the nearly four-decade history of speech-to-speech translation, little is known about the expectations and perceptions that users have about this service. The present study seeks to identify the perceptions of Brazilian users of speech-to-speech services offered by mobile applications from the App Store, the Apple's mobile application store. And, from this identification, this study briefly analyzes the user's expectations. The results show that several comments shared by the users convey contradictory expectations and perceptions and, no matter how much the user separates them into periods, the fact that these expectations and perceptions compose the same comment can lead the observer to have difficulties with their interpretation. Users' expectations and perceptions point to the need for applications to improve the services offered by them, especially regarding the translation of texts in images offered by Google Translate and the offline translation functionality offered by Voice Translator. In addition, the data show that user expectations and perceptions deserve attention from both speech-to-speech app developers and human interpreters, as the parameters that must be considered when judging the quality of automatic translations and interpretations are not always within the reach of that user who has never consumed that kind of service.

Keywords: speech-to-speech translation; users' expectations and perceptions; machine translation.

SUMÁRIO

CAPÍTULO 1: INTRODUÇÃO	1
CAPÍTULO 2 - BREVE HISTÓRICO DA INTERPRETAÇÃO AUTOMÁTICA	6
2.1 Do primeiro protótipo aos sistemas e aplicativos de IA atuais	6
2.2 A linha de produção: como funcionam os sistemas de interpretação automática?	13
CAPÍTULO 3 – DO DESIGN ÀS EXPECTATIVAS E PERCEPÇÕES DOS USUÁRIOS	18
3.1 Aspectos de <i>design</i>	18
3.2 Questões relevantes sobre interface humana	22
3.3 Expectativas e percepções dos usuários.....	23
CAPÍTULO 4 – ASPECTOS METODOLÓGICOS: DA DESCRIÇÃO DOS APLICATIVOS À ANÁLISE DAS EXPECTATIVAS E PERCEPÇÕES DOS USUÁRIOS	29
4.1 Os aplicativos selecionados para o estudo	29
4.2 A descrição dos aplicativos	34
4.3 Análise das expectativas dos usuários.....	35
4.3.1 Materiais e métodos de análise	36
CAPÍTULO 5 – OS APLICATIVOS E SEUS RECURSOS	39
5.1. Informações gerais	39
5.1.1 Datas de lançamento	40
5.1.2 Disponibilidade	40
5.1.3 Módulos de ASR, TA e TTS	41
5.1.4 Atualizações	42
5.1.5 Tamanho dos aplicativos.....	43
5.1.6 Passo a Passo de utilização e suporte.....	44
5.2 Configurando os aplicativos	44
5.2.1 Autorizações.....	44
5.2.2 Planos e assinaturas.....	46
5.2.3 Línguas e Pares linguísticos	48
5.2.4 Abrangência da interpretação automática.....	49
5.2.5 Personalização da voz sintetizada	52
5.2.6 Interpretações <i>offline</i>	52
5.2.7 Configurações de fala, sons e reprodução das vozes sintetizadas	53
5.2.8 Mostre-me e tradução de vozes de animais	54
5.3 Utilizando os aplicativos	55
5.3.1 Guias de conversação	57

5.3.2 Traduções e reconhecimento automático alternativos	57
5.3.3 Recursos para a aprendizagem de línguas	58
5.3.4 Outros recursos.....	59
CAPÍTULO 6 – AS EXPECTATIVAS E PERCEPÇÕES DOS USUÁRIOS	60
6.1 Percepções positivas e percepções negativas	63
6.2 Percepções sobre o desempenho.....	65
6.3 Percepções sobre a TA/IA.....	67
6.4 Percepções sobre o módulo ASR.....	70
6.5 Percepções sobre planos e assinaturas.....	72
6.6 Percepções sobre a utilidade.....	74
6.7 Percepções sobre outros aplicativos.....	76
6.8 Percepções sobre o módulo de TTS	78
6.9 Percepções sobre a interface.....	80
6.10 Percepções sobre o modo <i>offline</i>	82
6.11 Percepções sobre a tradução de texto em imagens	84
Considerações Finais.....	86
Referências.....	87

LISTA DE ILUSTRAÇÕES

Figura 1 - Arquitetura típica de um sistema de IA bidirecional	16
Figura 2 - Esquema das experiências e percepções	24
Figura 3 - Engajamento dos usuários	33
Figura 4 - Interface da ferramenta Taguette	37
Figura 5 - Modo de visualização highlights	38
Figura 6 - Acesso ao microfone.....	45
Figura 7 - Inconsistências no painel de planos e assinaturas.....	47
Figura 8 - Mensagem sobre recurso ainda não implementado	48
Figura 9 - Controle de velocidade e gênero da voz	52
Figura 10 - Recurso "mostre-me"	54
Figura 11 - Interpretação das vozes de cachorros e gatos	55
Figura 12 - Leiaute dos aplicativos.....	56
Figura 13 – Guia de conversação do iTranslate	57
Figura 14 - Reconhecimento automático alternativo.....	58
Figura 15 - Conjugação verbal	58
Figura 16 - Comentários por ano.....	61
Figura 17 – Distribuição das percepções	63
Figura 18 - Comentário de um usuário do Google Tradutor	64
Figura 19 – Distribuição das categorias PP e PN por aplicativo	65
Figura 20 - Distribuição da categoria PSD por aplicativo.....	66
Figura 21 – Distribuição das percepções negativas sobre o desempenho por aplicativo.....	66
Figura 22 – Distribuição das percepções positivas sobre o desempenho por aplicativo.....	67
Figura 23 - Distribuição da categoria PSTA/IA por aplicativo	68
Figura 24 – Distribuição das percepções negativas sobre a TA/IA por aplicativo.....	69
Figura 25 – Distribuição das percepções positivas sobre a TA/IA por aplicativo	69
Figura 26 - Distribuição da categoria PSMASR por aplicativo	70
Figura 27 – Distribuição das percepções negativas sobre o módulo ASR	71
Figura 28 - Percepções positivas sobre o módulo ASR.....	71
Figura 29 - Distribuição da categoria PSPA por aplicativo	72
Figura 30 – Distribuição das percepções negativas sobre os planos e assinaturas por aplicativo	73
Figura 31 – Distribuição das percepções positivas sobre os planos e assinaturas por aplicativo	73
Figura 32 - Distribuição da categoria PSU por aplicativo.....	74
Figura 33 – Distribuição das percepções positivas sobre a utilização por aplicativo.....	75
Figura 34 - Distribuição da categoria PSOA por aplicativo.....	76
Figura 35 – Distribuição das percepções negativas sobre outros aplicativos por aplicativo	77
Figura 36 – Distribuição das percepções positivas sobre outros aplicativos por aplicativo	77
Figura 37 - Distribuição da categoria PSMTTS por aplicativo.....	78
Figura 38 – Distribuição das percepções negativas sobre o módulo TTS por aplicativo.....	79
Figura 39 – Distribuição das percepções positivas sobre o módulo TTS por aplicativo.....	79
Figura 40 - Distribuição da categoria PSI por aplicativo	80
Figura 41 – Distribuição das percepções positivas sobre a interface por aplicativo.....	81

Figura 42 – Distribuição de percepções negativas sobre a interface por aplicativo.....	82
Figura 43 - Distribuição da categoria PSMO por aplicativo	83
Figura 44 – Distribuição das percepções negativas sobre o modo offline por aplicativo	83
Figura 45 - Distribuição da categoria PSTTI por aplicativo	84
Figura 46 – Distribuição das percepções negativas sobre a tradução de texto em imagens por aplicativo.....	85
Figura 47 – Distribuição das percepções positivas sobre a tradução de texto em imagens por aplicativo.....	85

LISTA DE TABELAS

Tabela 1 - Especificações técnicas do iPhone	30
Tabela 2 - Aplicativos com mais de mil classificações	31
Tabela 3 - Número de classificações e número de comentários dos aplicativos	32
Tabela 4 - Informações sobre os aplicativos selecionados	34
Tabela 5 - Data de Lançamento dos aplicativos na App Store	40
Tabela 6 - Disponibilidade dos aplicativos	41
Tabela 7 - Servidores de tradução automática	42
Tabela 8 - Atualizações em 2019	42
Tabela 9 - Tamanhos dos aplicativos e serviços oferecidos	43
Tabela 10 - Valores e número de serviços oferecidos	46
Tabela 11 - Pares linguísticos	48
Tabela 12 – Interpretação automática e variantes linguísticas	50
Tabela 13 - Distribuição das línguas	51
Tabela 14 - Línguas presentes em somente um aplicativo	51

LISTA DE ABREVIATURAS E SIGLAS

AIIC – Associação Internacional dos Intérpretes de Conferência
ALPAC – Automatic Language Processing Advisory Committee
API – Application Programming Interface
ASR – Automatic speech recognition
ASTAP – Asia-Pacific Telecommunity Standardization Program
ATR – Advanced Telecommunications Research
CMT – Center of machine translation
CMU – Carnegie Mellon University
C-STAR – Consortium for Speech Translation Advanced Research
EUA - Estados Unidos da América
IA – Interpretação automática
L₁ – Língua em que o discurso origina
L₂ – Língua para a qual o discurso é traduzido
LA – Linguística Aplicada
L_{ai} – Língua artificial intermediária
L_i – Língua inglesa
NEC - Nippon Electric Company
ONU - Organização das Nações Unidas
PLN – Processamento de Linguagem Natural
PN – Percepções negativas
PP – Percepções positivas
PSD – Percepções sobre o desempenho
PSI - Percepções sobre a interface
PSMASR – Percepções sobre o módulo de ASR
PSMO - Percepções sobre o modo offline
PSMTTS - Percepções sobre o módulo de TTS
PSOA - Percepções sobre outros aplicativos
PSPA - Percepções sobre planos e assinaturas
PSTA/IA – Percepções sobre a TA/IA

PSTTI - Percepções sobre a tradução de texto em imagens

PSU - Percepções sobre a utilidade

RAM – Random Access Memory

SL-TRANS – Spoken Language Translation System

TA – Tradução automática

TTS – Text-to-speech

UE – União Europeia

UIT – União Internacional de Telecomunicações

CAPÍTULO 1: INTRODUÇÃO

As últimas décadas não teriam sido as mesmas se a difusão massiva das tecnologias da informação e comunicação não tivesse se tornado um dos imperativos da Globalização. Em um mundo cada vez mais integrado, a tradução e a interpretação¹ são serviços essenciais para a manutenção das trocas socioeconômicas. Formar, equipar e contratar profissionais dessas áreas, entretanto, é um empreendimento cujo custo pode não estar ao alcance do orçamento de muitas empresas e instituições governamentais. Destarte, o desenvolvimento de sistemas e aplicativos que automatizam a tradução e a interpretação é uma das soluções mais promissoras para esse problema.

A interpretação automática (doravante IA²) é uma tecnologia da comunicação que traduz trechos de fala de uma língua para outra ($L_1 \rightarrow L_2$)³, de forma a viabilizar a comunicação oral entre dois ou mais indivíduos que não compartilham do mesmo sistema linguístico-cultural. Composta por três tecnologias integradas, a saber, reconhecimento automático de fala (ASR, de *automatic speech recognition*), tradução automática (TA) e síntese automática de voz (TTS, de *text-to-speech*), a IA começou a ser investigada pela Ciência da Computação na década de 1980. Desde então, inúmeros sistemas foram desenvolvidos por pesquisadores de várias instituições na Ásia, nos Estados Unidos da América (EUA) e na União Europeia (UE) (FREITAS; ESQUEDA, 2020).

Com o avanço tecnológico e o advento das lojas de aplicativos móveis, no entanto, a IA tem se tornado um serviço⁴ cada vez mais disponível. Ferramentas como o Google Tradutor e Bing Tradutor, por exemplo, foram inicialmente lançadas como tradutores automáticos. Com a adição das funcionalidades de reconhecimento automático de fala e síntese de voz essas ferramentas tornaram-se intérpretes automáticos (RAWLINSON, 2015).

Isto posto, o problema que orienta o presente estudo, que “surge de lacunas ou

¹ Neste trabalho, cada referência à “interpretação”, com inicial minúscula, diz respeito à atividade de tradução oral, ao passo que, no caso de inicial maiúscula, a referência é para a área de estudos.

² Chamo a atenção ao uso da sigla IA que aqui refere-se exclusivamente à “interpretação automática”. Toda vez que julgarmos necessário fazer referência à expressão “inteligência artificial”, que em Ciência da Computação tem recebido a mesma sigla, escrevo a expressão por extenso.

³ L_1 refere-se à língua em que dado discurso é produzido e L_2 , à língua para a qual o discurso é traduzido. Nos Estudos da Tradução e da Interpretação a L_1 também é chamada de língua-fonte (*source language*) ou língua de partida e a L_2 também é chamada de língua-alvo (*target language*) ou língua de chegada.

⁴ Ao longo deste trabalho utilizo o substantivo “serviço” tanto no singular quanto no plural. Em ambos os casos, no entanto, faço referência aos diferentes serviços de IA oferecidos pelos aplicativos móveis de IA, que são descritos no **Capítulo 5**.

conflitos em função do quadro teórico existente” (PRODANOV; FREITAS, 2013, p.34), é o seguinte: apesar dos inúmeros sistemas desenvolvidos ao longo das cerca de quatro décadas de história da IA, pouco se sabe sobre as expectativas e percepções que os usuários de aplicativos móveis de IA têm a respeito desses serviços. Geralmente, as expectativas são definidas como fenômenos que surgem antes de um indivíduo ter contato com algum serviço e as percepções, fenômenos que surgem depois do contato (JOHNSTON; CLARCK, 2002). Em meu estudo considero que as expectativas e as percepções dos usuários são fenômenos intimamente relacionados: as expectativas surgem das experiências prévias, ou seja, das percepções com outros serviços; as percepções, portanto, carregam em si as expectativas acerca de um serviço. Um usuário que compartilha, por exemplo, a percepção “achei o programa muito lento”, prontamente, compartilha a expectativa de que o programa tivesse, de algum modo, um desempenho superior.

O objetivo geral do presente estudo é identificar as percepções dos usuários brasileiros de serviços de IA oferecidos por aplicativos móveis da App Store, a loja de aplicativos móveis da Apple. E, a partir dessa identificação, tratar brevemente das expectativas desses usuários.

Em relação aos objetivos específicos:

- 1) Coletar os comentários compartilhados pública e anonimamente por usuários brasileiros desse tipo de tecnologia;
- 2) Analisar os comentários a fim de identificar as percepções neles veiculadas;
- 3) Classificar as percepções e, a partir disso, identificar as principais expectativas dos usuários.

Com base em minha experiência com sistemas e aplicativos de IA, tanto como pesquisador quanto como usuário, desde o ano de 2016, quando iniciei as minhas investigações sobre o tema e considerando que, assim como outras lojas de aplicativos, a App Store não oferece nenhum tipo de orientação ou moderação de comentários, estando, portanto, os usuários livres para expor (ou não) as suas percepções da forma que desejarem, levantei as seguintes hipóteses, que serão confirmadas após a coleta dos comentários e identificação das percepções: i. algumas percepções de um mesmo usuário poderão se mostrar contraditórias e ii. considerando o estado da arte da IA, as percepções dos usuários apontarão para a necessidade de os aplicativos melhorarem os serviços de IA por eles oferecidos.

O presente estudo possui caráter descritivo (SALDANHA; O'BRIEN, 2014), com base em dados tanto quantitativos quanto qualitativos obtidos nas páginas eletrônicas dos aplicativos de IA, hospedadas nos *hiperlinks* da App Store. De acordo com o mapeamento proposto por Williams e Chesterman (2014), a pesquisa se enquadra na área de Tradução e Tecnologia, pois tem por objeto de estudo ferramentas específicas de tradução e interpretação. Tanto do ponto de vista teórico, quanto do ponto de vista do material que descrevo e analiso, o presente estudo insere-se no contexto da Linguística Aplicada (LA).

Um estudo deste teor encontra sua justificativa social nos potenciais benefícios que o serviço de IA proporciona em um mundo cada vez mais globalizado, em que o contato entre indivíduos que não compartilham do mesmo sistema linguístico-cultural, em viagens, transações comerciais internacionais ou outros tipos de interação, torna-se fundamental para a constituição do tecido social. Além disso, de acordo com dados de um relatório da agência de pesquisas Globe Newswire (2020), estima-se que, no período de 2020 a 2025, o mercado de interpretações automáticas alcançará um recorde de crescimento de 9,4%. Investigar as expectativas dos usuários desse serviço, portanto, pode lançar luz sobre a qualidade das interpretações automáticas produzidas por aplicativos móveis.

Em relação ao mercado atual de tecnologias de processamento linguístico, o presente estudo vai ao encontro do crescente interesse pela IA, que, aliado ao surgimento dos dispositivos inteligentes (*smartphones*, *smart TVs*, *smartwatches* etc.) e popularização de aplicativos móveis, tem feito com que a tecnologia avance para além do meio acadêmico.

Do ponto de vista acadêmico, o estudo encontra sua justificativa no fato de que, apesar da constatação do aumento de pesquisas sobre a IA nas últimas décadas, principalmente em cenário internacional, as expectativas dos usuários desses serviços estão por ser investigadas, sobretudo pela LA e pelos Estudos da Tradução e da Interpretação (PÖCHHACKER, 2016). Após vários trabalhos publicados sobre o tema julgo que a existência de sistemas de IA levanta questões “com relevância social suficiente para exigirem respostas teóricas que tragam ganhos a práticas sociais e a seus participantes, no sentido de uma melhor qualidade de vida, num sentido ecológico” (ROJO, 2006, p. 258), ou seja, o tipo de questão com que a LA se preocupa. Do mesmo modo, os Estudos da Tradução e da Interpretação possuem uma vasta bagagem teórica construída ao longo de várias décadas investigando vários fenômenos linguísticos no campo da Tradução e da Interpretação (ARROJO, 1998; PÖCHHACKER, 2016).

Em relação à IA, no entanto, os trabalhos afiliados aos Estudos da Interpretação, além da morosidade, caracterizam-se por uma superficialidade sistemática. Pöchhacker (2016), por exemplo, dedica poucas linhas à discussão sobre a automação do processo de interpretação e aponta para o considerável desinteresse dos intérpretes profissionais pela IA, cuja investigação estaria relegada à Ciência da Computação e suas subáreas.

Julgo, portanto, que a relevância do presente estudo reside no fato de que ao lançar vistas sobre as expectativas dos usuários de serviços de IA em aplicativos móveis, posso contribuir para mitigar essa superficialidade e, quiçá, contribuir para a consolidação da área e para o refinamento das práticas e dos resultados de pesquisas futuras, principalmente no âmbito dos Estudos da Tradução e da Interpretação. Espero que, com o presente estudo, seja possível apresentar informações que contribuam tanto aos usuários da tecnologia de IA quanto os intérpretes humanos a se posicionarem sobre tais tecnologias. Os dados poderão também ser úteis para os desenvolvedores de tecnologias e para os formadores de intérpretes, que poderão compreender melhor questões relativas ao uso de tecnologia na interpretação de textos orais.

Além disso, posicionando-me como pesquisador e usuário desse tipo de tecnologia e, portanto, interessado em seus desdobramentos, no presente estudo dou continuidade a uma iniciativa com a qual estou trabalhando desde estudos anteriores, nos quais abordo i. o estado da arte da IA (FREITAS, 2016); ii. os conceitos, as definições e a arquitetura básica de um sistema de IA (FREITAS; ESQUEDA, 2017); e iii. a extração da intenção discursiva por sistemas de IA (ESQUEDA; FREITAS, 2019). Mais recentemente, publiquei, junto de minha orientadora, o livro Tradução e Interpretação Automáticas: Origens (FREITAS; ESQUEDA, 2020), no qual traço um histórico das tecnologias precursoras da IA, abordo os principais estudos sobre IA em escala mundial e discuto as limitações e potencialidades da tecnologia.

Esta dissertação, que se soma a todos esses trabalhos anteriormente elaborados, é composta de seis capítulos. No **Capítulo 1**, introduzo o tema, descrevo os objetivos geral e específicos, as hipóteses, o problema que orienta a pesquisa e a justificativa para o trabalho. No **Capítulo 2**, traço um breve histórico da interpretação automática e descrevo o funcionamento de sistemas de interpretação automática. No **Capítulo 3**, trato dos principais aspectos de *design* que servem de base para o desenvolvimento de sistemas de IA, exponho questões relevantes sobre interface humana e discorro sobre a investigação de expectativas e percepções de usuários.

No **Capítulo 4** exponho os materiais e a metodologia empregados na descrição dos aplicativos de IA selecionados para o estudo, o processo por meio do qual seleciono os aplicativos e os procedimentos empregados para a coleta, classificação e análise dos comentários fidedignos compartilhados pelos usuários. No **Capítulo 5**, descrevo os aplicativos de IA sobre os quais os comentários coletados versam, suas respectivas funções e vários quadros comparativos a fim de mostrar seus diferenciais mais relevantes. Por fim, no **Capítulo 6**, exponho os resultados das análises dos comentários fidedignos e discuto os dados levantados à luz do referencial teórico mobilizado nos dois primeiros capítulos.

CAPÍTULO 2 - BREVE HISTÓRICO DA INTERPRETAÇÃO AUTOMÁTICA

No presente estudo, investigo as expectativas e percepções sobre os serviços de IA em aplicativos móveis a partir de comentários fidedignos compartilhados espontaneamente por usuários brasileiros. Essa investigação retrata, ainda que indiretamente, o estado da arte dessa tecnologia. Cada avanço, cada aprimoramento na tecnologia gera impactos, tanto positivos quanto negativos, na percepção dos usuários e, por conseguinte, em suas expectativas, deixando assim transparecer o estado atual da tecnologia.

A partir desses avanços, os usuários passam a exigir padrões de qualidade superiores àqueles até então alcançados e, baseados nessas demandas, os desenvolvedores procuram meios de satisfazer as reivindicações dos usuários. Assim, as expectativas e percepções dos usuários, atreladas ao empenho dos desenvolvedores de tecnologia em oferecer serviços de qualidade cada vez mais superior, faz mover um ciclo de aprimoramentos tecnológicos. De quando em quando, o ciclo se fecha e tem-se, então, as diferentes gerações de *software* e *hardware* de uma tecnologia.

Devido ao caráter cíclico das expectativas e percepções dos usuários e o seu papel nas diferentes gerações tecnológicas, entendo ainda que as expectativas e percepções, por mais atuais que sejam, trazem consigo questões de cunho histórico. Assim, o tratamento que os pesquisadores e desenvolvedores de sistemas de IA deram aos problemas mais relevantes da tecnologia em seus primórdios encontra-se refletido nas expectativas e percepções dos usuários. À vista disso, apresento agora um breve panorama histórico da IA.

2.1 Do primeiro protótipo aos sistemas e aplicativos de IA atuais

O desenvolvimento de sistemas de IA só se tornou possível após vários anos de estudos teóricos e empíricos em LA e avanços tecnológicos nas áreas da Ciência da Computação e Linguística Computacional (FREITAS; ESQUEDA, 2020). Relativamente jovem se comparada a outras tecnologias da tradução, a IA é obra do espírito inventivo de pesquisadores de várias nacionalidades.

Inicialmente ocupando papel de pouco destaque em conferências sobre Processamento de Linguagem Natural (PLN), nas últimas três décadas, a IA tem obtido

maior visibilidade entre as comunidades científicas da Ásia, da Europa e da América do Norte. Segundo Waibel e Fügen (2008), o crescente interesse e entusiasmo entorno do tema podem ser explicados pela convergência de tecnologias com maior poder de processamento e uma crescente valorização da comunicação internacional. E essa convergência merece atenção, pois ela vai ao encontro de um dos maiores anseios por parte dos pesquisadores e desenvolvedores envolvidos com a IA desde as primeiras décadas: o desenvolvimento de sistemas capazes de realizar interpretações automáticas instantâneas e livres de erros. E como é exposto no **Capítulo 6**, esse anseio parece refletir-se nas expectativas e percepções dos usuários de IA, que valorizam o desempenho⁵ dos aplicativos acima de vários outros aspectos analisados.

Iniciada na primeira metade da década de 1980, a história da IA é dividida, por Nakamura (2009), em três gerações que correspondem às três fases do projeto para implementação de sistemas de interpretação via telefone do instituto internacional ATR, fundado em 1986 no Japão (KUREMATSU *et al.*, 1991). Em suas origens, portanto, a IA deriva de estudos e experimentos com a interpretação via telefone, serviço que chegou a ser implementado em alguns países, mas que logo perdeu fôlego.

A divisão em gerações não se refere necessariamente aos sistemas nelas desenvolvidos, mas às versões desses sistemas. Alguns sistemas de primeira geração, por exemplo, foram aprimorados e receberam versões de segunda e terceira gerações, como é o caso dos sistemas SL-TRANS (KITANO, 1991) e JANUS (WAIBEL *et al.*, 1991). Além disso, os atributos que caracterizam cada geração refletem tendências, fatos e acontecimentos políticos e sociais de escala mundial.

Desenvolvidos ao longo da década de 1980 e começo da década de 1990, os sistemas de primeira geração começaram a surgir após estudos iniciais terem sugerido a viabilidade do acoplamento dos módulos ASR, TA e TTS. Desde o início, no entanto, os pesquisadores entenderam que a viabilidade seria persistentemente desafiada pelas

⁵ Ao longo deste trabalho a questão do desempenho é mencionada diversas vezes. Para além das questões computacionais, o desempenho está relacionado à capacidade de os sistemas e aplicativos de IA produzirem interpretações de qualidade, ou seja, interpretações que contenham quantidades de inconsistências linguísticas cada vez menores. E é nesse sentido que acredito que a IA possa contribuir de forma mais evidente para os estudos sobre interpretação automática. Em meu trabalho monográfico de 2016, elaborado para a conclusão da graduação em Tradução, submeti o discurso de posse do 44º presidente dos EUA, Barack Obama, ao Google Tradutor e analisei o desempenho dessa ferramenta em relação às características linguísticas dos trechos traduzidos. A partir do experimento que realizei, constatei que o bom desempenho do Google Tradutor está relacionado à qualidade sonora do trecho submetido à ferramenta e, também, ao componente tecnológico analisado, ou seja, um mesmo trecho reconhecido pelo componente de ASR de modo eficaz pode vir a ser traduzido com algumas inconsistências linguísticas e vice-versa (FREITAS, 2016).

limitações tecnológicas da época e pela complexidade vinculada à tarefa de PLN exigida para a IA.

Apesar da escassa documentação sobre o episódio, o primeiro protótipo de IA teria sido apresentado publicamente em 1983 pela empresa japonesa NEC Corporation durante uma conferência internacional de telecomunicações que ocorreu em Genebra, na Suíça (NAKAMURA, 2009). Assim como qualquer tecnologia em estágio experimental, o protótipo possuía um desempenho muito baixo e era capaz de reconhecer, no máximo, 500 palavras (LEE, 2015).

Stentiford e Steer (1987), dois dos primeiros pesquisadores a publicarem estudos sobre o tema, mencionam três diferentes abordagens de análise linguística empregadas no desenvolvimento dos sistemas de IA de primeira geração e apresentam aquela que, segundo eles, seria a melhor solução para os erros e omissões decorrentes do ASR, que naquela época mostravam-se especialmente difíceis de serem resolvidos.

Amano, Nogami e Miike (1988) descrevem um sistema com o qual foi feito o primeiro experimento de IA via satélite. Apresentado durante uma conferência internacional de telecomunicações em 1987, o experimento possibilitou a comunicação entre oito pessoas em um escritório da empresa Toshiba, no Japão, e vários visitantes da conferência, na Suíça. O sistema descrito pelos pesquisadores não era um intérprete automático convencional, pois não possuía os módulos ASR e TTS. De acordo com a aceção dos autores, o que distingue a IA da TA é a presença de interação entre dois interlocutores, seja por meio de texto ou por meio de fala. Nos anos e décadas seguintes, todavia, o acoplamento dos módulos ASR e TTS ao módulo de TA mostrou-se tão amplamente difundido que, atualmente, a aceção que se tem da IA é praticamente inseparável da tríade ASR+TA+TTS, tornando o conceito veiculado por Amano, Nogami e Miike (1988) pouco frequente na literatura.

Tomita, Tomabechi e Saito (1990) descrevem o sistema SpeechTrans, desenvolvido pelo Centro de Tradução Automática da Universidade Carnegie Mellon (CMT/CMU), nos EUA. Apresentado pela primeira vez em junho de 1988, o sistema interpretava automaticamente trechos de fala contínua sem a necessidade de treinamento, ou seja, qualquer usuário poderia utilizá-lo sem ter que fazer pausas entre uma palavra e outra. Os sistemas anteriores exigiam que todo novo usuário tivesse que ter as características de sua voz mapeadas e inseridas nas especificações do módulo de ASR.

Responsável também pelo desenvolvimento do sistema Φ DmDialog, testado em várias ocasiões e apresentado ao público em março de 1989 (KITANO, 1994), o

CMT/CMU, fundado em 1986, foi pioneiro no lançamento de projetos de IA colaborativos. Os módulos de ASR e TTS utilizados no desenvolvimento do SpeechTrans e do Φ DmDialog, por exemplo, foram cedidos pelo instituto de pesquisas Matsushita e pela empresa Digital Equipment Corporation, respectivamente (TOMITA; TOMABECHI; SAITO, 1990). Segundo Kitano (1991), o CMT/CMU simboliza o retorno científico das pesquisas em TA nos EUA, após o relatório emitido pelo comitê ALPAC⁶, de 1966, ter quase posto fim às pesquisas sobre tradução automática naquele país.

Nos últimos anos da primeira geração começaram a surgir sistemas que realizavam interpretações indiretas por meio da construção de uma língua artificial intermediária ($L_1 \rightarrow L_{ai} \rightarrow L_2$), ou seja, uma interlíngua⁷, além do surgimento dos primeiros sistemas de domínios especializados. E, ainda que novas abordagens de TA tenham surgido nos anos finais, a abordagem baseada em regras gramaticais mostrou-se a mais utilizada (NAKAMURA, 2009).

Ao longo da década de 1980 várias instituições e centros de pesquisa investigavam a IA separadamente. Contudo, a filosofia que tem orientado as pesquisas e o desenvolvimento da tecnologia, expressa por Nakamura (2009) como “a derrubada das barreiras linguísticas” que impossibilitam a comunicação em um mundo cada vez mais globalizado, logo exigiu que os pesquisadores ampliassem seus horizontes, reunindo seus esforços em projetos colaborativos de larga escala. Assim, a partir de 1991, pesquisadores dos EUA, Alemanha, França, Japão, Coreia e Itália, começaram a constituir organizações de forma voluntária, por meio das quais eles colaboravam entre si compartilhando resultados de pesquisas independentes e estabelecendo objetivos para o desenvolvimento de sistemas de IA (SENDA, 2018).

A primeira organização desse tipo a ser formada foi o consórcio C-STAR (Consortium for Speech Translation Advanced Research), constituído por pesquisadores do instituto ATR, sediado no Japão; da CMU, nos EUA; da Universidade Técnica de Karlsruhe e da empresa de tecnologia Siemens, ambas da Alemanha. O principal objetivo do C-STAR era demonstrar a viabilidade da IA em serviços telefônicos. Segundo Gehrke e Schmidbauer (1993), os pesquisadores membros do consórcio mantinham uma relação

⁶ Para mais detalhes sobre o relatório ALPAC, conferir o HUTCHINS, John. ALPAC: the (in) famous report. *Readings in machine translation*, v. 14, p. 131-135, 2003.

⁷ Sistema linguístico desenvolvido naturalmente por aprendizes de uma segunda língua ou produzida artificialmente para representar estruturas de duas línguas (FERREIRA JUNIOR, 2007).

colaborativa e, ao mesmo tempo, competitiva. Os membros apoiavam-se mutuamente, concentrando-se em pares de línguas diferentes e competiam entre si ao adotar diferentes abordagens tecnológicas.

Finalizado em janeiro de 1993, o primeiro protótipo desenvolvido pelos pesquisadores membros do C-STAR interpretava trechos de diálogo entre pessoas que buscavam informações sobre uma conferência internacional ou que desejava inscrever-se na conferência e uma secretária do evento. Na primeira demonstração pública do protótipo, transmitida ao vivo por veículos de imprensa, pesquisadores do instituto ATR, da Siemens e da CMU tiveram suas conversas interpretadas para o alemão, inglês e japonês (GEHRKE; SCHMIDBAUER, 1993).

O surgimento das organizações de pesquisadores possibilitou a troca de conhecimento linguístico relacionado a um maior número de línguas, possibilitando que aos sistemas de primeira geração fossem adicionados novos pares linguísticos. Além disso, novas abordagens começaram a ser discutidas e aplicadas aos sistemas de primeira geração, fazendo surgir, a partir de 1993, os sistemas de segunda geração.

Desenvolvidos até o final da década de 1990, eles eram capazes de interpretar expressões utilizadas em determinados contextos e trechos de fala com alguns erros gramaticais. Pouco empregada ao longo da primeira geração, a especialização em domínios tornou-se um método amplamente aceito e aprimorado na segunda geração (RAYNER; BOUILLON; CARTER, 1995; LEVIN *et al.*, 1998).

Os sistemas ASURA, SLT e Verbmobil são exemplos de sistemas de segunda geração. Desenvolvido por pesquisadores do instituto ATR, o ASURA surgiu de algumas modificações no sistema SL-TRANS (KITANO, 1991) e a adição do par linguístico japonês-alemão. Assim como os outros sistemas desenvolvidos pelos pesquisadores membros do C-STAR, o ASURA foi testado em várias ocasiões por meio de conexões remotas que possibilitaram a comunicação entre pessoas de diferentes localidades (MORIMOTO *et al.*, 1993).

O sistema SLT (Spoken Language Translator), desenvolvido por pesquisadores suecos e ingleses, interpretava perguntas e respostas sobre viagens aéreas, do inglês para o sueco. Segundo os pesquisadores, apesar de os primeiros testes com o sistema terem sido encorajadores, muita pesquisa ainda seria necessária para que a tecnologia fosse empregada sem restrições. Erros de ASR e TA afetavam a IA em vários níveis e, a fim de conhecer melhor esses erros e poder solucioná-los, os pesquisadores organizaram em categorias os possíveis erros e o grau com que cada um deles afetava a comunicação.

Assim, mesmo que o sistema não fosse perfeito, os pesquisadores podiam investigar as soluções dos erros mais prejudiciais e, aos poucos, alcançavam melhoras no desempenho do sistema (RAYNER *et al.*, 1993).

Desenvolvido em duas fases ao longo de oito anos, o Verbmobil é um dos sistemas mais famosos e foi desenvolvido pelos membros do C-STAR. Iniciada em janeiro de 1993, a primeira fase culminou no desenvolvimento do protótipo denominado Verbmobil Demonstrator. A tecnologia de ASR do protótipo reconheceu e processou trechos de fala espontânea referentes a agendamento de reuniões de negócios e os interpretou do alemão para o inglês. Em março do mesmo ano, o protótipo foi apresentado ao público durante uma feira internacional de tecnologias que aconteceu em Hanôver, na Alemanha.

Na segunda metade de 1996, o próximo passo foi dado com o desenvolvimento de outro protótipo, apresentado na edição de 1997 da mesma feira internacional. Na ocasião, além de interpretar trechos de fala espontânea do alemão para o inglês, o novo protótipo interpretou trechos de fala espontânea do japonês para o inglês (WAHLSTER, 2013).

Iniciada em janeiro de 1997 e concluída em setembro de 2000, a segunda fase do Verbmobil resultou no desenvolvimento de um terceiro protótipo. Bem mais completo que os demais, ele foi inspecionado e finalizado em 1998. Segundo Wahlster (2000), em sua versão final, o Verbmobil podia ser acionado por comandos de voz, o que fazia com que o sistema fosse mais prático do que os sistemas desenvolvidos até então. Além disso, ele era um sistema robusto que incorporava PLN por meio de uma abordagem multimotor que fazia uso simultâneo de cinco mecanismos de tradução automática, algo até então inédito.

Ao contrário de sistemas anteriores, que traduziam os trechos de fala isoladamente, o Verbmobil, por meio de uma memória de diálogo e conhecimento de domínio, conseguia relacionar assuntos que foram citados em trechos anteriores e questões relativas ao tópico em discussão. Impelindo o processamento de fala espontânea a um nível superior, ele era capaz de gerar hipóteses de tradução para trechos que continham inconsistências e escolher a que melhor se adequava ao tópico do diálogo, interpretando o que o interlocutor tentou dizer e não o que ele realmente disse (WAHLSTER, 2013).

Um dos projetos de IA mais longo e conhecido no meio científico, os estudos desenvolvidos ao longo do projeto Verbmobil, deram origem a mais de 400 publicações (WAHLSTER, 2013). Sendo o artigo Verbmobil: Translation of face-to-face dialogs

(WAHLSTER, 1993) e o livro *Verbmobil: Foundations of Speech-to-Speech Translation* (WAHLSTER, 2013) as publicações mais citadas, de acordo com o mecanismo de busca Google Acadêmico.

Outro passo importante dado nos anos finais da segunda geração foi o estabelecimento de programas de padronização de especificações para sistemas de IA. O ASTAP (Asia-Pacific Telecommunity Standardization Program), por exemplo, iniciou suas atividades de padronização na região Ásia-Pacífico em 1998. Senda (2018) afirma que a iniciativa foi tão exitosa que, em pouco tempo, ficou claro que essa atividade de padronização deveria ser feita em escala global. As atividades do ASTAP foram então transferidas para o grupo de estudos multimídia do setor de padronização de telecomunicações da UIT (União Internacional de Telecomunicações), agência da ONU (Organização das Nações Unidas) especializada em tecnologias de informação e comunicação.

Apesar de os sistemas de primeira geração terem dado o pontapé inicial para a especialização em domínios, os sistemas de segunda geração exploraram mais a fundo o método e aprimoraram a abordagem de tradução automática baseada em exemplos. Segundo Nakamura (2009), vários dos sistemas de segunda geração utilizaram o inglês como língua intermediária.

Os sistemas de terceira geração, por sua vez, eram capazes de reconhecer trechos de fala com ruídos de fundo e traduzir expressões típicas de vários assuntos. Desenvolvidos ao longo da década de 2000, a abordagem de TA predominante nesses sistemas passou a ser a estatística e a tradução, que na primeira geração se dava por meio de uma linguagem intermediária artificial ($L_1 \rightarrow L_{ia} \rightarrow L_2$) e na segunda geração se dava por meio do inglês ($L_1 \rightarrow L_i \rightarrow L_2$), passou a ser feita de forma direta ($L_1 \rightarrow L_2$) (NAKAMURA, 2009).

Além das três gerações citadas por Nakamura (2009), com os avanços tecnológicos da última década, nomeadamente o acesso a vastas quantidades de dados disponíveis *online* por parte da indústria de tecnologia; o acesso a redes de dados móveis, por parte dos usuários; e a emergência do mercado de aplicativos móveis, a existência de uma quarta geração de sistemas de IA parece fazer sentido. Diferentemente das primeiras três gerações, definidas de acordo com as características linguísticas e fonológicas dos trechos de fala interpretados pelos sistemas desenvolvidos em cada uma delas e as abordagens de TA neles empregadas, a quarta geração destaca-se pela portabilidade,

conexão em rede e enriquecimento da interação dos seres humanos com outros seres humanos e com o mundo.

A partir da quarta geração, cujo início julgo ter acontecido na década de 2010, os sistemas de IA passaram a ser adaptados para dispositivos móveis, plataformas em que eles assumem a forma de aplicativos disponibilizados em lojas virtuais. Sob o ponto de vista da TA, os sistemas e aplicativos da quarta geração mesclam várias abordagens de TA, fazem uso de redes neurais artificiais (SPERBER; NIEHUES; WAIBEL, 2017), empregam inteligência artificial para o reconhecimento de padrões e possibilitam a tradução de trechos de fala, textos, imagens, reconhecimento de objetos em imagens e o compartilhamento de traduções em redes sociais.

A portabilidade, no entanto, é uma espada de dois gumes, pois ao mesmo tempo em que expande as possibilidades de utilização da tecnologia e facilita o seu manuseio, requer o emprego de abordagens que viabilizem o desenvolvimento de sistemas mais compactos e a existência de dispositivos móveis com processadores potentes, memória interna suficiente para o armazenamento de dados ou amplo acesso a redes de dados móveis. Apesar disso, alguns sistemas de IA adaptados para dispositivos móveis atingem desempenho semelhante ao de sistemas instalados em computadores de mesa e servidores (ZHOU *et al.*, 2003).

O primeiro sistema de IA portátil a ser comercializado foi desenvolvido em 2006 pela empresa japonesa NEC. Instalado em um PDA, ele interpretava trechos de fala do japonês para o inglês. O primeiro sistema portátil instalado em um telefone celular foi lançado em 2007 pelo instituto japonês ATR (NAKAMURA, 2009). Lançado em 2009, o Jibbigoi foi o primeiro aplicativo de IA instalado em um *smartphone*. Desenvolvido para ser utilizado em viagens, campanhas militares, serviços médicos e missões humanitárias, o Jibbigoi oferecia interpretações automáticas *offline* e possuía uma ampla variedade de idiomas (ECK *et al.*, 2010).

2.2 A linha de produção: como funcionam os sistemas de interpretação automática?

O nível tecnológico ao qual a humanidade chegou é admirável, aviões e navios cargueiros possibilitam que milhares de pessoas e mercadorias circulem em nível global todos os dias, desafiando os antigos conceitos de fronteira e realizando em poucos dias

tarefas que antes levavam meses e demandavam mão de obra muitas vezes escassa. Dispositivos móveis com conexões de dados cada vez mais velozes permitem que pessoas compartilhem suas experiências com amigos, familiares e até desconhecidos distantes a milhares de quilômetros.

A IA é uma dessas tecnologias que tem um potencial revolucionário: nutre a esperança em um futuro em que as barreiras linguísticas terão sido superadas. Um futuro em que empresas de todo mundo poderão comunicar entre si de forma livre e instantânea, trazendo prosperidade aos residentes de regiões distantes dos grandes centros de onde emana o poder. Um futuro em que turistas poderão visitar países remotos sem receio de não serem compreendidos e não poderem viver a experiência de forma plena. Dissertar sobre a IA, portanto, é lançar-se em uma jornada cheia de desafios e caminhos ainda pouco percorridos, o que não deixa de ser instigante.

Como explicitado na Introdução, os sistemas de IA, definidos como uma tecnologia da comunicação que traduz trechos de fala de uma língua para outra ($L_1 \rightarrow L_2$), são compostos por três tecnologias integradas: ASR, TA e TTS. Enquanto as pesquisas sobre TA remontam à década de 1940, as pesquisas sobre ASR começaram no início da década de 1950 e os primeiros sistemas TTS eletrônicos surgiram no final daquela mesma década.

O uso de tecnologia na prática profissional da interpretação teve seu início na década de 1920, quando o desenvolvimento de sistemas de transmissão de som eletroacústico permitiu o surgimento da interpretação simultânea (BAIGORRI-JALÓN, 1999). Nas últimas décadas do século XX, no entanto, avanços nas telecomunicações e nas técnicas de processamento de dados digitais provocaram o afluxo de novas tecnologias que transformaram profundamente a maneira com que a interpretação é praticada (PÖCHHACKER, 2016).

Em seu trabalho sobre interpretação e tecnologia, Braun (2020, p. 271) distingue quatro tipos básicos de tecnologias aplicadas à interpretação: i. tecnologias usadas para fornecer serviços de interpretação e ampliar seu alcance, como a interpretação mediada por tecnologia (*technology-mediated interpreting*) ou interpretação à distância; ii. tecnologias que podem ser aplicadas para auxiliar ou facilitar a preparação, o desempenho e o fluxo de trabalho do intérprete, como a interpretação assistida por tecnologia (*technology-supported interpreting*); iii. tecnologias desenvolvidas para substituir os intérpretes humanos; e iv. tecnologias híbridas.

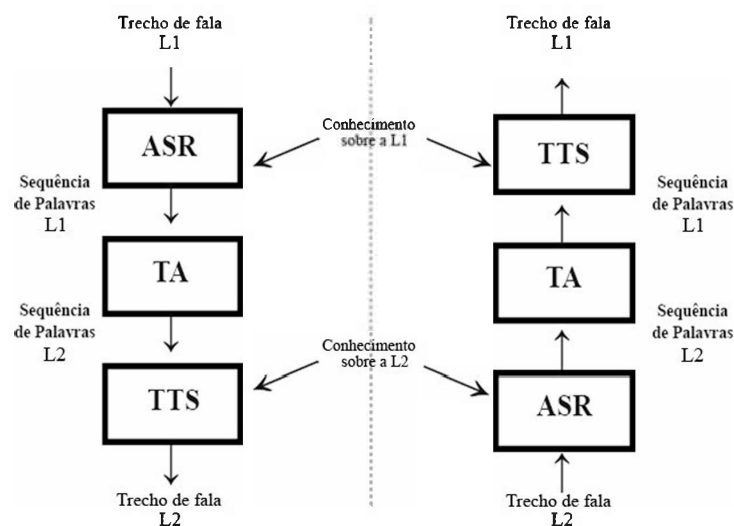
Pertencente ao terceiro tipo de tecnologia, a IA tem recebido cerca de 14 diferentes denominações na literatura sobre o tema em língua inglesa, sendo “speech translation” e “speech-to-speech translation” as mais frequentes. Em língua portuguesa, a IA também é chamada de “tradução de fala”, por pesquisadores portugueses e “tradução de voz”, por pesquisadores brasileiros (FREITAS; ESQUEDA, 2020). O termo “interpretação automática” foi empregado pela primeira vez como tradução de “speech-to-speech translation” por Freitas e Esqueda (2017).

A substituição de humanos por máquinas tem sido tema da Literatura e do Cinema desde há muito tempo. Em relação à IA, Seligman e Waibel (2019) afirmam que a ideia por trás dessa tecnologia utópica remonta às origens da computação na década de 1950. Segundo os autores, dispositivos portáteis de IA são frequentemente relacionados ao “tradutor universal” e ao “Babel fish”, dispositivos fictícios utilizados na comunicação com alienígenas, retratados pelas séries Star Trek e O Guia do Mochileiro da Galáxia, respectivamente.

Do ponto de vista linguístico, os sistemas de IA podem ser divididos em unidirecionais e bidirecionais. Os sistemas unidirecionais realizam interpretações apenas da L_1 para a L_2 , ainda que possuam mais de uma combinação de pares linguísticos (sistemas multilinguísticos). As primeiras versões do sistema Verbmobil, por exemplo, realizavam interpretações apenas do alemão e do inglês para o japonês, não possibilitando a interpretação do japonês para o alemão ou para o inglês (HAHN; ANTRUP, 1996). Os sistemas bidirecionais, por outro lado, realizam interpretações da L_1 para a L_2 e vice-versa ($L_1 \leftrightarrow L_2$).

A Figura 1 mostra os elementos que compõem a arquitetura de um sistema de IA bidirecional. Sistemas unidirecionais, por outro lado, são compostos pelos elementos de apenas um dos lados do diagrama.

Figura 1 - Arquitetura típica de um sistema de IA bidirecional



Fonte: adaptado de LEE (2015).

Organizados de forma linear, os três módulos realizam PLN semelhante a uma linha de produção industrial. Diferentemente das indústrias que transformam matérias primas em produtos, entretanto, a indústria da IA transforma enunciados da L_1 em enunciados da L_2 .

A primeira etapa da linha de produção é o ASR, processo por meio do qual o sinal acústico emitido pelo falante da L_1 , ou seja, as ondas sonoras emitidas pelo aparelho fonador daquele indivíduo, é captado e transformado em uma representação fonética das unidades mínimas constituintes das palavras. Uma vez concluída, a representação fonética é transformada em uma representação textual, também conhecida como transcrição da fala.

De posse dessa transcrição, o módulo de TA dá continuidade ao PLN de acordo com a abordagem de TA adotada. Se for a abordagem baseada em regras gramaticais, por exemplo, o módulo irá classificar as palavras da sentença conforme as regras gramaticais da L_1 . Se for a abordagem estatística, o módulo irá pesquisar em um *corpus* a ocorrência daquelas palavras e calcular seus dados estatísticos. Feito isso, o módulo de TA fará a conversão das regras gramaticais ou dos dados estatísticos da L_1 para as regras gramaticais ou dados estatísticos da L_2 . O resultado dessa conversão é chamado de tradução automática e, assim como a transcrição da fala, apresenta-se na forma de texto e geralmente é apresentado ao usuário na interface do sistema (SELIGMAN; WAIBEL, 2019).

Na etapa final, o módulo TTS recebe a tradução em formato textual e converte-a em uma representação fonética da L2. Uma vez concluída, a representação fonética é transformada em sinal acústico, ou seja, em ondas sonoras sintetizadas, que serão captadas pelo aparelho auditivo do interlocutor. De certa forma, o PLN empreendido pelos sistemas de IA também pode ser comparado a um procedimento de engenharia reversa, que conta com a captação de sinal acústico, representações fonéticas e textuais, transformações gramaticais ou estatísticas e reproduções de sinal acústico.

Assim como acontece na linha de produção de uma indústria, ao longo das etapas de processamento podem ocorrer erros e inconsistências de vários tipos. A existência desses erros, no entanto, tem levado os pesquisadores a fazerem escolhas de *design* que sejam capazes de solucionar, ou pelo menos amenizar, os impactos dessas falhas no produto final. As escolhas de *design* são muito importantes para a experiência dos usuários. No presente estudo, julgo que a explicitação dos aspectos de *design* que norteiam o desenvolvimento de sistemas de IA é crucial para a compreensão da descrição dos aplicativos e análise das expectativas dos usuários, apresentadas no **Capítulo 5** e no **Capítulo 6**, respectivamente.

CAPÍTULO 3 – DO DESIGN ÀS EXPECTATIVAS E PERCEPÇÕES DOS USUÁRIOS

Em Ciência da Computação, o *design*⁸ é definido pelo conjunto de elementos por meio dos quais um produto é configurado, concebido e elaborado para determinados fins. Trata-se de uma atividade estratégica, técnica e criativa que possibilita a produção em série ou em escala industrial (DOOLEY, 2011).

A complexidade do PLN necessário para o ASR e para a TA tem feito com que a IA progrida em etapas. Em cada uma delas, busca-se reduzir a complexidade da tarefa sob vários aspectos de *design*: arquitetura; captação e reprodução de áudio; domínio; estilo de fala; modo de interpretação; procedência dos componentes; sintaxe; velocidade; latência etc. Como foi dito, para melhorar o desempenho dos sistemas e tornar a IA uma tecnologia utilizável, os desenvolvedores têm que, necessariamente, fazer escolhas de *design*. A seguir, discorro sobre os principais aspectos *design*, com base no trabalho de Seligman e Waibel (2019).

3.1 Aspectos de *design*

Arquitetura: os sistemas de interpretação automática podem ser acoplados ao dispositivo em que eles serão utilizados ou instalados em servidores remotos. Geralmente, bom desempenho e qualidade são mais facilmente obtidos em sistemas remotos, pois, assim, recursos de computação mais robustos podem ser utilizados. Além disso, se as políticas de privacidade permitirem, os sistemas remotos também podem coletar dados de utilização que são empregados posteriormente para a melhoria do serviço. No entanto, quando o acesso a redes de dados não está disponível, é muito caro ou inseguro, os sistemas remotos tornam-se inviáveis. Para a interpretação de palestras ou para a transmissão de notícias, os sistemas remotos funcionam bem; em contrapartida, em aplicativos de interpretação automática para viagens, missões humanitárias e intervenções militares, a tecnologia acoplada é melhor.

⁸ A palavra *design* é usada nesse trabalho em referência à concepção física e funcionamento dos sistemas e aplicativos de IA.

Captação e reprodução de áudio: o uso de microfones próximos aos usuários em ligações telefônicas e palestras ou microfones de lapela em sistemas de IA melhora a qualidade da captação de áudio. Em reuniões em que há mais de uma pessoa discursando, por exemplo, pode haver sobreposição de vozes e em espaços abertos, onde há a interferência de ruídos do ambiente, provocados pela aglomeração de pessoas ou por agentes naturais como o vento, a captação de áudio pode ser comprometida. Por isso, pode ser necessário o emprego de microfones individuais e/ou a disponibilização de fones de ouvidos para os usuários.

Domínio: os sistemas de IA podem especializar-se em determinados domínios, ou seja, limitar o tema sobre o qual os usuários podem se comunicar. Sistemas de agendamento de reuniões, por exemplo, possibilitam que os usuários se comuniquem apenas para o agendamento de reuniões. Teoricamente, os usuários podem dizer o que quiserem, desde que seja sobre o domínio para o qual o sistema foi desenvolvido. Tornar os sistemas especializados em domínios facilita o trabalho dos desenvolvedores, pois o ASR e a TA passam a ser feitos com base nos padrões típicos do domínio. Embora esta abordagem tenha sido superada por várias outras técnicas de análise linguística, tradicionalmente, para especializar um sistema, é necessário a compilação de *corpora* relativos a um determinado tópico e, a partir disso, a aplicação de análises linguísticas específicas. O substantivo inglês “interest”, por exemplo, pode ser traduzido como “interesse”, mas em um *corpus* sobre finanças, também pode ser traduzido como “juros”.

A especialização em domínios começou a ganhar força após os primeiros testes com a IA terem demonstrado que, quanto maior o vocabulário, mais árdua a tarefa de manter a qualidade das interpretações. Este método parte do pressuposto de que apesar de os seres humanos falarem de forma espontânea e, geralmente, não conseguirem lidar com a imposição de restrições linguísticas, há várias tarefas do cotidiano em que os domínios discursivos são limitados. E explorar esses limites foi a forma encontrada para melhorar a qualidade das interpretações (RICCI, 2002).

Há sistemas, no entanto, que não possuem restrições de domínios, são os chamados sistemas de propósito geral, ou intérpretes automáticos universais. Por meio de tais sistemas, é possível se comunicar sobre qualquer tema, o que é importante para a interpretação automática de noticiários, palestras, discursos políticos, seminários e chamadas telefônicas. Para desenvolver tais sistemas, entretanto, é preciso lidar com vocabulários e conjuntos de conceitos amplos, tendo, muitas vezes, que processar fluxos de fala contínua, cuja segmentação nem sempre é regular.

Estilo de fala: para os sistemas de propósito geral, um problema fundamental é a qualidade da fala. Em um extremo está a fala espontânea e no outro, a fala elaborada⁹, entre ambas, vários graus de articulação que definem a dificuldade do reconhecimento automático do trecho. A fala de um jornalista, por exemplo, normalmente é elaborada com antecedência e lida em teleprompters e, por isso, não possui hesitações ou vícios de linguagem, sendo reconhecida com alta precisão. Palestrantes, no entanto, geralmente não leem discursos elaborados palavra por palavra e, dependendo do grau de familiaridade com o assunto ou habilidade de discursar em público, as falas de alguns deles podem apresentar-se fragmentadas e cheias de hesitações.

Modos de interpretação: assim como na interpretação humana, há dois modos básicos de IA: o modo consecutivo e o simultâneo. Na interpretação consecutiva, após falar o usuário faz uma pausa para que o sistema produza a interpretação. Na interpretação simultânea, a IA acontece enquanto a fala vai sendo articulada, sem pausas. Geralmente, a interpretação consecutiva é menos trabalhosa, pois as pausas do usuário são parâmetros utilizados pelo sistema para determinar o fim de um enunciado. Além disso, a pronúncia é geralmente mais clara e os falantes tendem a cooperar com o sistema. Na simultânea, os usuários ou quem quer que esteja falando nem sempre estão cientes da existência do sistema e, por isso, tendem a cooperar menos.

Procedência dos componentes: dado o caráter abrangente da área, os desenvolvedores de sistemas de IA podem desenvolver sistemas sem necessariamente ter que criar do zero os módulos de ASR, TA e TTS. Um desenvolvedor pode, por exemplo, construir uma interface que capture a fala, envia para um serviço de PLN (por exemplo, Nuance, Google e Microsoft) para realizar o ASR, envia o resultado para outro serviço para realizar a TA; e, finalmente, envia para um terceiro serviço de síntese de voz. Os cinco aplicativos do presente estudo, por exemplo, empregam esse tipo de abordagem.

Um sistema embutido pode ser construído de forma semelhante usando componentes licenciados. Em qualquer arquitetura, os custos de desenvolvimento e implementação podem (ou não) ser reduzido por meio de refinamentos de interface, personalização, combinação de linguagens ou plataformas etc. Segundo Seligman e Waibel (2019), esta abordagem reduz as dificuldades de entrar no mercado para desenvolvedores menores, mas cria uma dependência dos fornecedores de componentes. As vantagens da integração entre os componentes do sistema também podem ser perdidas.

⁹ Trechos de fala que foram elaborados previamente. Geralmente esses trechos são lidos ou decorados pelo orador. O oposto de fala espontânea.

Sintaxe: além das restrições de domínio, é possível limitar o sistema a determinadas frases, ou seja, limitar a sintaxe do sistema. Guias de conversação ativados por voz, por exemplo, aceitam apenas frases específicas ou algumas de suas variantes. Em casos assim, o módulo de ASR só precisa escolher uma das palavras ou frases que seja compatível com as frases que compõem o sistema. Frases muito diferentes daquelas que são aceitas, no entanto, não são interpretadas, o que torna esse tipo de sistema bastante limitado. O desenvolvimento de sistemas de IA de sintaxe restrita é menos trabalhoso que o desenvolvimento de sistemas especializados em domínios.

Velocidade e latência: nos sistemas de IA, a latência, ou seja, o tempo entre um comando dado ao sistema e a resposta a esse comando, é uma variável bastante importante, pois o intervalo entre a fala do locutor e o início da interpretação automática não pode ser muito grande. Na IA simultânea, por exemplo, a interpretação deve começar alguns segundos depois do início da fala. A velocidade com que a voz sintetizada é pronunciada e a latência dependem do tipo de material que será interpretado. Durante uma palestra, discurso parlamentar ou noticiário ao vivo, por exemplo, baixa latência é essencial. Mas se esses mesmos eventos forem gravados para visualização posterior, a interpretação automática poderá ser feita com maior latência e o material pode ser utilizado para contextualizar a interpretação.

De acordo com Seligman e Waibel (2019), a influência de todos esses aspectos no desenvolvimento de sistemas de IA torna a comparação direta entre sistemas uma tarefa difícil, não havendo uma resposta simples para a pergunta: “qual a qualidade da interpretação automática?”. Segundo os autores, levando em consideração cada aspecto de *design*, a resposta pode variar de “ótima! Problema fácil, já resolvido!” a “não é tão boa. Problema difícil, pesquisa em andamento” (p. 221). Por isso, ao definirmos os critérios para a seleção dos cinco aplicativos, tive o cuidado de incluir tanto dados quantitativos quanto qualitativos, conforme descrevo no **Capítulo 4**.

Em resposta aos desafios técnicos apresentados por cada aspecto, ao longo de sua história, a IA tem progredido em etapas, de guias de conversação ativados por voz a intérpretes automáticos de domínio especializado; de intérpretes automáticos de domínio especializado a intérpretes automáticos de propósito geral; de intérpretes automáticos de fala elaborada a intérpretes automáticos de fala espontânea, de sistemas acoplados a sistemas remotos; e de intérpretes automáticos consecutivos a intérpretes automáticos simultâneos (SELIGMAN; WAIBEL, 2019).

3.2 Questões relevantes sobre interface humana

As interfaces para sistemas de IA devem equilibrar objetivos concorrentes: por um lado, velocidade máxima e transparência (interferência mínima), e por outro, o máximo de precisão e naturalidade. Cada vez mais os pesquisadores entendem que a perfeição é algo complexo de se atingir, até porque mesmo os intérpretes humanos normalmente gastam um tempo considerável explicando determinados termos ou expressões. Enquanto a perfeição permanece algo distante, avanços têm sido feitos em mecanismos de correção de erros. A primeira etapa é permitir que os usuários reconheçam os erros, tanto no ASR quanto na TA.

Erros de ASR podem ser reconhecidos por usuários alfabetizados se as transcrições forem exibidas na tela do dispositivo. Para usuários analfabetos ou cegos, a reprodução sonora da transcrição pode ser usada. Para corrigir os erros do ASR, alguns sistemas podem permitir que os usuários editem ou escrevam as palavras. Em vez disso, podem ser fornecidos recursos para correção por voz. Outra opção é pronunciar o trecho de fala outra vez; mas então os erros podem se repetir ou novos erros podem surgir. Finalmente, as resoluções multimodais podem ser aplicadas, por exemplo envolvendo a seleção manual de um erro em uma interface gráfica seguida de correção por meio de voz.

Depois que quaisquer erros de ASR forem corrigidos, a transcrição pode ser passada para a TA (em sistemas cujos componentes ASR e TA são separados). Então, o reconhecimento e a correção da interpretação tornam-se mais fáceis.

Vários sistemas de IA auxiliam no reconhecimento de erros de TA, fornecendo traduções hipotéticas. Outros sistemas fornecem retrotraduções, para que os usuários possam determinar se o trecho que ele pronunciou continua compreensível após o processo de tradução. A retrotradução, no entanto, pode gerar traduções errôneas. Alguns sistemas minimizam erros gerando retrotraduções diretamente de representações semânticas neutras; e outros sistemas melhoram a precisão forçando o mecanismo de TA a reutilizar durante a retrotradução os elementos semânticos usados na tradução direta.

3.2.1 Tradutores Multimodais

A comunicação entre línguas pode envolver uma ampla gama de modalidades além de texto e fala. Os sistemas podem traduzir não apenas a fala, mas também

mensagens de texto, postagens, imagens de sinais de trânsito e documentos (YANG *et al.*, 2001; WAIBEL, 2002; ZHANG *et al.*, 2002) e até mesmo a leitura labial por meio do movimento dos músculos da face, medido por sensores eletromiográficos (MAIER-HEIN *et al.*, 2005). Ao que tudo indica, a multimodalidade é necessária para melhor capturar e transmitir elementos humanos da comunicação: emoções, gestos e expressões faciais ajudarão a transmitir a intenção dos falantes no contexto da cultura, relacionamentos, ambiente e status social.

A multimodalidade também varia de acordo com a situação. Em palestras, por exemplo, a reprodução de voz de várias fontes pode causar desconforto nos ouvintes, então pode-se empregar fones de ouvido, alto-falantes de áudio direcionado etc. Pode-se, também, transformar o intérprete automático em um tradutor que reconhece a os trechos de fala e os transforma em texto, apresentando a transcrição da fala na forma de legendas em um telão, em dispositivos pessoais, em óculos etc.

3.3 Expectativas e percepções dos usuários

Geralmente, as expectativas são definidas como representações mentais elaboradas antes de um indivíduo ter contato com algum serviço ou experiência e as percepções, representações mentais elaboradas depois do contato (JOHNSTON; CLARCK, 2002). De acordo com o dicionário Michaelis, a palavra “expectativa” possui as seguintes acepções: i. situação de quem espera um acontecimento em tempo anunciado ou conhecido; ii. esperança baseada em supostos direitos, probabilidades ou promessas; e iii. estado de quem espera um bem que se deseja e cuja realização se julga provável (MICHAELIS, 2015). Em relação à palavra “percepção”, dentre as acepções oferecidas pelo mesmo dicionário constam: i capacidade de distinguir por meio dos sentidos ou do intelecto; ii. representação mental das coisas; e iii. qualquer sensação física manifestada através da experiência (MICHAELIS, 2015). No âmbito dos Estudos da Interpretação, Pöchhacker (2016, p.430) afirma que as expectativas são “crenças sobre o que é provável que alguém obtenha ou gostaria de obter” a partir da interpretação. E, nesse sentido, de acordo com Pöchhacker, as expectativas são cruciais para avaliar a qualidade da prestação de serviços e dependem da perspectiva de um determinado usuário.

Com base nessas definições, parece razoável afirmar que expectativas e percepções estão intimamente relacionadas: as expectativas surgem das experiências prévias, ou seja, das percepções com outros serviços; as percepções, por sua vez, deixam

expostas as expectativas acerca de um serviço. Um usuário que compartilha, por exemplo, a percepção “achei o programa muito lento”, automaticamente, compartilha a expectativa de que o programa tivesse, de algum modo, um desempenho superior. Do mesmo modo, um usuário que diz esperar que um “serviço y” seja mais rápido que um “serviço x”, expõe sua percepção sobre este último. Assim, as expectativas e as percepções são fenômenos que se retroalimentam, principalmente quando as experiências com vários serviços são consideradas, conforme ilustra o esquema da Figura 2.

Figura 2 - Esquema das experiências e percepções



Fonte: elaborado pelo autor.

Estudos sobre as expectativas e percepções dos usuários são frequentemente conduzidos por áreas como o marketing (TOMASSINI; AQUINO; CARVALHO, 2008), engenharia de produção (SOUZA; GRIEBELER; GODOY, 2007), administração (MASANO, 2006) e diversas outras áreas que se enveredam a investigar a qualidade de determinados serviços. De acordo com Grönroos (1993) e Zeithaml e colaboradores (1990), a relação entre as expectativas e percepções dão origem à qualidade percebida pelos usuários, relação essa que foi originalmente pesquisada por Oliver (1980) em estudos a respeito da satisfação do usuário. Segundo Garcez e Rados (2002), os usuários costumam ter expectativas que podem ser consideradas bastante básicas, por isso, “é de grande importância que exista um desempenho básico e que as promessas sejam cumpridas. Os usuários esperam um bem ou serviço de qualidade em um ambiente amigável e cortês” (p.15).

3.3.1 Pesquisas sobre expectativas e percepções dos usuários no âmbito da interpretação

Na qualidade de prestadores de serviços, os intérpretes procuram atender às expectativas dos indivíduos envolvidos na comunicação ou às expectativas da pessoa ou instituição que paga pelo serviço. Geralmente, os indivíduos envolvidos na comunicação são chamados de “usuários” e a pessoa ou instituição que paga pelo serviço é chamada de “cliente”. Segundo Pöchhacker (2016), as expectativas, entendidas como “crenças sobre o que é provável que alguém obtenha ou gostaria de obter” (p. 430), são cruciais para avaliar a qualidade da prestação de serviços e dependem das perspectivas dos usuários em uma sessão de interpretação.

Iniciadas no final dos anos 1980, no âmbito da interpretação de conferências, as pesquisas sobre as expectativas dos usuários também têm despertado interesse de estudiosos da interpretação comunitária, na qual a dinâmica de interação comunicativa envolve necessidades e perspectivas divergentes tanto a respeito da situação comunicativa quanto a respeito do intérprete. No contexto da interpretação comunitária, no entanto, os intérpretes e os pesquisadores falam em “clientes” em vez de “usuários”, e suas expectativas se relacionam principalmente às qualificações, à função e ao comportamento dos intérpretes, e não à sua produção linguística (HALE, 2007).

Para descobrir as expectativas dos usuários, intérpretes e pesquisadores têm abordado os usuários por meio de entrevistas e questionários. Inicialmente empregado no contexto da interpretação de conferências, esse método foi lançado indiretamente por Bühler (1986), que solicitou a 47 membros da Associação Internacional de Intérpretes de Conferência (AIIC) que avaliassem a importância de 16 critérios para avaliar a qualidade dos intérpretes e a qualidade da interpretação. O estudo chegou à conclusão de que os critérios apontados pelos intérpretes de conferência refletiam as exigências dos usuários.

Em seu estudo, Kurz (1993) testou a validade dos critérios de Bühler (1986) em participantes de três conferências técnicas em que ocorreram sessões de interpretação simultânea, pedindo-lhes que avaliassem, em uma escala de quatro pontos, a importância de vários critérios de qualidade. Kurz descobriu que os participantes da conferência eram, em geral, menos exigentes do que os intérpretes profissionais pesquisados por Bühler (1986). Tanto intérpretes quanto usuários, no entanto, atribuíam maior importância a critérios como “sentir consistência com a mensagem original”¹⁰ e “coesão lógica do enunciado”¹¹, seguida de “uso da terminologia correta” e “fluência”. Os critérios “voz

¹⁰ Sense consistency with original message (Tradução minha).

¹¹ Logical cohesion of utterance (Tradução minha).

agradável” e “sotaque nativo” receberam as mais baixas classificações de importância tanto por parte dos usuários quanto por parte dos intérpretes.

As pesquisas de Kurz (1993) levaram vários outros intérpretes a conduzir estudos semelhantes sobre as expectativas dos usuários em vários países e domínios, adotando frequentemente os oito critérios relacionados à produção usados por Kurz (1993) e Bühler (1986) e concentrando-se na perspectiva do ouvinte no modo simultâneo. Collados-Aís e colaboradores (2007) replicaram o método em várias etapas com quase 200 professores de direito na Espanha, com resultados amplamente consistentes. Kurz (1993), no entanto, encontrou algumas divergências nos padrões de expectativa de diferentes grupos de usuários (médicos, engenheiros e delegados em uma conferência sobre Educação organizada pelo Conselho da Europa), sugerindo que as preferências relacionadas ao domínio podem influenciar a avaliação dos critérios de qualidade. Isso ficou mais evidente no contexto da interpretação simultânea em ambientes de mídia, onde âncoras de TV e chefes de programas de língua alemã deram prioridade a critérios como “fluência” e “voz agradável” e não à “completude”, critério priorizado pela audiência (KURZ; PÖCHHACKER, 1995).

A variabilidade das preferências do usuário foi confirmada em uma escala mais ampla em um estudo encomendado pela AIIC (MOSER, 1996). Liderados por um cientista social, 94 intérpretes de conferências realizaram 201 entrevistas com participantes em 84 eventos diferentes em vários países. Verificou-se que as expectativas dos usuários em relação a alguns aspectos da qualidade do desempenho variam de acordo com o tamanho e o tipo da reunião (grande/pequena, geral/técnica) e em relação a fatores como idade, gênero e experiência prévia em conferências. Desse modo, desconsiderando a preferência quase que unânime por interpretações fiéis ao significado, parece não haver um padrão mais ou menos estável de expectativas dos usuários. Além disso, as expectativas hipotéticas dos usuários em relação a um determinado critério de qualidade podem não corresponder ao impacto desse critério nos julgamentos dos usuários sobre um desempenho real de interpretação, como ocorre, por exemplo, em relação aos critérios “entonação monótona” (COLLADOS-AÍIS, 1998) e “sotaque não nativo” (CHEUNG, 2013).

Segundo Pöchhacker (2016), pesquisas sobre expectativas dos usuários sofrem de limitações metodológicas, pois critérios como “gramática correta” e “voz agradável” podem significar coisas diferentes para diferentes entrevistados. Essa incerteza conceitual

já havia sido apontada por Bühler (1986) e foi minuciosamente investigada nas entrevistas de Collados-Aís e colaboradores (2007).

No geral, as pesquisas sobre as expectativas dos usuários de interpretação de conferências demonstram que há uma ordem estável de preferências, em que aspectos de conteúdo como “sensação de consistência” e “coesão lógica” são complementados por “fluência” e “terminologia correta” e todos esses se sobressaindo aos critérios relacionados à forma, como “entonação vigorosa”, “voz agradável” e “sotaque nativo”. Levando em consideração as limitações metodológicas e a considerável variabilidade, no entanto, o conjunto de dados levantados acerca das expectativas dos usuários fornece pouca orientação específica para a tomada de decisões dos intérpretes de conferência em uma determinada situação. Isso também se aplica ao contexto da interpretação comunitária, onde as expectativas, principalmente as dos prestadores de serviços institucionais, podem variar conforme o domínio, o país e os padrões de qualidade estabelecidos pelas empresas que contratam os serviços de interpretação.

Além disso, Pöchhacker (2016) afirma que os documentos projetados para registrar o que os clientes podem esperar, na forma de códigos de ética ou conduta profissional, permanecem vagos, muitas vezes confiando em critérios como “exatidão”¹² e “comportamento imparcial”. Por outro lado, intérpretes comunitários que desenvolvem uma relação de trabalho próxima com um provedor de serviços ou clientes específicos (uma pessoa surda, por exemplo) podem desenvolver expectativas muito mais específicas.

A investigação das expectativas dos usuários de serviços de interpretação é importante para o aprimoramento do serviço, o desenvolvimento de códigos de ética pautados na percepção dos usuários e a elaboração de protocolos de controle de qualidade. No âmbito da IA, não tenho conhecimento de nenhum estudo, até o presente momento, tanto em língua portuguesa quanto em língua inglesa, que tenha investigado as expectativas e percepções dos usuários.

O presente estudo, portanto, tem por objetivo lançar luz sobre essa questão, buscando nas percepções compartilhadas publicamente pelos usuários as suas expectativas a respeito desse tipo de serviço. No presente estudo, no entanto, devido às limitações impostas pelo anonimato dos comentários compartilhados publicamente pelos

¹² Accuracy (Tradução minha).

usuários de aplicativos de IA, será possível obter apenas uma visão generalista sobre as expectativas e percepções dos usuários.

CAPÍTULO 4 – ASPECTOS METODOLÓGICOS: DA DESCRIÇÃO DOS APLICATIVOS À ANÁLISE DAS EXPECTATIVAS E PERCEPÇÕES DOS USUÁRIOS

Como foi dito anteriormente, esta pesquisa tem por objetivo geral identificar as percepções dos usuários brasileiros de serviços de IA oferecidos por aplicativos móveis da App Store e, a partir dessa identificação, tratar brevemente das expectativas desses usuários. Trata-se de um estudo exploratório, pois tem por objetivo definir e classificar fatos variáveis (SALOMON, 1978). A amostragem utilizada foi a acidental ou por acaso (não probabilística), porque conta com a extração de elementos de um dado universo de forma totalmente aleatória e não especificada.

O estudo foi desenvolvido em duas etapas: i. descrição dos aplicativos e ii. identificação das percepções e breve análise das expectativas dos usuários brasileiros. Trata-se, portanto, de uma pesquisa de caráter descritivo e experimental (SALDANHA; O'BRIEN, 2014), com base em dados tanto quantitativos quanto qualitativos, obtidos nas páginas eletrônicas dos aplicativos, hospedadas nos *hiperlinks* da App Store.

4.1 Os aplicativos selecionados para o estudo

Para realizar o estudo, selecionei aplicativos de IA disponíveis na App Store. Lançada em 2008, a App Store é uma das maiores e mais acessadas lojas de aplicativos móveis da atualidade. Em janeiro de 2020, o número total de aplicativos disponíveis era de 1,8 milhões e o número de instalações em junho de 2017 chegou a 180 bilhões. Embora possua uma política de compartilhamento de dados mais rígida que outras empresas de tecnologia, a Apple disponibiliza uma série de dados sobre os aplicativos vendidos na App Store, entre eles as classificações dos usuários, por meio das quais é feito um ranqueamento dos aplicativos; os comentários dos usuários, que podem servir de apoio para futuros usuários/consumidores que desejam saber quais as opiniões de quem já utilizou os aplicativos; e informações importantes sobre as atualizações dos aplicativos. Todas essas informações são disponibilizadas publicamente nas *webpages* dos aplicativos e podem ser acessadas tanto por meio de computadores de mesa quanto por meio dos dispositivos móveis da Apple.

Para a seleção dos aplicativos, acessei a App Store por meio de um iPhone SE, cujas especificações técnicas estão resumidas na Tabela 1. No mecanismo de busca da App Store, foram aplicadas as palavras-chave “tradutor”, “tradutor de fala”, “tradutor de voz”, “intérprete” e “interpretação automática”. Buscas com as duas últimas palavras-chave, no entanto, não incluíram aplicativos de meu interesse. Após as buscas iniciais, verifiquei que a palavra-chave “tradutor de voz” era a que mais recuperava aplicativos de interpretação automática: 52, no total. Diante do grande número de aplicativos encontrados e a impossibilidade de investigar cada um deles, parti para a análise dos dados iniciais que tinha a minha disposição.

Tabela 1 - Especificações técnicas do iPhone

Memória interna	128 Gb
Memória RAM	2 Gb
Modelo	SE
Versão do sistema operacional	13.3.1

Fonte: Elaborado pelo autor.

Assim que os aplicativos são recuperados, a App Store oferece uma visualização contendo informações resumidas sobre os serviços oferecidos, a classificação atribuída pelos usuários e o número de usuários que classificaram os aplicativos. Concentrando-me nas classificações, elaborei uma planilha contendo as principais informações de cada um dos aplicativos. Logo ficou claro que muitos aplicativos possuíam boas classificações, mas um baixo número de classificações. Um aplicativo, por exemplo, recebeu cinco estrelas, mas havia sido classificado por apenas uma pessoa e não havia recebido nenhum comentário, enquanto outro aplicativo, com 4.5 estrelas, havia sido classificado por mais de nove mil pessoas.

Para delimitar o *corpus* de estudo, defini que seriam selecionados apenas os aplicativos que tivessem recebido mais de mil classificações. Dos 52 aplicativos, oito se encaixam no parâmetro estabelecido, conforme mostra a Tabela 2, organizada de forma decrescente, do aplicativo com a melhor classificação para o aplicativo com a pior classificação. O Tradutor da Microsoft e o SayHi são os aplicativos com a melhor classificação e o Fala & Tradução é o aplicativo com o maior número de classificações recebidas.

Tabela 2 - Aplicativos com mais de mil classificações

Aplicativo	Classificação (estrelas)	Número de classificações
Tradutor da Microsoft ¹³	4.8	8.600
SayHi ¹⁴	4.8	3.600
iTranslate ¹⁵	4.7	24.300
Fala & Tradução ¹⁶	4.6	38.622
Traduzir Agora	4.7	3.400
Translate Me ¹⁷	4.5	9.100
Google Tradutor ¹⁸	4.5	6.800
Tradutor de Voz ¹⁹	4.4	6.900

Fonte: Elaborado pelo autor.

As classificações e o número de classificações, no entanto, oferecem apenas dados estatísticos. Além disso, o sistema de classificação por meio da atribuição de cinco estrelas é problemático, pois, na prática, os usuários tendem a atribuir notas extremas a experiências que envolvem muito mais fatores do que uma estrela e cinco estrelas podem sugerir (BREINLINGER; HAGIN; WRIGHT, 2019). Portanto, levando em consideração o teor quali-quantitativo da pesquisa, julguei necessário acrescentar às classificações e aos números de classificações um parâmetro que pudesse oferecer dados qualitativos sobre os aplicativos e, a partir daí, procedi à seleção dos aplicativos.

Para entender o mecanismo por trás da classificação adotada pela Apple e saber quais parâmetros ofereciam os dados de que necessitava, realizei leituras sistemáticas das *webpages* da empresa e fóruns sobre classificação de produtos. Assim como em outras plataformas móveis, o convite à classificação de um produto surge automaticamente algum tempo depois de os usuários buscarem os aplicativos, instalá-los e utilizá-los. A classificação dos aplicativos da App Store é feita de duas formas: i. atribuição de uma nota que varia de uma a cinco estrelas e ii. elaboração de comentários que expressem suas opiniões, nessa ordem. Ao usuário são dadas três opções: i. recusar-se a classificar; ii. classificar; ou iii. classificar e elaborar um comentário (APPLE, 2020).

¹³ Disponível em: <<https://apps.apple.com/br/app/tradutor-da-microsoft/id1018949559>>. Acesso em: 28 jul. 2020.

¹⁴ Disponível em: <<https://itunes.apple.com/br/app/wwdc/id437818260?mt=8>>. Acesso em: 28 jul. 2020.

¹⁵ Disponível em: <<https://apps.apple.com/br/app/id288113403?ign-mpt=uo%3D4>>. Acesso em: 28 jul. 2020.

¹⁶ Disponível em: <<https://apps.apple.com/br/app/id804641004>>. Acesso em: 28 jul. 2020.

¹⁷ Disponível em: <<https://apps.apple.com/br/app/tradutor-de-voz-translate-me/id1231853258>>. Acesso em: 28 jul. 2020.

¹⁸ Disponível em: <<https://apps.apple.com/br/app/google-tradutor/id414706506>>. Acesso em: 28 jul. 2020.

¹⁹ Disponível em: <<https://apps.apple.com/br/app/tradutor-de-voz/id1036725690>>. Acesso em: 28 jul. 2020.

Ao contrário das classificações superficiais obtidas por meio da atribuição de estrelas, os comentários oferecem dados qualitativos sobre a experiência dos usuários. Além disso, os comentários demandam tempo e disposição, sugerindo um engajamento maior por parte do usuário (GOEL, 2018). Partindo desses pressupostos, verifiquei que três aplicativos possuíam poucos comentários em relação ao número de classificações recebidas, conforme ilustra a Tabela 3, cujos itens foram ordenados de forma decrescente de acordo com o número de classificações de cada aplicativo.

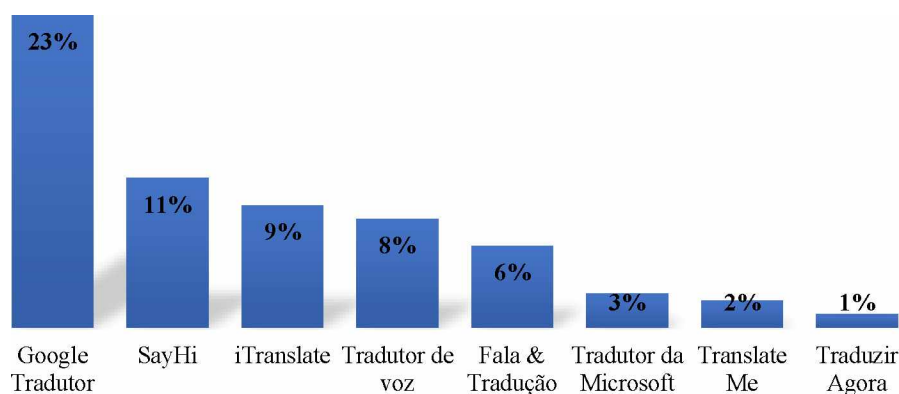
Tabela 3 - Número de classificações e número de comentários dos aplicativos

Aplicativo	Número de classificações	Número comentários
Fala & Tradução	38.622	2.230
iTranslate	24.300	2.207
Translate Me	9.100	149
Tradutor da Microsoft	8.600	221
Tradutor de Voz	6.900	572
Google Tradutor	6.800	1.567
SayHi	3.600	410
Traduzir Agora	3.400	34

Fonte: Elaborado pelo autor.

A razão entre o número de comentários e o número de classificações são as variáveis por meio dos quais é possível determinar o grau de engajamento com o produto (VAN DOORN *et al.*, 2010). Se, por exemplo, um aplicativo recebeu um número hipotético de 10 comentários e 20 classificações, o grau de engajamento do usuário para este aplicativo é de 0,5, ou seja, um comentário para cada duas classificações recebidas. Na Figura 3, o engajamento dos usuários dos aplicativos de IA da App Store é expresso em termos percentuais, do aplicativo cujo engajamento dos usuários é maior ao aplicativo com o menor engajamento.

Figura 3 - Engajamento dos usuários



Fonte: Elaborado pelo autor.

O engajamento dos usuários dos aplicativos Tradutor da Microsoft, Translate Me e Traduzir Agora não chegam nem a 5%, indicando que o número de comentários é baixo em relação ao número de classificações. O aplicativo Traduzir Agora, cujo engajamento dos usuários é o menor, possui somente 34 comentários enquanto o Google Tradutor, cujo engajamento dos usuários é o maior, possui mais de mil comentários.

Segundo Goel (2018), para “cada usuário que se empenha em reclamar, cerca de 26 outros permanecem em silêncio” (não paginado)²⁰. Os comentários, no entanto, não necessariamente contêm apenas reclamações, mas o fato de a maioria dos usuários permanecer em silêncio mesmo quando lhes é concedido o direito de exigir a melhoria do produto adquirido sugere a importância do número de comentários relativos a determinado serviço. Em função disso, decidi não incluir nas análises e descrições do presente estudo os aplicativos Tradutor da Microsoft, Translate Me e Traduzir Agora, cujos engajamentos dos usuários mostraram-se muito inferior aos demais. Estes, portanto, foram os critérios utilizados na seleção dos cinco aplicativos com os quais trabalho no presente estudo e cujas informações estão resumidas na Tabela 4.

²⁰ For every customer who bothers to complain, nearly 26 others remain silent.

Tabela 4 - Informações sobre os aplicativos selecionados

Aplicativos	Google Tradutor / SayHi / iTranslate / Tradutor de Voz / Fala & Tradução
Número total de comentários	6.577
Anos dos comentários	janeiro/2012 a março/2020
Idioma	português brasileiro

Fonte: Elaborado pelo autor.

4.2 A descrição dos aplicativos

Como foi dito, a qualidade é um conceito relativo e multidimensional (PÖCHHACKER, 2016; GILE; ESQUEDA; FREITAS, 2019), por isso, julgo que o primeiro passo antes de se estabelecer a qualidade de um serviço é saber quais são as suas principais características. Os estudos de Matsuda e colaboradores (2013) e Zheng e colaboradores (2010) são alguns dos poucos estudos que descrevem aplicativos de IA desenvolvidos para plataformas móveis. Além disso, pouco tem sido publicado sobre o modo com que esses aplicativos são explorados pela indústria de tecnologia para plataformas móveis.

A tarefa de descrição dos aplicativos mostrou-se desafiadora desde o início, pois a interatividade que os aplicativos oferecem faz com que a experiência de utilização seja bastante diversificada: diante de tantas funcionalidades, quais caminhos tomar? Dunleavy (2003) problematiza diferentes tipos de abordagens descritivas e afirma que descrições que tomam por base a organização espacial e a ordem cronológica de acontecimentos, ou seja, elementos externos ao pesquisador, tendem a ser difíceis de serem mentalmente mapeadas e processadas pela memória. A recomendação do autor é que os elementos sejam apresentados de forma analítica, ou seja, por meio de categorias elaboradas pelo pesquisador com base em sua experiência com o objeto de estudo.

Nesse sentido, partilho da ideia de que a experiência do pesquisador não deve ser colocada à margem da pesquisa (MOCH; GATES, 2000). Por isso, a descrição analítica aqui proposta baseia-se no uso experimental que tenho feito da tecnologia de interpretação automática desde a graduação, quando investiguei o estado da arte dos sistemas de IA (FREITAS, 2016).

A descrição contemplará informações gerais sobre os aplicativos, aspectos relacionados à configuração e os principais recursos oferecidos aos usuários. A descrição

conterá ilustrações, capturas de telas e tabelas. Por motivos de espaço, alguns elementos são ilustrados com as capturas de telas de apenas um aplicativo, com exceção dos casos em que o objetivo é evidenciar o modo específico com que cada aplicativo disponibiliza esses elementos em suas respectivas interfaces.

4.3 Análise das expectativas dos usuários

Na qualidade de prestadores de serviços, os intérpretes humanos procuram atender às expectativas dos indivíduos envolvidos na comunicação ou às expectativas das pessoas ou instituições que pagam pelo serviço. Na IA, os indivíduos que participam da comunicação são também as pessoas que pagam pelo serviço. Nem todos os aplicativos, no entanto, são pagos. Por isso, no presente estudo, não faço distinção entre os indivíduos que pagam e os que não pagam, referindo-me a todos como “usuários”.

De modo geral, a qualidade da IA tem sido avaliada por meio de métodos avaliativos automáticos ou manuais e raros estudos contam com a presença de seres humanos. Os poucos que envolvem seres humanos se dão em ambientes controlados em que a qualidade é avaliada tendo por base dados obtidos a partir de questionários aplicados a usuários hipotéticos.

Como já foi mencionado, na interpretação humana, para averiguar quais são as expectativas dos usuários, os intérpretes e pesquisadores abordam os usuários e clientes reais por meio de entrevistas e questionários (GRBIC, 2015). Abordar os usuários da IA em aplicativos móveis, no entanto, é uma tarefa complexa e muitas vezes impossível, pois geralmente os usuários não são identificados ou não se deixam identificar. Além disso, as ocasiões em que essa tecnologia é utilizada são bastante intimistas e para as quais não há divulgação prévia e nem há a presença de um mediador que possa facilitar o acesso, diferentemente do que ocorre na Interpretação de Conferências, por exemplo.

No caso da IA oferecida por aplicativos para dispositivos móveis, os usuários expõem suas percepções sobre os aplicativos em fóruns ou *webpages* específicas por meio de comentários espontâneos, vlogs ou resenhas. De acordo com Chesterman (2007), os comentários e demais *feedbacks* públicos compartilhados pelos usuários fazem parte de um processo cognitivo que se inicia quando os usuários entram em contato com uma tradução. É uma resposta que evidencia se suas expectativas foram atendidas ou não. Transportando essa noção para a Interpretação, Kurz (2002) afirma que a pesquisa de

expectativas e preferências dos usuários é crucial para uma profissão cujo objetivo é estabelecer uma comunicação eficaz entre as partes envolvidas.

Isto posto, a análise das expectativas dos usuários veiculadas em comentários fidedignos compartilhados espontaneamente pelos usuários dos cinco aplicativos de IA selecionados mostra-se pertinente para o estudo que proponho.

4.3.1 Materiais e métodos de análise

Aos cinco aplicativos foram atribuídos mais de seis mil comentários. Ante essa grande quantidade de comentários e a impossibilidade de analisar qualitativamente seus respectivos conteúdos, devido às restrições de tempo impostas pelo mestrado, utilizei a ferramenta *online* Random Integer Generator²¹ para a seleção aleatória dos comentários a serem analisados. Para manter uma amostragem proporcional, selecionei aleatoriamente 40 comentários de cada aplicativo, totalizando 200 comentários fidedignos. Para a seleção aleatória dos comentários, no entanto, foi necessário coletar todos os comentários nas *webpages* dos cinco aplicativos e os colocar em planilhas a fim de que cada comentário recebesse uma numeração que possibilitasse a sua posterior identificação.

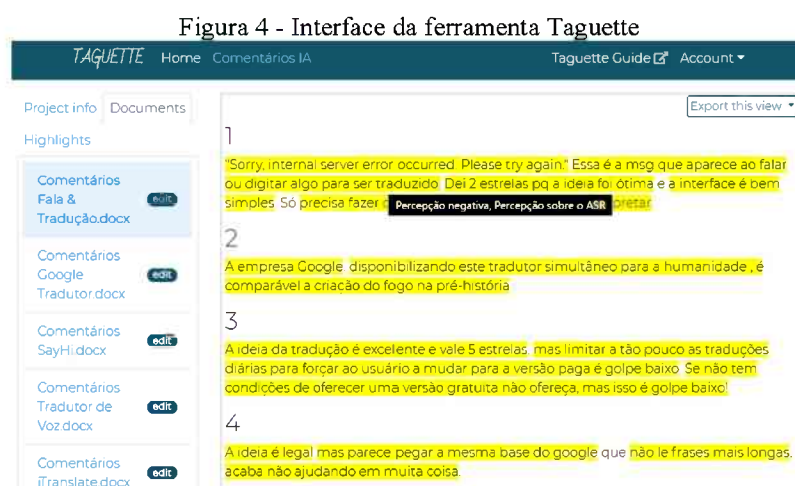
Após a seleção aleatória, os comentários foram reunidos em arquivos txt e submetidos à ferramenta gratuita Taguette²², por meio da qual os comentários foram categorizados conforme as percepções veiculadas em cada trecho dos comentários. A ferramenta Taguette suporta vários formatos de arquivos, mas em testes anteriores com arquivos em formato .pdf e .docx, ocorreram interferências de metadados e desconfiguração dos comentários, por isso, optei por reunir os comentários selecionados em arquivos txt, caracterizados pela ausência de formatação de estilo e metadados.

Desenvolvida por Rémi Rampin, engenheiro da computação da Universidade de Nova York, a Taguette é uma ferramenta que pode tanto ser instalada no computador pessoal do pesquisador quanto utilizada *online*. Voltada para a análise de dados qualitativos por meio da criação de etiquetas e realces em trechos de texto que o pesquisador julga relevantes para a sua pesquisa, a ferramenta mostrou-se interessante para a minha pesquisa, pois além de gratuita, é fácil de usar, suporta vários formatos de arquivo e possibilita a exportação dos dados etiquetados.

²¹ Disponível em: <<https://www.random.org/integers/>>. Acesso em: 13 jul. 2021.

²² Disponível em: <<https://www.taguette.org>>. Acesso em: 15 ago. 2021.

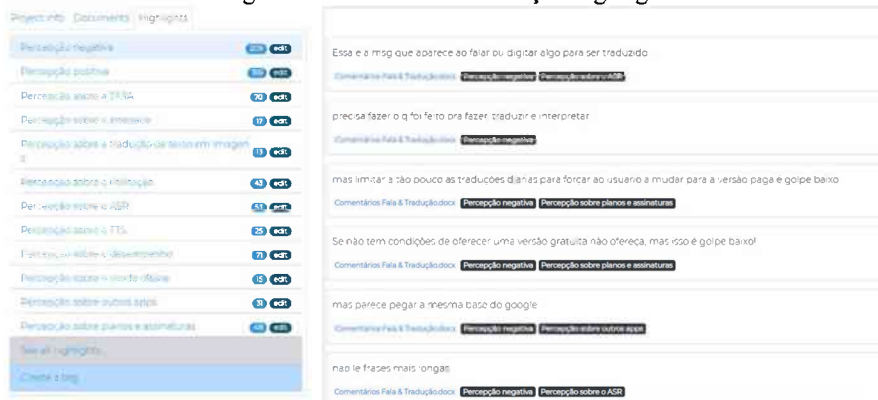
A interface da ferramenta Taguette é composta por quatro painéis principais e oferece dois modos de visualização. No modo *Documents*, ilustrado na Figura 4, o pesquisador cria etiquetas e as atribui aos trechos de texto com os quais está trabalhando. Uma vez atribuídas aos trechos, as etiquetas assumem a forma de *highlights*, ou seja, realces que marcam toda a extensão do trecho definido pelo pesquisador. Várias etiquetas podem ser atribuídas ao mesmo trecho, mas no modo de visualização *Documents*, os rótulos das etiquetas só podem ser visualizados se o usuário passar o cursor sobre o realce, como é mostrado pela tarja preta da Figura 4, em cujo rótulo lê-se “percepção negativa, percepção sobre o ASR”.



Fonte: Interface *online* Taguette.

No modo de visualização *highlights*, ilustrado na Figura 5, os trechos são agrupados conforme as etiquetas atribuídas. Esse modo de visualização possibilita ao pesquisador uma visualização mais detalhada dos trechos e das etiquetas que coocorrem em um mesmo trecho, caso isso aconteça. Além disso, esse modo de visualização permite que o pesquisador exporte os trechos realçados de acordo com as etiquetas, possibilitando a seleção dos trechos que julgar relevante para a sua pesquisa.

Figura 5 - Modo de visualização highlights



Fonte: Interface *online* Taguette.

A simplicidade da ferramenta Taguette exige pouco tempo para que o pesquisador aprenda a utilizar suas principais funcionalidades. A sua simplicidade, no entanto, faz com que a ferramenta careça de alguns recursos de personalização das etiquetas que podem facilitar a vida do pesquisador. Para a análise dos comentários dos usuários dos aplicativos de interpretação automática, por exemplo, a impossibilidade de criar subetiquetas demanda a criação de etiquetas cujo nível de hierarquia concorre com as etiquetas principais. Essa limitação, no entanto, não afeta a análise dos dados se o pesquisador realizar o gerenciamento hierárquico das etiquetas de forma manual. Além disso, o fato de a ferramenta ser gratuita faz com que a limitação na personalização das etiquetas seja relativizada frente a existência de ferramentas semelhantes cuja utilização requer a compra de planos de utilização que podem ser onerosos dependendo do orçamento da pesquisa.

CAPÍTULO 5 – OS APLICATIVOS E SEUS RECURSOS

Neste capítulo descrevo cinco aplicativos de IA disponíveis na App Store: Google Tradutor, SayHi, iTranslate, Tradutor de Voz e Fala & Tradução. Os critérios utilizados para a seleção dos aplicativos foram apresentados no **Capítulo 4**. A descrição apresenta as principais características e funcionalidades que os aplicativos oferecem aos usuários.

Partindo do pressuposto de que a experiência do pesquisador não deve ser colocada à margem da pesquisa (MOCH; GATES, 2000), proponho uma descrição baseada no uso experimental dessa tecnologia que tenho feito desde a graduação, quando investiguei o estado da arte dos sistemas de IA e realizei experimentos com o Google Tradutor (FREITAS, 2016). Trata-se de uma descrição analítica (DUNLEAVY, 2003) que abrange informações obtidas desde o momento em que realizei a busca pelos aplicativos na App Store, passando pela instalação, configuração, até o momento em que os utilizei para realizar os experimentos e avaliações preliminares, cujos resultados são expostos adiante. Dei prioridade, no entanto, aos elementos relacionados às etapas de configuração e utilização, haja vista que a etapa de instalação se dá em apenas um clique e, por ser muito simples, não oferece dados que sejam relevantes à descrição.

5.1. Informações gerais

Assim que surgiu, o slogan da App Store era: “There’s an app for that”²³, que retratava o fato de a loja oferecer um aplicativo para praticamente qualquer situação. De fato, o portfólio de aplicativos disponíveis na App Store inclui uma ampla variedade de aplicativos, de jogos bastante simples e rudimentares a aplicativos de monitoramento de batimentos cardíacos e até mesmo cartomancia. As lojas de aplicativos agregam informações importantes criadas tanto pelos desenvolvedores de aplicativos quanto pelos usuários. Nas *webpages* dos aplicativos, os desenvolvedores descrevem os aplicativos por meio de blocos de textos contendo apresentações resumidas e listas dos recursos oferecidos, compartilham informações sobre atualizações e apresentam os valores e especificidades dos planos e assinaturas disponíveis.

²³ “Há um app para isso” (Tradução minha).

Os cinco aplicativos selecionados para o presente estudo foram recuperados na App Store por meio das palavras-chave “tradução de voz” e “tradução de fala”.

5.1.1 Datas de lançamento

Dos cinco aplicativos, apenas o Google Tradutor, lançado mundialmente em 2006, foi desenvolvido para a tradução de texto, pela qual tornou-se mais conhecido. A funcionalidade de interpretação automática do Google Tradutor foi adicionada em 2011, mesmo ano em que foi lançado na App Store, como mostra a Tabela 5.

Tabela 5 - Data de Lançamento dos aplicativos na App Store

Aplicativo	Ano de lançamento na App Store
Google Tradutor	2011
iTranslate	2011
SayHi	2012
Fala & Tradução	2014
Tradutor de Voz	2015

Fonte: Elaborado pelo autor.

O Google Tradutor e o iTranslate são os aplicativos disponíveis há mais tempo na App Store e o Tradutor de Voz, lançado quatro anos depois, em 2015, é o aplicativo com menos tempo de App Store. O aplicativo SayHi, que até 2018 era pago, foi lançado em 2012 e o Fala & Tradução, em 2014.

5.1.2 Disponibilidade

Atualmente, existem duas grandes plataformas móveis: Android e iOS, a primeira pertencente à Google e a segunda, à Apple. Instaladas em tablets, *smartphones* e *smartwatches*, juntas elas são responsáveis por cerca de 91% das vendas de *smartphones* e cerca de 96% das vendas de tablets em todo o mundo (RASMUSSEN, 2019). A Tabela 6 mostra a disponibilidade dos aplicativos em diferentes plataformas. A “web” é a mais ampla de todas, pois pode ser acessada tanto de *smartphones*, tablets, notebooks e computadores de mesa.

Tabela 6 - Disponibilidade dos aplicativos

Aplicativo	Plataformas	Produtos da Apple
Google Tradutor	iOS, Android, web e API	iPhone e iPad
iTranslate	iOS, Android, web	iPhone, iPad, Apple Watch e iMessage
SayHi	iOS, Android	iPhone e iPad
Fala & Tradução	iOS	iPhone, iPad e Apple Watch
Tradutor de Voz	iOS	iPhone

Fonte: Elaborado pelo autor.

Os itens da Tabela 6 estão ordenados de acordo com a disponibilidade dos aplicativos nas plataformas. O Google Tradutor está disponível em diversas plataformas e o iTranslate é o aplicativo mais presente nos produtos da Apple. Os aplicativos Fala & Tradução e Tradutor de Voz só estão disponíveis para usuários de produtos da Apple. O aplicativo SayHi pode ser instalado tanto em dispositivos iOS quanto em dispositivos Android. Relacionando as datas de lançamento dos aplicativos (Tabela 5) às suas respectivas disponibilidades (Tabela 6), percebe-se que os aplicativos lançados há mais tempo são os mais disponíveis.

5.1.3 Módulos de ASR, TA e TTS

Além da disponibilidade, outro aspecto importante na caracterização dos aplicativos são os módulos de ASR, TA e TTS. Apesar de o iOS possuir um módulo de ASR integrado, que é utilizado pelos aplicativos Tradutor de Voz, Google Tradutor e iTranslate, o ASR dos aplicativos SayHi e Fala & Tradução são oferecidos pela Nuance Communications.

Quanto ao módulo de TA, dois grandes servidores podem ser destacados: as APIs²⁴ Google Cloud²⁵ e Microsoft Azure²⁶. Na Tabela 7, é possível visualizar que a Google Cloud é responsável pelas traduções automáticas de quatro aplicativos, a

²⁴ Application Programming Interface, trata-se de um conjunto de rotinas e padrões de programação para acesso a um aplicativo de *software* ou plataforma baseado na web. As APIs são criadas quando uma empresa de *software* tem a intenção de que outros criadores de *software* desenvolvam produtos associados ao seu serviço.

²⁵ Disponível em: <<https://cloud.google.com/translate>>. Acesso em: 28 jul. 2020.

²⁶ Disponível em: <<https://azure.microsoft.com/pt-br/>>. Acesso em: 28 jul. 2020.

Microsoft Azure é responsável pelas traduções automáticas de dois aplicativos, sendo que no aplicativo Fala & Tradução, ambas APIs trabalham em conjunto.

Tabela 7 - Servidores de tradução automática

Aplicativo	Módulo ASR	Módulo TA
Fala & Tradução	Nuance	Google Cloud e Microsoft Azure
Tradutor de Voz	Apple ASR	Google Cloud
Google Tradutor	Apple ASR	Google Cloud
iTranslate	Apple ASR	Google Cloud
SayHi	Nuance	Microsoft Azure

Fonte: *Webpages* dos aplicativos.

Diante da pouca variedade de tecnologias envolvidas no ASR e tradução automática pode-se pensar que as interpretações automáticas realizadas pelos cinco aplicativos não se distanciem tanto em questões de qualidade. Nenhum dos aplicativos informam qual o módulo de TTS utilizado para a sintetização de voz.

5.1.4 Atualizações

Na App Store, é possível encontrar um resumo das atualizações que cada aplicativo implementou nos últimos anos. As atualizações são importantes, pois melhoram a experiência de uso e demonstram comprometimento com os usuários (FLEISCHMANN *et al.*, 2016). A Tabela 8 mostra o número de atualizações de cada aplicativo no ano de 2019.

Tabela 8 - Atualizações em 2019

Aplicativo	Número de atualizações em 2019
Google Tradutor	8
SayHi	8
Tradutor de Voz	12
Fala & Tradução	15
iTranslate	33

Fonte: *Webpages* dos aplicativos.

Os aplicativos com o menor número de atualizações são o Google Tradutor e o SayHi, ambos gratuitos, mas ambos com os maiores engajamentos por parte dos usuários. Os aplicativos Tradutor de Voz e Fala & Tradução foram atualizados mais de 10 vezes

no ano de 2019 e o iTranslate, o aplicativo mais presente nos produtos da Apple, como já mencionado, foi atualizado mais de 30 vezes no mesmo período.

5.1.5 Tamanho dos aplicativos

Atualmente, os usuários de dispositivos móveis podem escolher entre milhões de aplicativos. Em janeiro de 2020, por exemplo, o número total de aplicativos disponíveis na App Store era de 1,8 milhões (COSTELLO, 2020). No Brasil, *smartphones* e outros dispositivos móveis têm em média 63 aplicativos instalados, dos quais, 22 são usados diariamente (AVAZU, 2003). Em um cenário como esse, é de se esperar que o tamanho deles seja um fator importante a ser considerado antes de sua instalação ou compra.

Quanto maiores, mais espaço eles ocupam, impedindo assim que outros sejam instalados. E embora as capacidades de armazenamento tenham aumentado constantemente e o preço da tecnologia tenha se mantido estável, o tamanho deles gera impactos no consumo de energia dos dispositivos, fazendo com que os usuários tenham que recarregá-los com maior frequência (NGUYEN, 2013). A Tabela 9 mostra que, com exceção do Fala & Tradução, os tamanhos dos aplicativos estão relacionados à quantidade de serviços oferecidos.

Tabela 9 - Tamanhos dos aplicativos e serviços oferecidos

Aplicativo	Tamanho dos aplicativos	Serviços oferecidos
SayHi	26,3 MB	Interpretação automática Tradução de texto
Google Tradutor	94,4 MB	Interpretação automática Tradução de texto Tradução de texto em imagens
Fala & Tradução	139,2 MB	Interpretação automática Tradução de texto
Tradutor de Voz	194,2 MB	Interpretação automática Tradução de texto Tradução de texto em imagens Guia de conversação
iTranslate	586 MB	Interpretação automática Tradução de texto Tradução de texto em imagens Guia de Conversação Tradução de sites Teclado de tradução Conjugações verbais

Fonte: *Webpages* dos aplicativos.

O SayHi é o que ocupa menos espaço e oferece os serviços de IA e tradução de texto. O Fala & Tradução oferece os mesmos serviços, mas ocupa cerca de cinco vezes mais espaço. O iTranslate oferece 3,5 mais serviços e ocupa cerca de 22 vezes mais espaço que o SayHi. O Google Tradutor, que oferece os serviços de IA, tradução de texto e tradução de texto em imagens, exige que o usuário tenha pelo menos 94,4 MB de memória livre e o Tradutor de Voz, que oferece os mesmos serviços e um guia de conversação, exige pelo menos 194,2 MB de memória livre.

5.1.6 Passo a Passo de utilização e suporte

O primeiro contato do usuário com o aplicativo pode dificultar a utilização de recursos que não estejam disponibilizados na tela ou configurações avançadas. Por isso, é importante que os desenvolvedores disponibilizem instruções de utilização. Dos cinco aplicativos, no entanto, apenas o SayHi oferece um passo a passo de utilização integrado à sua interface. Os demais redirecionam os usuários a páginas de suporte contendo perguntas e respostas, explicações sobre como utilizar os recursos e tutoriais em formato de vídeo.

5.2 Configurando os aplicativos

A etapa de configuração dos aplicativos não demanda muito tempo. Além disso, a maioria dos aplicativos oferece ao usuário a opção de configurar mais tarde ou simplesmente deixar de configurar e avançar para outras etapas. Na etapa de configuração, os usuários são levados a fazer uma série de escolhas que afetam a experiência de utilização dos recursos oferecidos.

5.2.1 Autorizações

Após a instalação, os aplicativos Google Tradutor, Tradutor de Voz e iTranslate, que fazem uso do módulo de ASR da Apple, solicitam que os usuários os autorizem a acessar o módulo e informam que o reconhecimento de fala só será ativado quando o usuário tocar no botão de voz, conforme ilustrado pela captura de tela do aplicativo SayHi na Figura 6. Do mesmo modo, o Google Tradutor, o Tradutor de Voz e o iTranslate, que

oferecem o serviço de tradução de texto em imagens, solicitam autorização de acesso à câmera do dispositivo. Todos os aplicativos solicitam autorização para o envio de notificações.

Caso opte por não conceder acesso ao microfone e à câmera, o usuário não pode usufruir dos recursos que fazem uso desses módulos. Diferentemente das solicitações de acesso ao módulo de ASR e à câmera, o acesso à central de notificações é uma escolha que não afeta a usabilidade dos aplicativos, pois os recursos continuam funcionando normalmente.

Figura 6 - Acesso ao microfone



Fonte: Interface do aplicativo SayHi.

Ainda na etapa de autorização, os aplicativos apresentam informações sobre os termos de utilização, propriedade intelectual e política de privacidade. De modo geral, eles parecem utilizar um modelo padrão de documentação, com poucas alterações de um para outro. Aos usuários é vetado o uso dos aplicativos para fins comerciais, alteração do código fonte, engenharia reversa e depreciação de pessoas. Dos cinco, apenas o Google Tradutor disponibiliza sua documentação em língua portuguesa, os demais somente em inglês.

5.2.2 Planos e assinaturas

Nos aplicativos que cobram pelos serviços, assim que as autorizações são concedidas, os usuários são redirecionados a um painel de planos e assinaturas. No iTranslate, Tradutor de Voz e Fala & Tradução, os usuários podem utilizar todas as funcionalidades por um período de testes que varia de três a sete dias, após o qual o valor do plano selecionado passa a ser cobrado automaticamente. Caso a assinatura seja cancelada antes de o período de teste terminar, os usuários podem realizar traduções e interpretações até o fim do período de teste estipulado. Depois disso, o uso das funcionalidades é restringido a determinado número de frases ou a interface do aplicativo passa a mostrar *banners* de propagandas de outros aplicativos e serviços do Google AdSense²⁷.

Dos cinco aplicativos, três são pagos e dois são gratuitos. Os aplicativos que cobram pelos serviços oferecem planos anuais e mensais, somente o aplicativo Tradutor de Voz oferece um plano semanal. A Tabela 10 mostra os valores que cada um deles cobra e a relação de serviços oferecidos referentes ao primeiro semestre de 2020. Os aplicativos Google Tradutor e SayHi são gratuitos.

Tabela 10 - Valores e número de serviços oferecidos

Aplicativo	Planos e valores	Número de Serviços
Google Tradutor	Gratuito	3
SayHi	Gratuito	2
iTranslate	Plano mensal: R\$ 19,90 Plano anual: R\$ 159,90	7
Tradutor de Voz	Plano semanal: R\$ 25,90 Plano mensal: R\$ 40,90 Plano anual: R\$ 209,90	4
Fala & Tradução	Plano mensal: R\$ 22,90 Plano anual: R\$ 129,90	2

Fonte: Webpages dos aplicativos.

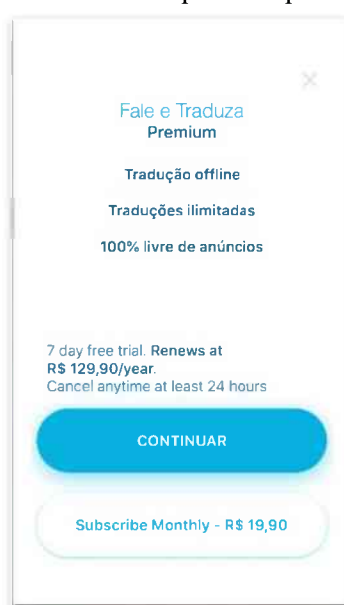
Dos aplicativos pagos, o iTranslate é o que oferece mais serviços e o seu plano mensal é o mais acessível. O Fala & Tradução tem o plano anual mais acessível, mas oferece apenas os serviços de IA e tradução de texto.

Durante a etapa de configuração, duas inconsistências ocorreram com os aplicativos Fala & Tradução e Tradutor de Voz. No painel de planos e assinaturas do aplicativo Fala & Tradução, ilustrado na Figura 7, as informações sobre os recursos da

²⁷ Plataforma de divulgação paga de produtos e serviços.

assinatura *premium* são apresentadas em português, já as informações sobre o período de teste, o valor a ser cobrado após esse período e as informações do botão de assinatura mensal são apresentadas em inglês. Conforme mencionado, esses aplicativos são desenvolvidos para atender às necessidades de pessoas monolíngues, então informações importantes relativas a valores de compra e condições de uso que não podem ser compreendidas podem ter um impacto negativo na experiência do usuário. Além disso, parece ser contraditório que para instalar um aplicativo de tradução, o usuário tenha que recorrer a outro aplicativo de tradução.

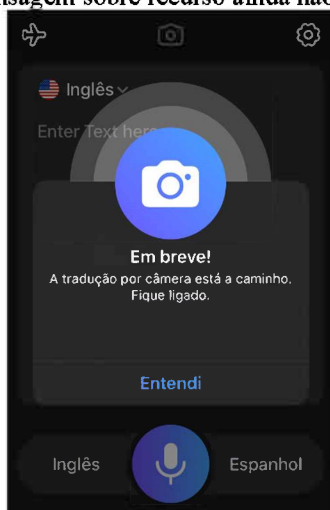
Figura 7 - Inconsistências no painel de planos e assinaturas



Fonte: Aplicativo Fala & Tradução.

A segunda inconsistência diz respeito ao fato de o aplicativo Tradutor de Voz, que em sua *webpage* diz oferecer tradução de texto em imagens, não ter implementado o recurso na interface do aplicativo, como mostra a Figura 8. Se os usuários pautarem suas decisões de compra nas informações fornecidas pelos aplicativos em suas *webpages*, a incompatibilidade entre o que é prometido e o que realmente é oferecido pode gerar frustrações (GORLA *et al.*, 2014). Essa inconsistência, no entanto, ocorreu somente com o aplicativo citado e parece não ser algo recorrente nos demais aplicativos.

Figura 8 - Mensagem sobre recurso ainda não implementado



Fonte: Aplicativo Tradutor de Voz.

5.2.3 Línguas e Pares linguísticos

Diferentemente dos intérpretes humanos, que interpretam para um número limitado de línguas, os intérpretes automáticos oferecem inúmeras combinações linguísticas, fazendo com eles possam ser considerados multi-intérpretes. Definir as línguas e os pares linguísticos é uma das etapas de configuração mais importantes. O número de línguas disponibilizadas e as combinações que podem ser feitas dependem de vários fatores: o serviço a ser utilizado (interpretação automática, tradução de texto, tradução de imagens etc.), os módulos de ASR e TA à disposição, a conexão de dados e fatores como os planos e assinaturas contratados pelos usuários. A Tabela 11 mostra o número de línguas disponíveis para cada funcionalidade dos cinco aplicativos e os pares linguísticos que podem ser utilizados. O número de pares linguísticos resulta da combinação de duas línguas disponíveis, não importando a direcionalidade.

Tabela 11 - Pares linguísticos

Aplicativo	Interpretação automática		Tradução de texto	
	Número de línguas	Pares linguísticos	Número de línguas	Pares linguísticos
Google Tradutor ²⁸	42	861	103	5.253
SayHi	33	528	54	1.431
Fala & Tradução ²⁹	32	496	108	5.778

²⁸ Disponível em: <<https://translate.google.com/intl/en-GB/about/languages/>>. Acesso em: 28 jul. 2020.

²⁹ Disponível em: <https://www.apalon.com/speak_and_translate.html>. Acesso em: 28 jul. 2020.

iTranslate ³⁰	29	406	91	4.095
Tradutor de Voz	26	325	94	4.371

Fonte: Webpages dos aplicativos.

O Google Tradutor é o aplicativo que oferece o maior número de combinações de pares linguísticos para a interpretação automática e o Fala & Tradução, o que oferece o maior número de combinações para a tradução de texto. O Google Tradutor oferece 10 línguas a mais para a interpretação automática do que o Fala & Tradução e, como foi mostrado, ambos utilizam o mesmo API de traduções, mas módulos de ASR diferentes. O que pode explicar o menor número de línguas disponíveis para a interpretação automática no aplicativo Fala & Tradução.

Além disso, o Google Tradutor, o iTranslate e o Tradutor de voz fazem uso dos mesmos módulos de ASR e TA, mas, ainda assim, oferecem quantidades diferentes de línguas para cada funcionalidade. O número de pares linguísticos revela a abrangência dos aplicativos, quanto maior o número de combinações, maior o número de indivíduos e comunidades abrangidas. O aplicativo Fala & Tradução oferece cinco línguas a mais para a tradução de texto que o Google Tradutor, cujo único módulo de TA é o Google Cloud; e 45 línguas a mais para a tradução de texto que o SayHi, cujo único módulo de TA é o Microsoft Azure, mostrando que a combinação de módulos de TA, pelo menos no que diz respeito à tradução de texto, torna o aplicativo mais abrangente.

A tradução de texto é um importante recurso oferecido pelos aplicativos. O foco desse estudo, no entanto, é a interpretação automática, por isso, as análises seguintes consideram apenas o serviço de interpretação automática oferecido pelos cinco aplicativos de IA disponíveis na App Store.

5.2.4 Abrangência da interpretação automática

Outro aspecto relevante para a abrangência dos aplicativos é a disponibilidade de interpretações automáticas para diferentes línguas. Mesmo que pertençam à mesma família, as línguas diferem entre si sob vários aspectos, por isso os módulos de ASR precisam ser treinados para trabalhar com diferentes regras fonológicas (BAHL *et al.*, 1980). A Tabela 12 mostra que o Google Tradutor é o aplicativo que abrange o maior

³⁰ Disponível em: <<https://www.itranslate.com/languages>>. Acesso em: 28 jul. 2020.

número de línguas, oferecendo interpretações automáticas para nove línguas de 61 países. O Fala & Tradução é o de menor abrangência, oferecendo interpretações automáticas para cinco línguas de 13 países.

Tabela 12 – Interpretação automática e variantes linguísticas

Aplicativo	Língua	Variantes linguísticas
Google Tradutor	árabe	Arábia Saudita, Argélia, Barém, Egito, Emirados Árabes Unidos, Israel, Jordânia, Kwait, Líbia, Marrocos, Omã, Palestina, Qatar, Tunísia
	bengali	Bangladesh, Índia
	espanhol	América Latina, Argentina, Bolívia, Chile, Colômbia, Costa Rica, El Salvador, Espanha, EUA, Guatemala, Honduras, México, Nicarágua, Panamá, Paraguai, Peru, Porto Rico, Rep. Dominicana, Uruguai, Venezuela
	francês	Canadá, França
	inglês	África do Sul, Austrália, Canadá, EUA, Filipinas, Gana, Índia, Nigéria, Nova Zelândia, Quênia, Reino Unido, Tanzânia
	mandarim	China, Hong Kong, Taiwan
	português	Brasil, Portugal
	suaíli	Quênia, Tanzânia
SayHi	tâmil	Índia, Malásia, Singapura, Sri Lanka
	árabe	Arábia Saudita, Argélia, Barém, Egito, Emirados Árabes Unidos, Iémen, Iraque, Jordânia, Kwait, Líbia, Marrocos, Omã, Palestina, Qatar, Tunísia)
	espanhol	Argentina, Bolívia, Chile, Colômbia, Costa Rica, El Salvador, Espanha, EUA, Guatemala, Honduras, México, Nicarágua, Panamá, Paraguai, Peru, Porto Rico, Rep. Dominicana, Uruguai, Venezuela
	francês	Canadá, França
	italiano	Itália, Suíça
	inglês	África do Sul, Austrália, Canadá, EUA, Índia, Nova Zelândia, Reino Unido
	mandarim	China, Taiwan
	português	Brasil, Portugal
iTranslate	árabe	Arábia Saudita, Emirados Árabes Unidos, Egito
	espanhol	Espanha, EUA, México
	francês	Canadá, França
	inglês	Reino Unido, EUA, Austrália
	mandarim	China, Taiwan
	português	Brasil, Portugal
Tradutor de Voz	espanhol	América Latina, Espanha, México
	francês	Canadá, França
	inglês	Austrália, Canadá, EUA, Índia, Reino Unido
	mandarim	China, Taiwan
	português	Brasil, Portugal
Fala & Tradução	espanhol	Espanha, México
	francês	Canadá, França
	inglês	Austrália, Canada, EUA, Índia, Reino Unido
	mandarim	China, Taiwan
	português	Brasil, Portugal

Fonte: Elaborado pelo autor.

Ao todo, os cinco aplicativos realizam interpretações automáticas para 49 línguas, desse montante, 24 estão presentes em todos os aplicativos e 18 estão presentes em apenas um aplicativo, conforme mostra a Tabela 13. A maioria das línguas presentes em todos os aplicativos são línguas europeias e, das línguas presentes somente em um aplicativo, oito são asiáticas e oito são europeias, apenas uma, o africânder, é africana.

Tabela 13 - Distribuição das línguas

Línguas presentes em todos os aplicativos	alemão, catalão, coreano, dinamarquês, eslovaco, espanhol, finlandês, francês, grego, híndi, inglês, japonês, mandarim, neerlandês, norueguês, polaco, português, romeno, russo, sueco, tailandês, tcheco, turco, árabe
Línguas presentes em 4 aplicativos	hebraico, húngaro, indonésio, italiano
Línguas presentes em 2 aplicativos	croata, vietnamita
Línguas presentes em 1 aplicativo	africânder, armênio, azerbaijano, basco, bengali, cambojano, cingalês, esloveno, galego, islandês, letão, malaio, moldávio, nepalês, suaíli, sérvio, tâmil, ucraniano
Total de línguas	49

Fonte: Elaborado pelo autor.

A Tabela 14 mostra mais detalhadamente a distribuição das línguas presentes em somente um aplicativo. O Google Tradutor destaca-se dos demais aplicativos, pois oferece interpretação automática para 13 línguas.

Tabela 14 - Línguas presentes em somente um aplicativo

Google Tradutor	africânder, armênio, azerbaijano, bengali, cambojano, cingalês, esloveno, islandês, letão, nepalês, suaíli, sérvio, tâmil
Fala & Tradução	basco, galego, moldávio
SayHi	malaio, ucraniano

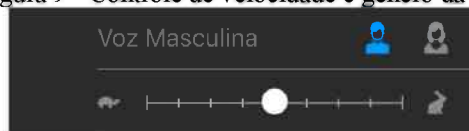
Fonte: Elaborado pelo autor.

A análise dos dados reportados acima mostra que os aplicativos possuem uma limitação de combinações de pares linguísticos que impossibilita a comunicação de determinadas comunidades: falantes monolíngues de azerbaijano, por exemplo, não conseguiriam se comunicar com falantes monolíngues de ucraniano, a menos que utilizassem o inglês como língua intermediária.

5.2.5 Personalização da voz sintetizada

Os aplicativos SayHi, iTranslate e Fala & Tradução possibilitam que os usuários alterem a velocidade e gênero da voz sintetizada. O aplicativo Tradutor de Voz possibilita somente a alteração da velocidade e o Google Tradutor não oferece nenhum dos dois recursos. As alterações são feitas por meio de escalas e ícones, conforme ilustra a Figura 9.

Figura 9 - Controle de velocidade e gênero da voz



Fonte: Aplicativo iTranslate.

5.2.6 Interpretações *offline*

Um dos grandes diferenciais dos dispositivos móveis é a possibilidade de estabelecer conexões de dados de alta mobilidade e em longo alcance da antena. A internet móvel, no entanto, ainda é um serviço caro se comparado à internet tradicional e muitas vezes não é uma opção para viajantes que se distanciem muito dos seus países de origem ou que estejam em localizações remotas de baixa cobertura. Uma alternativa a esses impeditivos é a disponibilidade de recursos *offline*, que podem ser acessados a qualquer hora e lugar independentemente de haver ou não conexão de dados. O download de dados requer que o dispositivo tenha memória extra suficiente.

Dos cinco aplicativos, apenas o SayHi não disponibiliza nenhum recurso *offline*. Os demais possibilitam que os usuários façam o download de determinadas línguas e o Fala & Tradução é o único que possibilita a sincronização de dados em nuvem por meio do iCloud, o serviço de armazenamento de dados em nuvem da Apple. O iTranslate, embora não use o iCloud, possibilita que o usuário crie um *login* e, a partir dele, sincronize e acesse o histórico de traduções e interpretações com um serviço remoto disponibilizado pelo próprio desenvolvedor.

5.2.7 Configurações de fala, sons e reprodução das vozes sintetizadas

Como tem sido mostrado, os cinco aplicativos possuem interfaces com vários recursos para a personalização e configuração dos serviços. Alguns aplicativos, no entanto, oferecem maior personalização do que outros. Vários recursos estão presentes na maioria dos aplicativos, mas a diferença é que alguns aplicativos possibilitam que esses recursos sejam ativados ou desativados. A detecção automática do fim da fala, por exemplo, pode fazer com que o módulo de ASR, por motivos diversos, pare de reconhecer antes de o usuário concluir seu raciocínio, o que compromete a comunicação. Por isso, ter a possibilidade de desativá-la dá ao usuário a liberdade de ele próprio decidir o tamanho do trecho a ser interpretado. Dos cinco aplicativos, apenas o Google Tradutor e o Tradutor de voz, que utilizam o módulo de ASR da Apple, não possibilitam que o usuário controle o ASR.

Interações face a face intermediadas por aplicativos de IA requerem que os usuários dividam suas atenções entre a interface do aplicativo e todos os demais elementos típicos de uma situação comunicativa: linguagem corporal, expressões faciais, o ambiente a sua volta etc. Em um cenário como esse, parece razoável considerar que quanto menos o aplicativo exigir a atenção do usuário, maiores as chances de a comunicação obter bons resultados.

Nesse sentido, os desenvolvedores podem se empenhar na elaboração de interfaces que exijam o mínimo de intervenção do usuário para que os recursos sejam ativados. Um aplicativo ideal empregaria a abordagem conhecida como *handsfree*, que contrasta com a abordagem *tap and hold*. Enquanto a primeira deixa o usuário completamente livre para a utilização de suas mãos para outros fins, a segunda exige que o usuário pressione botões (*tap*) e mantenha-os pressionados (*hold*) para que os recursos funcionem. Os cinco aplicativos de IA selecionados para o presente estudo, no entanto, estão a meio do caminho, empregando a abordagem *tap to translate*, que requer apenas um toque nos botões de comando.

Alertas sonoros e reproduções automáticas das vozes sintetizadas também contribuem para uma utilização mais intuitiva. Dos cinco aplicativos, apenas o Tradutor de Voz não sinaliza o início e o fim do processo de ASR com alertas sonoros. Todos os aplicativos reproduzem automaticamente as vozes sintetizadas, mas no SayHi e no Fala & Tradução esse recurso pode ser desativado.

5.2.8 Mostre-me e tradução de vozes de animais

Um recurso interessante que pode facilitar muito a comunicação entre dois indivíduos que estão se encontrando pela primeira vez e que não falam a mesma língua é o “mostre-me”, disponibilizado somente pelo aplicativo SayHi. Nele os usuários são instruídos a, antes de engajarem-se em uma interpretação automática, mostrarem uma explicação do que acontecerá durante a interação, evitando assim que a pessoa se sinta desconfortável pela presença de um dispositivo móvel gravando o diálogo, conforme ilustra a Figura 10.

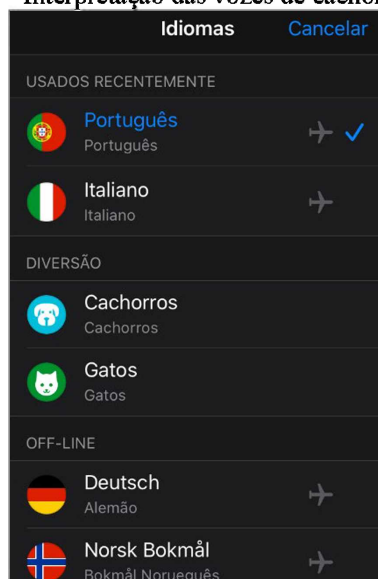
Figura 10 - Recurso "mostre-me"



Fonte: SayHi.

Além desses recursos, cuja utilidade torna-se evidente conforme se utilizam os aplicativos, um recurso disponibilizado somente pelo aplicativo Tradutor de Voz chama a atenção por seu caráter recreativo. Trata-se da interpretação automática das vozes de cachorros e gatos. Inseridas no mesmo painel que as línguas humanas, a interpretação de vozes de cachorros e gatos é apresentada como um recurso voltado para a diversão, conforme ilustra a Figura 11.

Figura 11 - Interpretação das vozes de cachorros e gatos



Fonte: Aplicativo Tradutor de Voz.

Assim como os humanos, os animais expressam suas emoções e desejos por meio de sons e vozes. Ikeda e Ishii (2008), por exemplo, demonstram a possibilidade de usar o reconhecimento das vocalizações produzidas por vacas para distinguir determinadas condições psicológicas. O Google apresentou sua própria versão de um intérprete automático de vozes de animais³¹ em 2010.

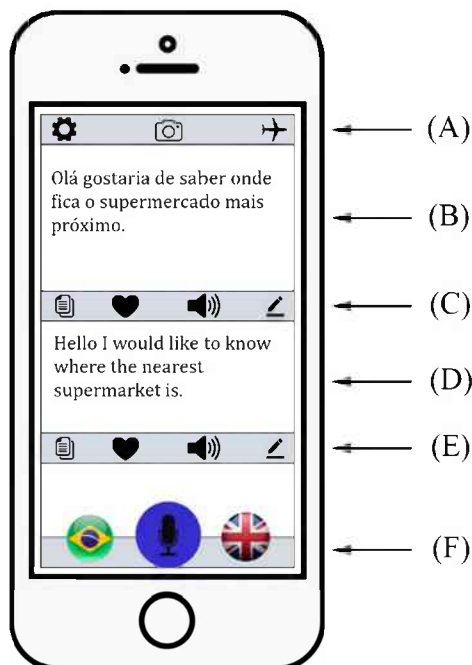
5.3 Utilizando os aplicativos

Na etapa de utilização dos aplicativos os usuários podem testar as escolhas feitas na etapa de configuração. Caso as interpretações e traduções automáticas não se enquadrem no perfil esperado, o usuário pode reconfigurar todos as definições escolhidas. Além disso, é na etapa de utilização que o usuário verifica se os recursos prometidos pelos desenvolvedores nas *webpages* dos aplicativos realmente são oferecidos. Por isso, a etapa de utilização é crucial tanto para o usuário quanto para os desenvolvedores dos aplicativos, pois é nela que o desempenho dos aplicativos se mostra de forma mais evidente, podendo levar o usuário a desinstalar o aplicativo, continuar utilizando uma versão mais básica ou até mesmo adquirir um plano semanal, mensal ou anual.

³¹ <https://archive.google.com/translateforanimals/>

Os leiautes dos cinco aplicativos, embora possuam esquemas de cores e ícones diferentes, são constituídos por quantidades semelhantes de elementos, organizados em seis campos verticais, conforme ilustra a Figura 12.

Figura 12 - Leiaute dos aplicativos



Fonte: Elaborado pelo autor.

Nos campos (A) e (F) são disponibilizados ícones relacionados à configuração dos recursos, seleção dos pares linguísticos, botões para a interpretação e tradução automáticas, tradução de texto em imagens, traduções *offline* e guias de conversação. Nos campos (B) e (D) são disponibilizadas as transcrições do trecho de fala reconhecido e do trecho de fala interpretado. Os campos (C) e (E) contêm ícones para a cópia dos trechos, seleção de trechos favoritos ou salvamento, reprodução das gravações e edição dos trechos.

Além dos recursos apresentados na etapa de configuração, novos recursos são oferecidos durante a utilização dos aplicativos. Em geral, esses recursos não podem ser configurados e são disponibilizados na tela inicial do aplicativo.

5.3.1 Guias de conversação

Dos cinco aplicativos, apenas o iTranslate e o Tradutor de Voz oferecem guias de conversação. O guia de conversação do aplicativo Tradutor de Voz possui 16 categorias e está disponível para 27 línguas. O guia de conversação do iTranslate está disponível para 18 línguas, é dividido em nove categorias e várias subcategorias, a depender do tópico, como ilustra a Figura 13.

Figura 13 – Guia de conversação do iTranslate



Fonte: Aplicativo iTranslate.

5.3.2 Traduções e reconhecimento automático alternativos

Disponibilizadas na interface do aplicativo, abaixo do campo em que são mostradas as transcrições, as traduções alternativas possibilitam que o usuário escolha a tradução que mais se adeque ao contexto, caso não haja compreensão por parte do interlocutor. Dos cinco aplicativos, apenas o Google Tradutor e o iTranslate disponibilizam traduções alternativas. No aplicativo SayHi, ao final do ASR, caso tenha ocorrido algum erro de reconhecimento, o usuário pode repetir o processo de reconhecimento antes de reproduzir o áudio ao interlocutor, conforme ilustra a Figura 14.

Figura 14 - Reconhecimento automático alternativo



Fonte: Aplicativo SayHi.

5.3.3 Recursos para a aprendizagem de línguas

Além dos serviços de tradução e interpretação automáticas, alguns aplicativos de IA oferecem também recursos para quem deseja aprender outras línguas. Dentre eles estão a pronúncia de palavras e frases, conjugações verbais, definições de palavras, transliterações e guias de conversação. Como a IA requer a pronúncia de palavras e frases, todos os cinco aplicativos apresentam esse recurso. As conjugações verbais são oferecidas apenas pelo iTranslate e as definições de palavras somente pelo Google Tradutor, para um número limitado de línguas. A Figura 15 ilustra o recurso de conjugação verbal do iTranslate.

Figura 15 - Conjugação verbal



Fonte: Aplicativo iTranslate.

As conjugações verbais do iTranslate são oferecidas nos modos indicativo e subjuntivo e em todos os tempos verbais. Além disso, no painel de conjugações o usuário pode ouvir as diferentes pronúncias de um mesmo verbo. Verbos de todos os idiomas disponíveis no iTranslate podem ser conjugados.

5.3.4 Outros recursos

Recursos como traduções de textos da área de transferência do dispositivo, histórico e salvamento de interpretações e traduções também são oferecidos pelos aplicativos. Todos os cinco aplicativos oferecem traduções de textos da área de transferência e histórico de interpretações e traduções. O salvamento de interpretações e traduções é feito na própria interface do aplicativo e é chamado de “favoritos”, o recurso é oferecido pelo Google Tradutor, iTranslate e Tradutor de Voz.

A tradução de textos da área de transferência possibilita que os usuários ampliem a aplicação dos serviços de interpretação e tradução para textos provenientes de outras fontes, do whatsapp ou de *websites*, por exemplo. O histórico e o salvamento de interpretações e traduções facilitam o acesso a traduções já realizadas e possibilita que os usuários criem seus próprios repertórios de traduções a semelhança dos guias de conversação.

CAPÍTULO 6 – AS EXPECTATIVAS E PERCEPÇÕES DOS USUÁRIOS

Neste capítulo apresento a análise dos comentários fidedignos compartilhados pelos usuários brasileiros dos cinco aplicativos de IA disponíveis na App Store descritos no **Capítulo 5**. A descrição dos recursos e serviços oferecidos pelos cinco aplicativos fundamentou-se em minha experiência com os aplicativos, nas informações disponíveis nas interfaces dos aplicativos e nas descrições oferecidas pelos desenvolvedores nas *webpages* dos aplicativos na App Store. Analiso nesta seção os comentários dos usuários com o objetivo de saber quais são as principais expectativas dos usuários brasileiros desse tipo de tecnologia.

Trata-se de uma análise tanto quantitativa quanto qualitativa que, diferentemente da descrição dos aplicativos, fundamenta-se na experiência dos usuários. Não somente na experiência de alguns usuários, mas nas percepções mais recorrentes citadas pelos usuários em 200 comentários³² fidedignos compartilhados espontaneamente por eles e selecionados aleatoriamente para a análise, conforme os critérios apresentados no **Capítulo 4**.

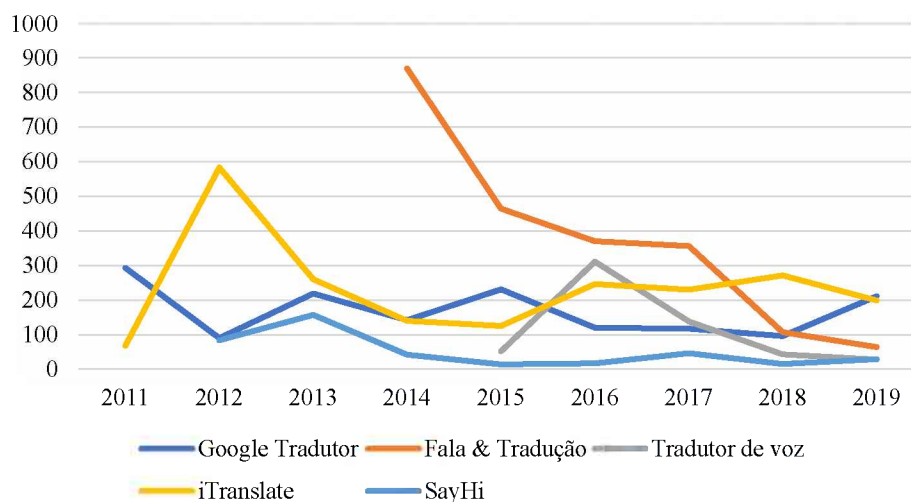
A Figura 16 mostra o número de comentários por aplicativo desde seus respectivos anos de lançamento na App Store até o ano de 2019. No gráfico é possível visualizar que nos últimos anos, o número de comentários tem diminuído para todos os aplicativos. Além disso, o Google Tradutor e o Fala & Tradução foram os únicos aplicativos cujos números de comentários foram maiores em seus respectivos anos de lançamentos do que em qualquer outro ano.

Dos cinco aplicativos, o iTranslate, o SayHi e o Tradutor de Voz receberam mais comentários no ano seguinte a seus lançamentos na App Store e o Google Tradutor e o Fala & Tradução receberam menos comentários no mesmo período. A queda no número de comentários dos usuários do aplicativo Fala & Tradução, que inclusive é o aplicativo que mais recebeu comentários em seu primeiro ano de lançamento, é a que mais se destaca

³² Os comentários podem ser acessados por meio das *webpages* dos aplicativos, mencionadas no **Capítulo 4** e inseridas abaixo para efeito de consulta: **SayHi**. Disponível em: <<https://itunes.apple.com/br/app/wwdc/id437818260?mt=8>>. Acesso em: 28 jul. 2020; **iTranslate**. Disponível em: <<https://apps.apple.com/br/app/id288113403?ign-mpt=uo%3D4>>. Acesso em: 28 jul. 2020; **Fala e Tradução**. Disponível em: <<https://apps.apple.com/br/app/id804641004>>. Acesso em: 28 jul. 2020; **Google Tradutor**. Disponível em: <<https://apps.apple.com/br/app/google-tradutor/id414706506>>. Acesso em: 28 jul. 2020; **Tradutor de Voz**. Disponível em: <<https://apps.apple.com/br/app/tradutor-de-voz/id1036725690>>. Acesso em: 28 jul. 2020.

no gráfico. Se a queda no número de comentários em todos os aplicativos mostrar-se uma tendência nos próximos anos, pode tornar-se ainda mais difícil usar os comentários como fonte de dados para a análise dos aplicativos.

Figura 16 - Comentários por ano



Fonte: Elaborado pelo autor.

Os comentários foram coletados das *webpages* e colocados em cinco documentos txt, nos quais para cada comentário foi atribuído um número a fim de facilitar a sua posterior localização. Para a análise, utilizei a ferramenta *online* gratuita Taguette³³, por meio da qual os comentários foram etiquetados conforme as percepções veiculadas em cada trecho dos comentários.

Assim, a etiquetagem do conteúdo dos comentários deu origem a 10 categorias, que são apresentadas a seguir em ordem alfabética:

1. Percepções sobre a interface (PSI);
2. Percepções sobre a TA/IA (PSTA/IA);
3. Percepções sobre a tradução de texto em imagens (PSTTI);
4. Percepções sobre a utilidade (PSU);
5. Percepções sobre o módulo de ASR (PSMASR);
6. Percepções sobre o modo offline (PSMO);
7. Percepções sobre o módulo de TTS (PSMTTS);
8. Percepções sobre outros aplicativos (PSOA);
9. Percepções sobre planos e assinaturas (PSPA);
10. Percepções sobre o desempenho (PSD);

³³ Disponível em: <<https://www.taguette.org>>. Acesso em 18 jun. 2019.

Devido às limitações da configuração do presente estudo, optei por aplicar à classificação das percepções um método de avaliação de confiabilidade intra-observador, por meio do qual as percepções foram re-etiquetadas 180 dias após a primeira etiquetagem por meio da mesma ferramenta de análise e pelo mesmo observador. Julgo que esse período de afastamento dos dados pode impactar na forma com que o observador classifica os comentários e os períodos que os compõem. A variação observada entre a primeira e a segunda etiquetagens pode servir para medir a confiabilidade da primeira etiquetagem e, assim, relativizá-la.

As maiores variações entre a primeira e a segunda etiquetagem ocorrem na percentagem de percepções positivas sobre a utilização no aplicativo Tradutor de Voz e na percentagem de distribuição da categoria PSMO, no aplicativo Google Tradutor. No primeiro caso, a variação se deu em 6% mais etiquetas atribuídas a percepções positivas sobre a utilização no aplicativo Tradutor de voz; no segundo caso, a variação se deu em 7% menos etiquetas atribuídas à categoria PSMO referente ao aplicativo Google Tradutor. Para as demais categorias, as variações ficaram entre 5% mais etiquetas atribuídas e 5% menos etiquetas, respectivamente. Das 150 análises feitas nas duas etiquetagens, em 34% delas não foram observadas quaisquer variações em cerca de 19% foi observada a alteração de 1% e em 12% foram observadas variações de -1%, ou seja, a ocorrência de perda de etiquetas por parte de alguns aplicativos e em 12% foram observadas variações de 2%, ou seja, a ocorrência de ganhos de etiquetas por parte de alguns aplicativos. O aplicativo iTranslate, que na primeira etiquetagem não havia recebido nenhuma etiqueta referente a percepções negativas sobre a interface, passou a configurar entre os aplicativos que receberam essa etiqueta. Tal fato não ocorreu com os demais aplicativos.

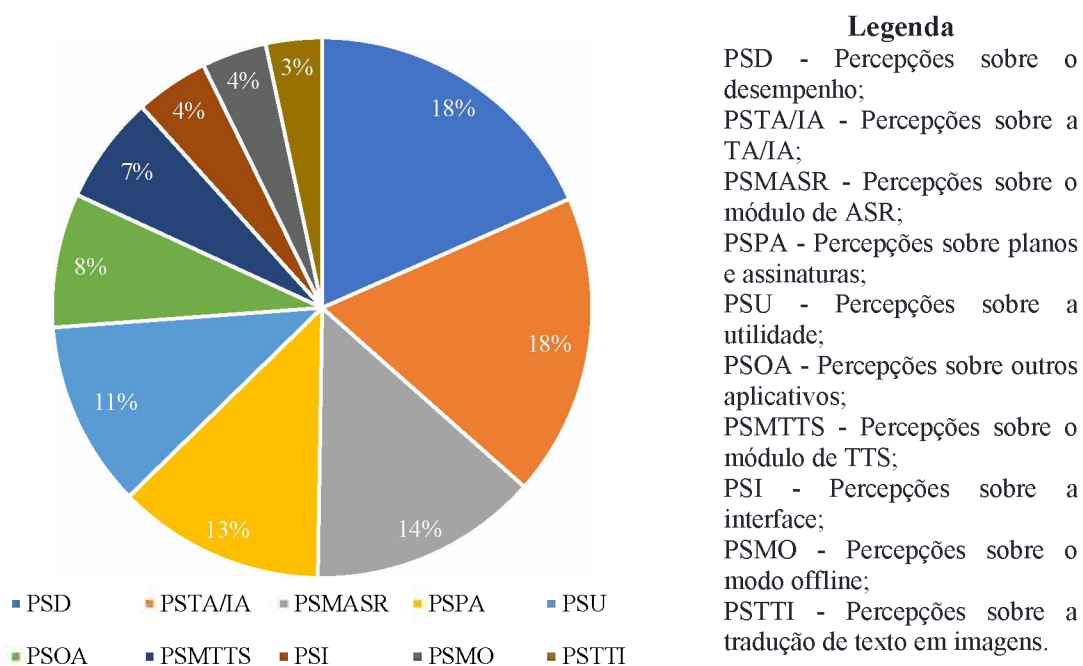
Ao compartilhar suas percepções, os usuários podem manifestar tanto suas satisfações quanto suas insatisfações, fazendo com que as percepções veiculadas por cada categoria possam ser positivas ou negativas. O comentário “tradução sem erros”, por exemplo, é uma percepção sobre a TA/IA, mas também é uma percepção positiva.

Devido às limitações tecnológicas da ferramenta *online* Taguette, que não possibilita a criação de subetiquetas, conforme já abordado no **Capítulo 4**, elaborei duas etiquetas especiais: percepções positivas (PP) e percepções negativas (PN). Essas duas categorias especiais podem coocorrer com as 10 outras categorias e mostram-se bastante úteis para comentários que veiculam percepções genéricas como as expressas nos

comentários “aplicativo excelente!” e “aplicativo péssimo!”, nos quais só é possível identificar se são positivos ou negativos.

Nos 200 comentários analisados foram identificadas 425 percepções que se enquadram nas 10 categorias antes listadas. A Figura 17 mostra a distribuição das categorias em todos os aplicativos, com exceção das categorias especiais PP e PN. As categorias mais frequentes são PSD e PSTA/IA, cada uma ocupando 18% das percepções compartilhadas pelos usuários. Em seguida, as categorias mais frequentes são PSMASR, PSPA e PSU, cujas frequências somadas ocupam 38% das percepções e as categorias menos frequentes são PSOA, PSTTS, PSI, PSMO e PSTTI, cujas frequências somadas ocupam 26% das percepções compartilhadas pelos usuários.

Figura 17 – Distribuição das percepções

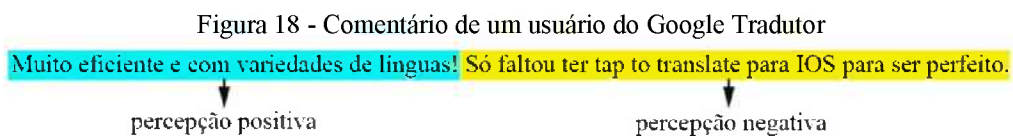


Fonte: Elaborado pelo autor.

6.1 Percepções positivas e percepções negativas

A maioria dos comentários veiculam tanto percepções positivas (PP) quanto percepções negativas (PN) para diferentes aspectos relacionados aos aplicativos. Na Figura 18, por exemplo, o usuário comenta que o Google Tradutor é muito eficiente e

possui várias opções de línguas, ou seja, expõe percepções positivas sobre o desempenho do aplicativo e a quantidade de línguas oferecidas. Apesar disso, no mesmo comentário, o usuário expõe uma percepção negativa sobre a interface do aplicativo, comentando que o recurso *tap to translate* é um item que está em falta.



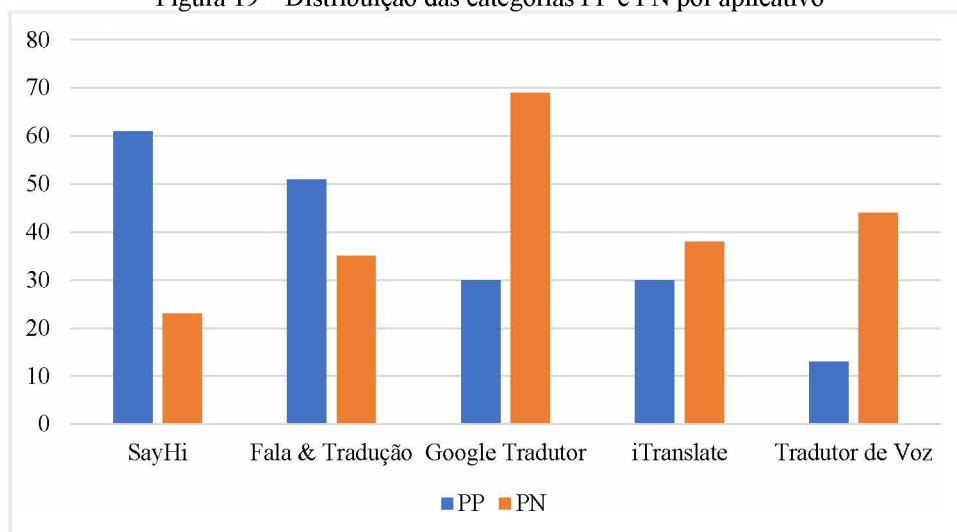
Fonte: Google Tradutor.

Apesar de a App Store não estabelecer critérios ou oferecer aos seus usuários quaisquer orientações sobre o modo com que os comentários devam ser feitos, a maioria dos comentários aleatórios analisados estão estruturados de forma que o desencadeamento de expectativas e percepções dos usuários encontra-se mais ou menos organizado dentro de períodos, simples ou compostos. Desse modo, a unidade sintática selecionada para a classificação e análise dos comentários foi o período, ou seja, o enunciado construído por uma ou mais orações e cujo significado pode ser delimitada do começo da primeira oração até a conclusão do período, por meio de pontuação específica.

Nos comentários analisados, a incidência de percepções positivas e negativas por aplicativo é apresentada na Figura 19. O SayHi é o aplicativo com o maior número de percepções positivas e menor número de percepções negativas. O Fala & Tradução é o aplicativo com o segundo maior número de percepções positivas e segundo menor número de percepções negativas. Os aplicativos Google Tradutor e iTranslate possuem o mesmo número de percepções positivas, mas o Google Tradutor é o aplicativo com o maior número de percepções negativas de todos os aplicativos analisados.

Se considerarmos a diferença entre o número de percepções positivas e percepções negativas, os aplicativos podem ser alinhados em uma escala que começa com o SayHi na primeira posição, o Fala & Tradução na segunda, o iTranslate na terceira, o Tradutor de voz na quarta e o Google Tradutor na última posição. Esses três últimos aplicativos receberam mais percepções negativas do que positivas.

Figura 19 – Distribuição das categorias PP e PN por aplicativo



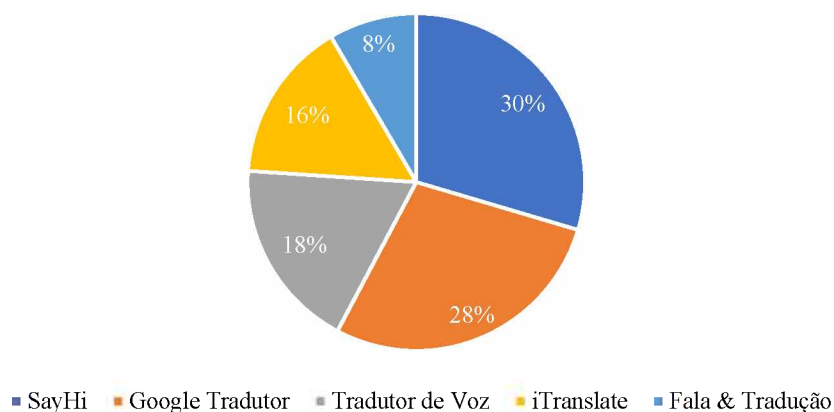
Fonte: Elaborado pelo autor.

6.2 Percepções sobre o desempenho

Desempenho é o modo com que o aplicativo se comporta levando em conta a sua eficiência, o tempo de resposta e o modo com que as tarefas a que se propõem são executadas (CORTELLESA; MARCO; INVERARDI, 2011). Nos comentários analisados, somente as percepções que se referem ao desempenho geral dos aplicativos são classificadas como percepções sobre o desempenho (PSD), percepções que se referem ao desempenho dos módulos ASR, TA e TTS são classificadas com as categorias correspondentes. No comentário “reconhecimento de palavras lento”, por exemplo, as categorias atribuídas são PSMASR e PN, pois trata-se de uma percepção negativa concedida ao módulo ASR.

As percepções do tipo PSD correspondem a 18% das percepções identificadas no conjunto de comentários analisado. Os aplicativos SayHi e Google Tradutor são os aplicativos com o maior número de percepções do tipo PSD e o Fala & Tradução é o aplicativo com o menor número de percepções desse tipo, como mostra a Figura 20.

Figura 20 - Distribuição da categoria PSD por aplicativo

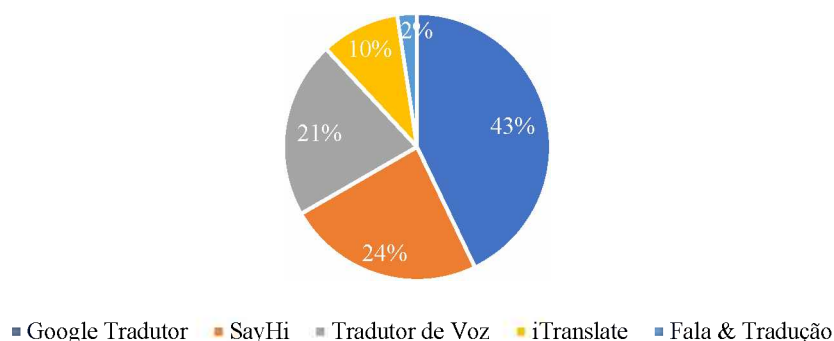


Fonte: Elaborado pelo autor.

Nos comentários compartilhados, os usuários referem-se ao desempenho dos aplicativos por meio de adjetivos como “lento”, “rápido”, “eficiente” e “veloz” e empregam verbos como “demorar” e “atrasar”, ou seja, percepções sobre o tempo de resposta em que os aplicativos entregam os serviços oferecidos. Além disso, os usuários usam também os verbos “travar”, “funcionar”, “carregar”, “abrir”, “fechar”, “sair” e “iniciar” para relatar comportamentos inesperados e o funcionamento ou não dos aplicativos.

A maioria das percepções sobre o desempenho estão vinculadas a percepções negativas sobre o mal funcionamento dos aplicativos e à demora/lentidão com que os comandos são executados. O aplicativo que mais recebeu comentários negativos a respeito do seu desempenho foi o Google Tradutor e o que menos recebeu comentários desse tipo foi o Fala & Tradução, conforme mostra a Figura 21.

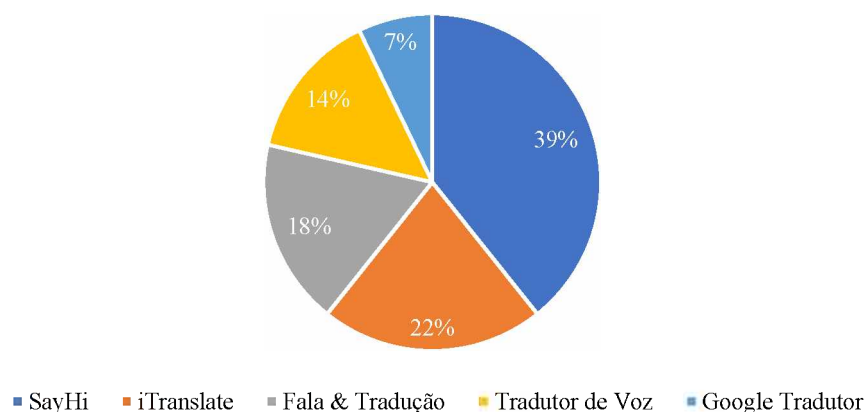
Figura 21 – Distribuição das percepções negativas sobre o desempenho por aplicativo



Fonte: Elaborado pelo autor.

As percepções positivas sobre o desempenho dos aplicativos estão vinculadas principalmente à rapidez com que os comandos são executados. O aplicativo que mais recebeu comentários positivos a respeito do seu desempenho foi o SayHi e o que menos recebeu comentários desse tipo foi o Google Tradutor, conforme mostra a Figura 22.

Figura 22 – Distribuição das percepções positivas sobre o desempenho por aplicativo



Fonte: Elaborado pelo autor.

A partir do conjunto de percepções positivas e negativas sobre o desempenho é possível identificar as principais expectativas dos usuários quanto ao desempenho dos aplicativos. Expressões como “traduções na hora” e “traduções num piscar de olhos” revelam que os usuários esperam que os aplicativos ofereçam respostas rápidas aos seus comandos, independentemente do tipo de conexão de dados, de modo fácil e intuitivo. Além disso, os usuários esperam que as atualizações aperfeiçoem o desempenho, sem retirar ou prejudicar os recursos existentes. Essas expectativas, portanto, condizem, como exposto no **Capítulo 2**, aos esforços dos pesquisadores e desenvolvedores envolvidos com a IA desde o surgimento dessa tecnologia.

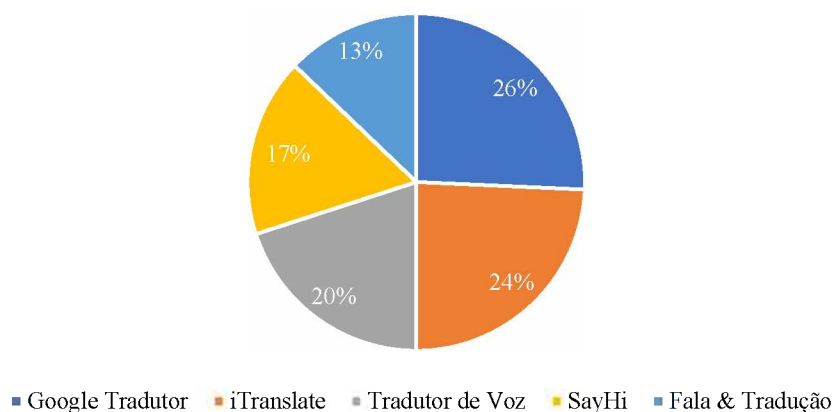
6.3 Percepções sobre a TA/IA

Os cinco aplicativos oferecem tanto o serviço de interpretação automática quanto o serviço de tradução automática. No conjunto de comentários, no entanto, não há nenhum comentário que faça uso da palavra “interpretação” ou da expressão “interpretação automática”. Ambos os serviços são mencionados como “tradução” e, exceto quando os usuários fazem menção ao ASR e à qualidade da voz sintetizada, é praticamente impossível identificar a que serviço o usuário se refere. Por isso, as

percepções sobre a TA/IA (PSTA/IA) referem-se tanto à tradução automática enquanto serviço oferecido pelos aplicativos quanto à tradução automática enquanto um dos três módulos tecnológicos que compõem a IA.

As percepções do tipo PSTA/IA correspondem a 18% das percepções identificadas no conjunto de comentários analisado. O aplicativo Google Tradutor é o aplicativo com o maior número de percepções do tipo PSTA/IA e o Fala & Tradução é o aplicativo com o menor número de percepções desse tipo, conforme mostra a Figura 23.

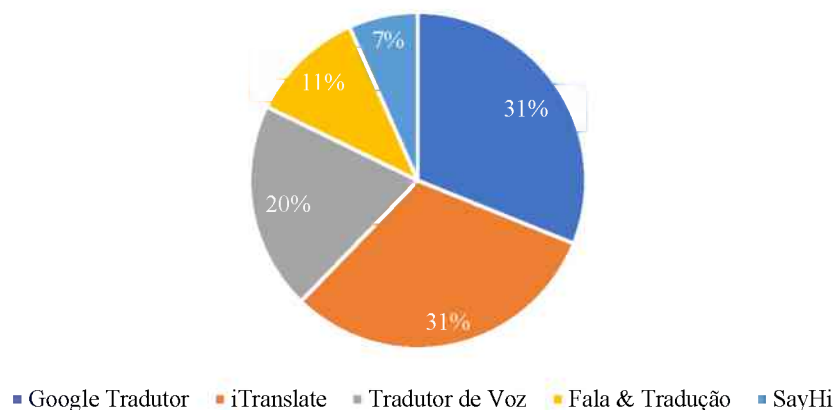
Figura 23 - Distribuição da categoria PSTA/IA por aplicativo



Fonte: Elaborado pelo autor.

Os usuários referem-se à tradução e interpretação automáticas por meio de adjetivos como “correta” e “errada”, alguns referem-se a traduções “assertivas” e outros a traduções “fiéis” e poucos a traduções “informais”. A maioria das percepções do tipo PSTA/IA estão vinculadas a percepções negativas sobre a estrutura das frases e a incapacidade de os aplicativos traduzirem expressões idiomáticas. Algumas percepções referem-se à ausência de formalidade e a impossibilidade de se escolher palavras sinônimas. Há apenas um comentário sobre tradução ofensiva. Os aplicativos que mais receberam comentários negativos sobre a TA/IA foram o Google Tradutor e o iTranslate e o que menos recebeu comentários desse tipo foi o SayHi, conforme mostra a Figura 24.

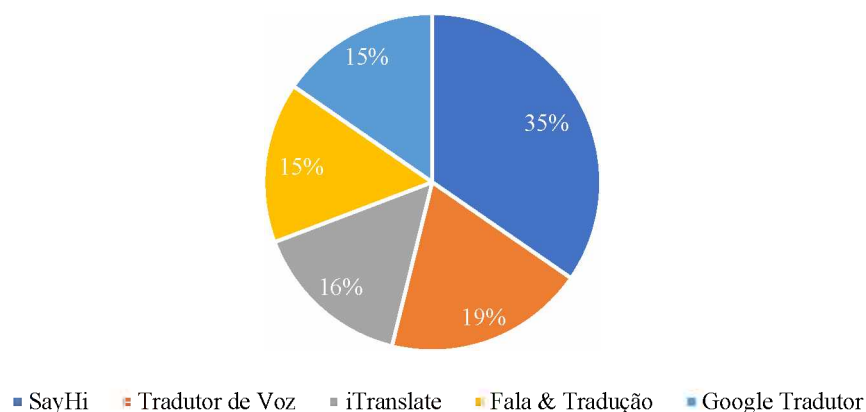
Figura 24 – Distribuição das percepções negativas sobre a TA/IA por aplicativo



Fonte: Elaborado pelo autor.

As percepções positivas sobre a TA/IA são, em sua grande maioria, de caráter elogioso e não trazem nenhum detalhamento sobre quais aspectos foram considerados para a avaliação positiva. Apenas uma percepção menciona o tamanho do vocabulário como algo positivo. O aplicativo que mais recebeu comentários positivos a respeito da TA/IA foi o SayHi e os que menos receberam comentários desse tipo foram o Fala & Tradução e o Google Tradutor, conforme mostra a Figura 25.

Figura 25 – Distribuição das percepções positivas sobre a TA/IA por aplicativo



Fonte: Elaborado pelo autor.

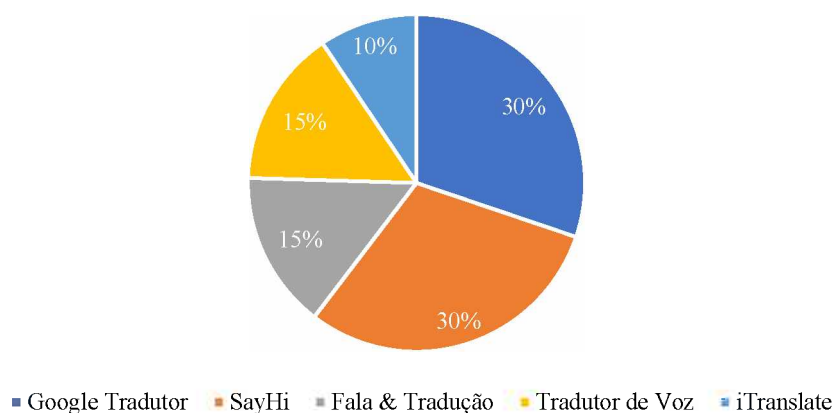
As percepções sobre a TA/IA revelam que os usuários esperam que os aplicativos produzam traduções livres de erros, que mantenham o grau de formalidade expresso pelo trecho original, que saibam colocar as palavras traduzidas na ordem correta para o idioma de chegada e que sejam capazes de traduzir expressões idiomáticas. Além disso, os usuários esperam que o módulo de TA seja capaz de oferecer traduções sinônimas.

6.4 Percepções sobre o módulo ASR

O módulo ASR é o primeiro de três módulos que compõem um sistema de IA, é responsável pelo reconhecimento automático do sinal acústico emitido pelo usuário, transformando-o em texto que é repassado para o módulo de TA. Desde os primórdios da IA, alguns estudos têm tentado identificar quais erros são tipicamente produzidos pelo módulo ASR e o impacto deles nas etapas de TA e TTS.

As percepções sobre o módulo ASR (PSMASR) correspondem a 14% das percepções identificadas no conjunto de comentários analisado. Os usuários referem-se ao módulo ASR por meio das expressões “captação de voz”, “entendimento de voz”, “compreensão de voz”, “identificação de voz”, “leitor de voz”, “mecanismo de conversação”, “fala em áudio” e “reconhecimento de palavras”. Os aplicativos Google Tradutor e SayHi são os aplicativos com o maior número de percepções do tipo PSMASR e o iTranslate é o aplicativo com o menor número de percepções desse tipo, conforme mostra a Figura 26.

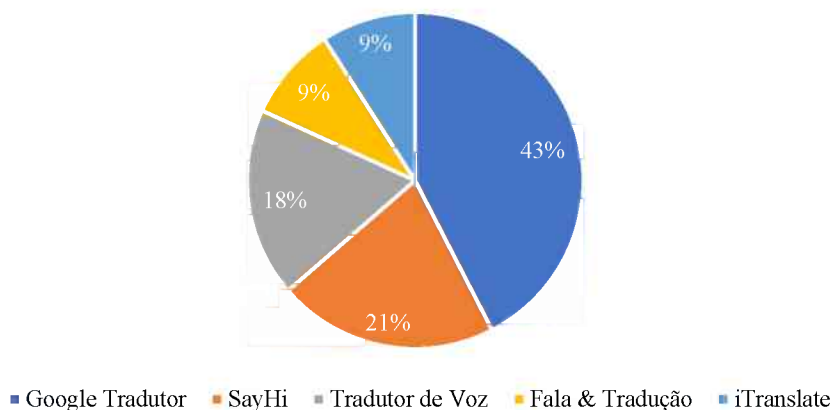
Figura 26 - Distribuição da categoria PSMASR por aplicativo



Fonte: Elaborado pelo autor.

A maioria das percepções do tipo PSMASR estão vinculadas a percepções negativas sobre os erros de reconhecimento automático e a incapacidade de o módulo reconhecer pausas, pontuar as orações e recuperar elementos prosódicos. Algumas percepções referem-se à demora/lentidão com que o reconhecimento de fala é feito. Há poucas percepções sobre a inversão da ordem das palavras. O aplicativo que mais recebeu comentários negativos sobre o módulo ASR foi o Google Tradutor e o que menos recebeu comentários desse tipo foi o iTranslate, conforme mostra a Figura 27.

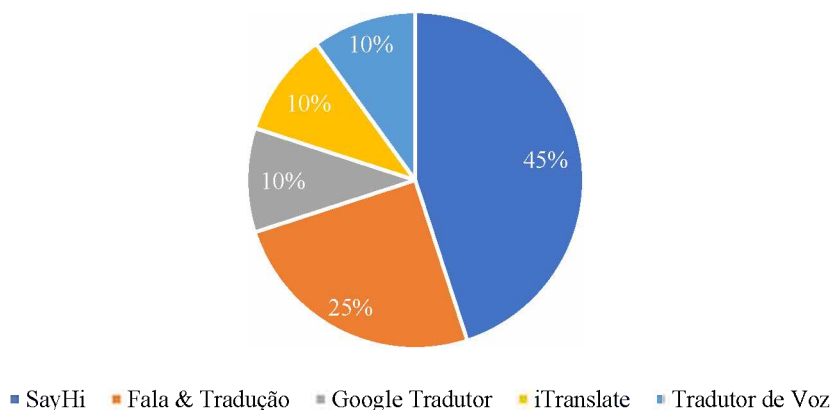
Figura 27 – Distribuição das percepções negativas sobre o módulo ASR



Fonte: Elaborado pelo autor.

As percepções positivas sobre o módulo ASR referem-se à rapidez com que as frases são reconhecidas e muitas delas trazem poucas informações sobre o motivo de o ASR ter sido considerado “perfeito”, “impecável” ou “assertivo”. O aplicativo que mais recebeu comentários positivos a respeito do módulo ASR foi o SayHi e os que menos receberam comentários desse tipo foram o Google Tradutor, o iTranslate e o Tradutor de Voz, conforme mostra a Figura 28.

Figura 28 - Percepções positivas sobre o módulo ASR



Fonte: Elaborado pelo autor.

A análise das percepções sobre o módulo ASR revela que os usuários esperam que o ASR seja rápido e “perfeito”, ou seja, totalmente livre de erros. Como o módulo ASR é o primeiro dos três módulos que compõem um sistema de interpretação automática, nas percepções compartilhadas pelos usuários parece haver uma expectativa maior de que ele seja mais eficiente do que os demais módulos. Os usuários também esperam que o módulo

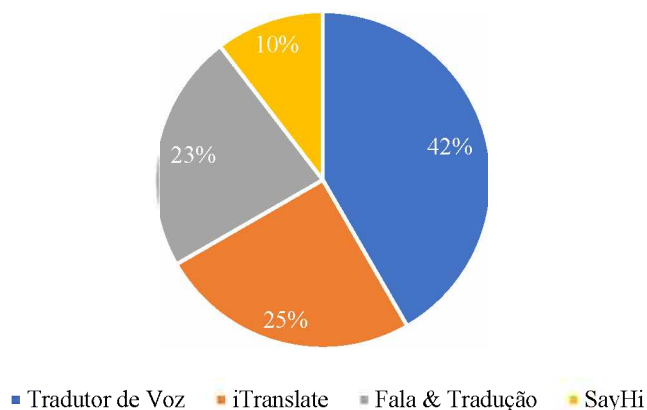
ASR seja capaz de reconhecer elementos prosódicos como pausas e hesitações e seja capaz de reconhecer até mesmo nomes próprios, jargões, regionalismos e neologismos.

6.5 Percepções sobre planos e assinaturas

Dentre as várias percepções veiculadas pelos comentários dos usuários, as percepções sobre planos e assinaturas (PSPA) referem-se a questões como os preços dos aplicativos e, em alguns casos, à relação custo-benefício observada pelos usuários. Nos comentários, essas percepções são expressas por meio de colocações como “muito caro”, “(não) vale a pena”, “preço absurdo/razoável”, “anuidade/mensalidade cara”, “aplicativo gratuito” e “custo zero”.

As percepções do tipo PSPA correspondem a 13% das percepções identificadas no conjunto de comentários analisado. O Tradutor de Voz é o aplicativo que recebeu o maior número de percepções do tipo PSPA e o SayHi é o aplicativo com o menor número de percepções desse tipo, conforme mostra a Figura 29. Tanto o SayHi quanto o Google Tradutor são aplicativos gratuitos, mas somente o primeiro recebeu percepções do tipo PSPA.

Figura 29 - Distribuição da categoria PSPA por aplicativo

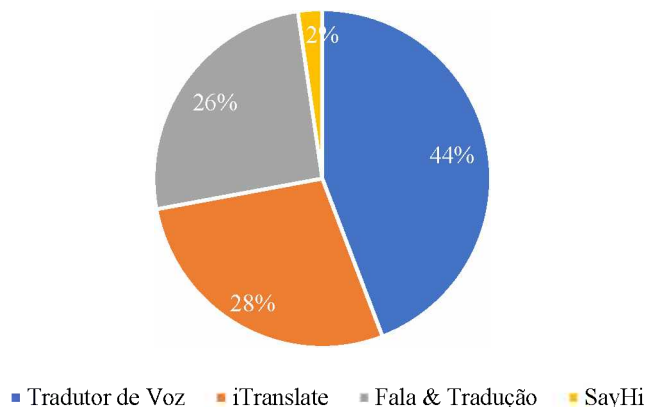


Fonte: Elaborado pelo autor.

No conjunto de comentários analisado, as percepções do tipo PSPA são as que mais veiculam impressões negativas, estando vinculadas a palavras de baixo calão e expressões que demonstram o grau de frustração dos usuários. O aplicativo que mais recebeu comentários negativos a respeito dos planos e assinaturas oferecidos foi o

Tradutor de Voz e o que menos recebeu comentários desse tipo foi o SayHi, conforme mostra a Figura 30.

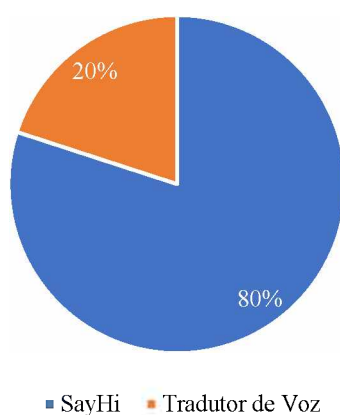
Figura 30 – Distribuição das percepções negativas sobre os planos e assinaturas por aplicativo



Fonte: Elaborado pelo autor.

No montante de comentários, somente cinco comentários veiculam percepções positivas sobre os planos e assinaturas, sendo essa, portanto, a percepção que menos ocorreu em toda a análise. As percepções positivas sobre os planos e assinaturas estão vinculadas principalmente ao custo-benefício percebido pelos usuários. Somente dois aplicativos receberam comentários positivos a respeito dos planos e assinaturas: o SayHi e o Tradutor de Voz, conforme mostra a Figura 31.

Figura 31 – Distribuição das percepções positivas sobre os planos e assinaturas por aplicativo



Fonte: Elaborado pelo autor.

De acordo com as percepções sobre planos e assinaturas, os usuários dos aplicativos de interpretação automática esperam um serviço barato e eficiente. Quando o Google Tradutor é colocado como ponto de referência, os usuários esperam que o serviço

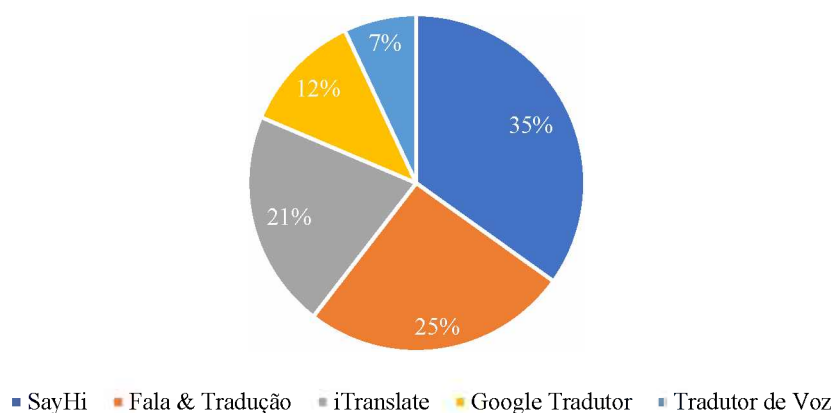
seja gratuito e eficiente, oferecendo traduções e interpretações para o maior número de pares linguísticos possível. Os usuários esperam poder testar os aplicativos por períodos maiores do que os oferecidos pelos aplicativos e esperam ser reembolsados caso não concordem com a cobrança efetuada em seus cartões de crédito.

6.6 Percepções sobre a utilidade

Os cinco aplicativos apresentam-se como tradutores e intérpretes automáticos que também podem ser utilizados na aprendizagem de idiomas. As percepções sobre a utilidade (PSU) referem-se a percepções nas quais os usuários compartilham para que fins específicos eles utilizam os aplicativos.

As percepções do tipo PSU correspondem a 11% das percepções identificadas no conjunto de comentários analisado. O SayHi é o aplicativo que recebeu o maior número de percepções do tipo PSU e o Tradutor de Voz é o aplicativo com o menor número de percepções desse tipo, como mostra a Figura 32.

Figura 32 - Distribuição da categoria PSU por aplicativo



Fonte: Elaborado pelo autor.

No conjunto de comentários analisado, além do compartilhamento da finalidade com a qual eles utilizam os aplicativos, alguns usuários compartilham as situações em que lançam mão de aplicativos desse tipo e se suas expectativas naqueles dados contextos foram atendidas. A maioria dos usuários afirmam ter usado os aplicativos em viagens para países onde o inglês não é uma língua oficial.

Vários usuários empregam os aplicativos no aprendizado novos idiomas e, enquanto alguns dizem que os aplicativos são como que uma ferramenta a mais para este fim, outros dizem que os aplicativos são as únicas ferramentas que eles utilizam para tal.

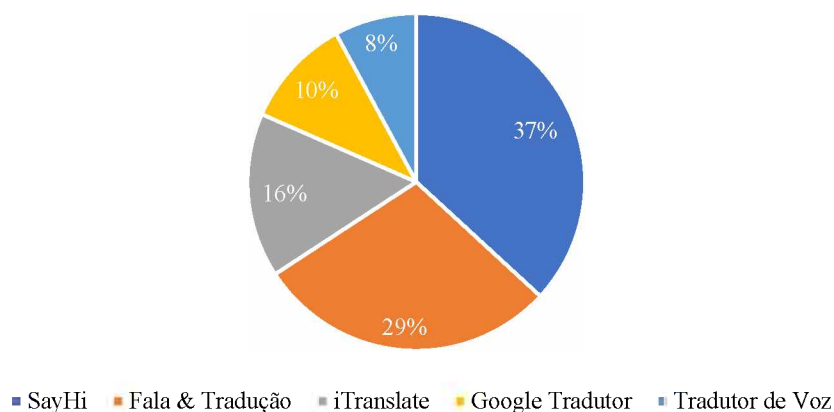
Poucos afirmam que os aplicativos podem substituir as escolas de idiomas. Vários usuários afirmam usar os aplicativos para praticar e tirar dúvidas sobre a pronúncia de palavras.

Alguns comentários sobre o contexto em que os aplicativos são empregados aparecem uma única vez. Dentre as várias finalidades apontadas pelos usuários, identifiquei as seguintes: comunicação com jogadores estrangeiros em partidas de jogos *online*; discussões em redes sociais; realização de compras no exterior; tradução de menus em restaurantes e tradução de páginas de livros por meio da câmera do dispositivo móvel.

A maioria das percepções sobre a utilidade dos aplicativos estão vinculadas a percepções positivas. O iTranslate, que recebeu três percepções negativas para a categoria PSU, foi o único aplicativo a receber percepções desse tipo. As três percepções negativas não foram detalhadas, constando apenas de comentários sobre o contexto de utilização e comentários do tipo “me deixou na mão”, “passei vergonha” e “não gostei”.

As percepções positivas estão vinculadas à possibilidade de comunicar-se com os nativos de determinada região em viagens internacionais e com o aprendizado de línguas por um custo módico e, no caso do SayHi e do Google Tradutor, de forma gratuita. O aplicativo que mais recebeu comentários positivos para a categoria PSU foi o SayHi e o que menos recebeu comentários desse tipo foi o Tradutor de Voz, conforme mostra a Figura 33.

Figura 33 – Distribuição das percepções positivas sobre a utilização por aplicativo



Fonte: Elaborado pelo autor.

As percepções sobre a utilidade dos aplicativos revelam que os usuários esperam que os aplicativos possam ser utilizados para gerar impactos positivos em suas experiências no exterior e promover o aprendizado de novas línguas. Expressões como “me deixou na mão” e “passei vergonha”, anteriormente citadas, parecem indicar que os

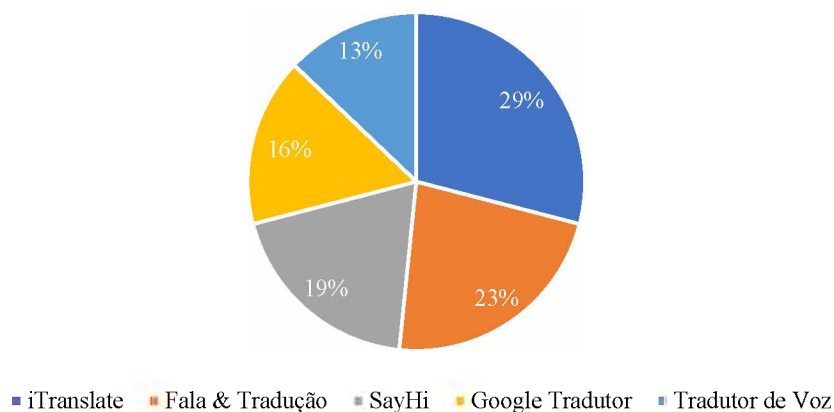
usuários esperam que os aplicativos não deixem de funcionar quando eles mais precisam e que a falha dos aplicativos deixa exposta suas incapacidades de se comunicar em determinadas línguas, algo que lhes gera constrangimento.

6.7 Percepções sobre outros aplicativos

Os comentários compartilhados pelos usuários de aplicativos de IA também são espaços em que os usuários podem emitir opiniões sobre outros aplicativos. Ao compartilhar suas percepções, alguns usuários comparam os aplicativos que instalaram a aplicativos utilizados anteriormente para traduzir e interpretar automaticamente.

As percepções sobre outros aplicativos (PSOA) correspondem a 8% das percepções identificadas no conjunto de comentários analisado. O aplicativo iTranslate é o aplicativo com o maior número de percepções do tipo PSOA e o Tradutor de Voz é o aplicativo com o menor número de percepções desse tipo, conforme mostra a Figura 34.

Figura 34 - Distribuição da categoria PSOA por aplicativo



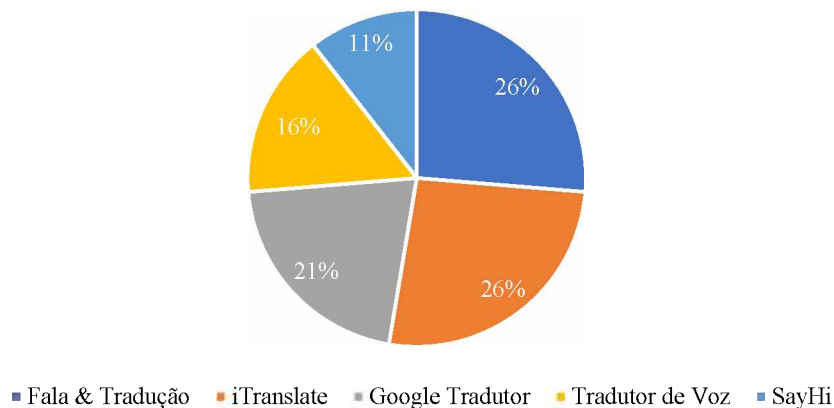
Fonte: Elaborado pelo autor.

Em comentários como “vou usar um aplicativo comum como o Google Tradutor, que é mil vezes melhor que esse” e “o Google Tradutor é bem melhor e é de graça!”, bastante frequentes no conjunto de dados analisado, o aplicativo Google Tradutor é visto como um aplicativo referência, pois, segundo os usuários, além de ser gratuito, oferece o maior número de pares linguísticos, tradução de texto em imagens, definições das palavras e outros recursos que os até mesmo os aplicativos pagos não oferecem.

A maioria das percepções do tipo PSOA estão vinculadas a percepções negativas sobre os aplicativos que os usuários estão utilizando quando estão compartilhando seus comentários. Essas percepções também estão vinculadas às percepções sobre o custo-

benefício dos aplicativos. O aplicativo que mais recebeu comentários negativos a respeito do seu desempenho em comparação a outros aplicativos foi o Fala & Tradução e o que menos recebeu comentários desse tipo foi o SayHi, conforme mostra a Figura 35.

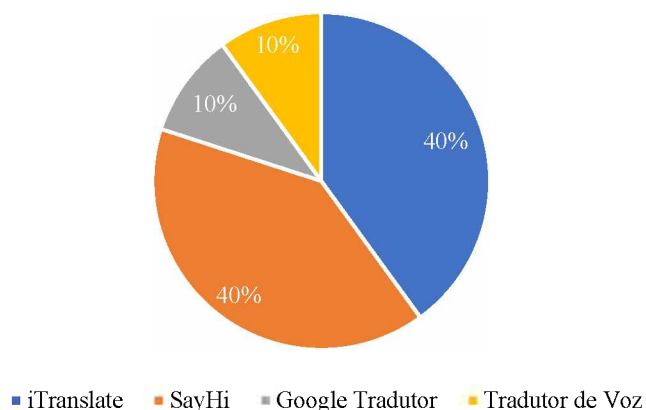
Figura 35 – Distribuição das percepções negativas sobre outros aplicativos por aplicativo



Fonte: Elaborado pelo autor.

As percepções positivas relacionadas à categoria PSOA estão vinculadas aos pontos fracos dos aplicativos que os usuários utilizaram em algum momento. O aplicativo que mais recebeu comentários positivos a respeito do seu desempenho em relação a outros aplicativos foi o iTranslate e o que menos recebeu comentários desse tipo foi o Tradutor de Voz, conforme mostra a Figura 36.

Figura 36 – Distribuição das percepções positivas sobre outros aplicativos por aplicativo



Fonte: Elaborado pelo autor.

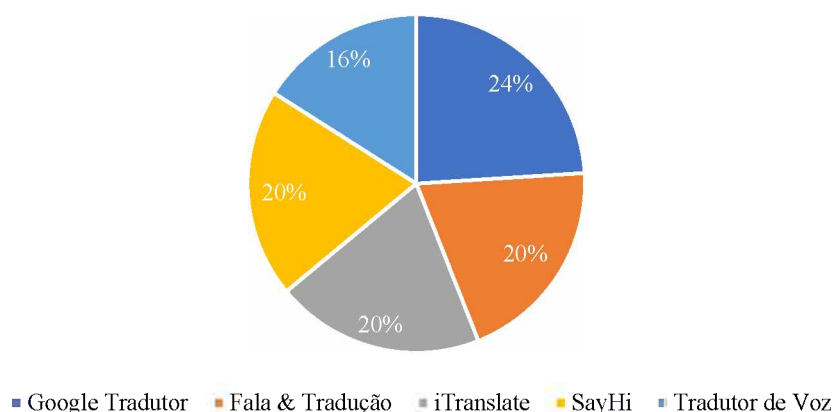
De acordo com as percepções sobre outros aplicativos, os usuários esperam que a mesma quantidade de serviços e recursos seja oferecida pelo mesmo preço e qualidade. Assim como nas percepções do tipo PSPA, quando o Google Tradutor é citado como referência, os usuários esperam que os sejam totalmente gratuitos e eficientes.

6.8 Percepções sobre o módulo de TTS

O módulo TTS corresponde à etapa final do PLN realizado pelos sistemas de interpretação automática. Na etapa de síntese, os trechos traduzidos provenientes do módulo de TA são convertidos para sinal acústico. As percepções sobre o módulo de TTS (PSMTTS) dizem respeito à qualidade da voz sintetizada reproduzida pelos cinco aplicativos analisados, à velocidade de reprodução da voz e, em alguns aplicativos, a impossibilidade de seleccionar o gênero masculino ou feminino.

As percepções do tipo PSMTTS correspondem a 7% das percepções identificadas no conjunto de comentários analisado. O aplicativo Google Tradutor é o aplicativo com o maior número de percepções do tipo PSMTTS e o Tradutor de Voz é o aplicativo com o menor número de percepções desse tipo, conforme mostra a Figura 37.

Figura 37 - Distribuição da categoria PSMTTS por aplicativo



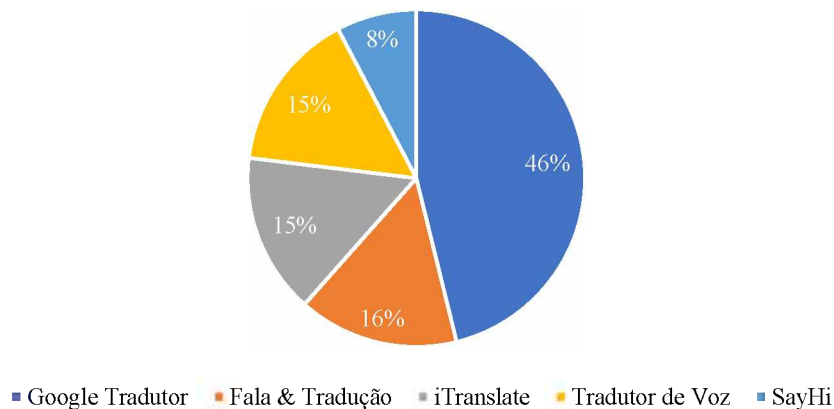
Fonte: Elaborado pelo autor.

Nos comentários compartilhados, os usuários relatam as dificuldades de o módulo TTS reproduzir os aspectos prosódicos do trecho de fala original. Elementos como pronúncia de má qualidade, entonação incompatível com a entonação original e ausência de sotaques naturais são observadas pelos usuários.

A maioria das percepções do tipo PSMTTS estão vinculadas a percepções negativas sobre a qualidade da voz sintetizada, descrita pelos usuários como “voz robótica”. Há também alguns comentários que veiculam percepções sobre a lentidão da voz sintetizada e a ausência de vozes sintéticas do gênero masculino.

O aplicativo que mais recebeu comentários negativos referentes ao módulo TTS foi o Google Tradutor e o que menos recebeu comentários desse tipo foi o Tradutor de Voz, conforme mostra a Figura 38.

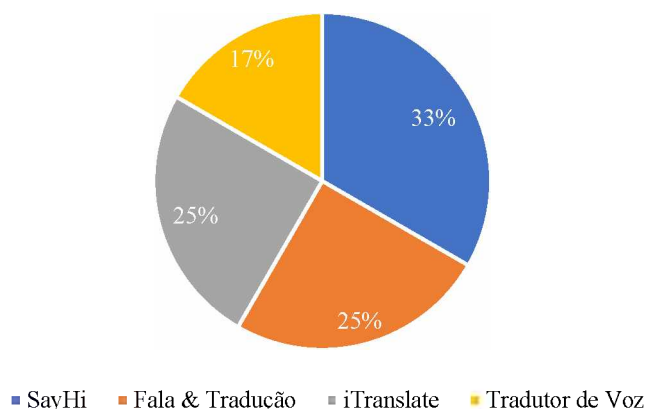
Figura 38 – Distribuição das percepções negativas sobre o módulo TTS por aplicativo



Fonte: Elaborado pelo autor.

As percepções positivas sobre o módulo TTS referem-se à naturalidade com que a voz sintética soa aos ouvidos dos usuários e a possibilidade de escolher a velocidade de reprodução dos trechos sintetizados e a escolha de gênero. O aplicativo que mais recebeu comentários positivos referentes à categoria PSMTSS foi o SayHi e o que menos recebeu comentários desse tipo foi o Tradutor de Voz, conforme mostra a Figura 39. O aplicativo Google Tradutor não recebeu nenhuma percepção positiva desse tipo.

Figura 39 – Distribuição das percepções positivas sobre o módulo TTS por aplicativo



Fonte: Elaborado pelo autor.

O conjunto de percepções do tipo PSMTTS releva que os usuários dos aplicativos de interpretação automática esperam que a voz sintetizada pelo módulo TTS seja

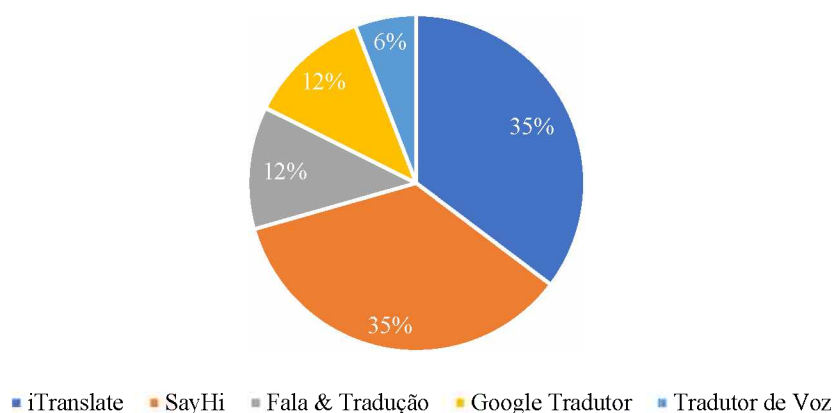
agradável aos ouvidos e soe com naturalidade semelhante à voz humana. Os usuários esperam que os aplicativos ofereçam a possibilidade de personalizar o gênero da voz sintetizada, assim como a velocidade de reprodução da voz. Além disso, os usuários esperam que as vozes sintetizadas sejam capazes de reproduzir os aspectos prosódicos presentes na voz do interlocutor que emitiu a fala original.

6.9 Percepções sobre a interface

A interface de um aplicativo é a área em que o usuário interage com o *software* a fim de usufruir dos recursos disponibilizados pelo aplicativo. A interação se dá por meio de comandos, ícones, menus, caixas de diálogos, sistema de ajuda e *feedback* etc. Os recursos disponibilizados pelos cinco aplicativos analisados são acionados predominantemente por meio de ícones, os elementos textuais são apresentados como o produto das traduções e transcrições realizadas.

Nos comentários analisados, os usuários referem-se à interface por meio de expressões como “aplicativo fácil/ simples de usar” “aplicativo prático”, “interface intuitiva/ bem elaborada” e “interface simples de manusear/ fácil de operar”. Além disso, a interface também é chamada de “tela”, “layout” e “visual”. As percepções sobre a interface (PSI) correspondem a 4% das percepções identificadas no conjunto de comentários analisado. O iTranslate é o aplicativo com o maior número de percepções do tipo PSI e o Tradutor de Voz é o aplicativo com o menor número de percepções desse tipo, conforme mostra a Figura 40.

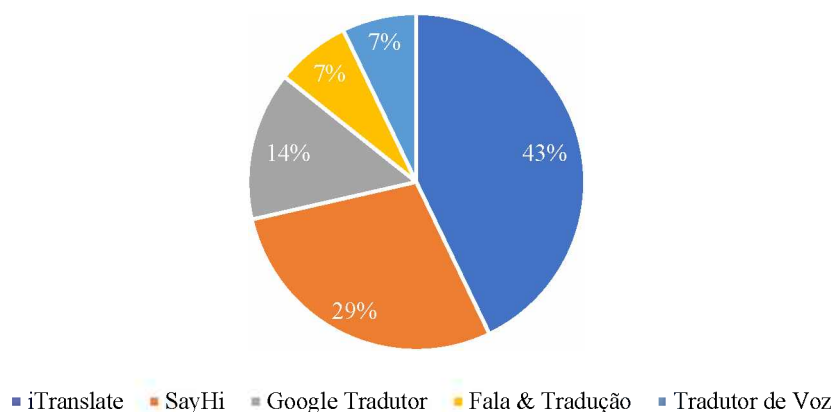
Figura 40 - Distribuição da categoria PSI por aplicativo



Fonte: Elaborado pelo autor.

A maioria das percepções sobre a interface estão vinculadas a percepções positivas sobre a facilidade de se utilizar os recursos oferecidos pelos aplicativos. Das 10 categorias identificadas, apenas a PSI está vinculada a mais percepções positivas do que a percepções negativas. O aplicativo que mais recebeu comentários positivos a respeito de sua interface foi o iTranslate e o que menos recebeu comentários desse tipo foi o Tradutor de Voz, conforme mostra a Figura 41.

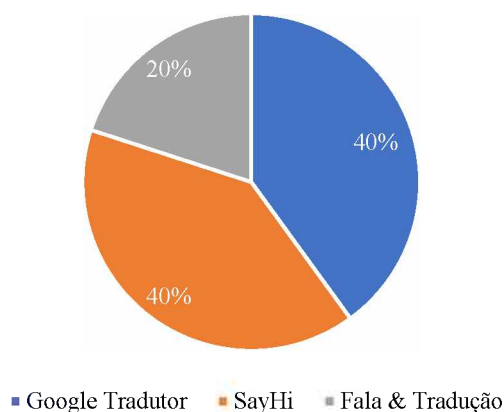
Figura 41 – Distribuição das percepções positivas sobre a interface por aplicativo



Fonte: Elaborado pelo autor.

As percepções negativas atribuídas à categoria PSI estão vinculadas a comportamentos inesperados tais como tela branca, desaparecimento e mal funcionamento de ícones. O aplicativo que mais recebeu comentários negativos a respeito de sua interface foi o Google Tradutor e o que menos recebeu comentários desse tipo foi o Fala & Tradução, conforme mostra a Figura 42. No conjunto de comentários analisado, nenhuma percepção negativa foi atribuída às interfaces dos aplicativos iTranslate e Tradutor de Voz.

Figura 42 – Distribuição de percepções negativas sobre a interface por aplicativo



Fonte: Elaborado pelo autor.

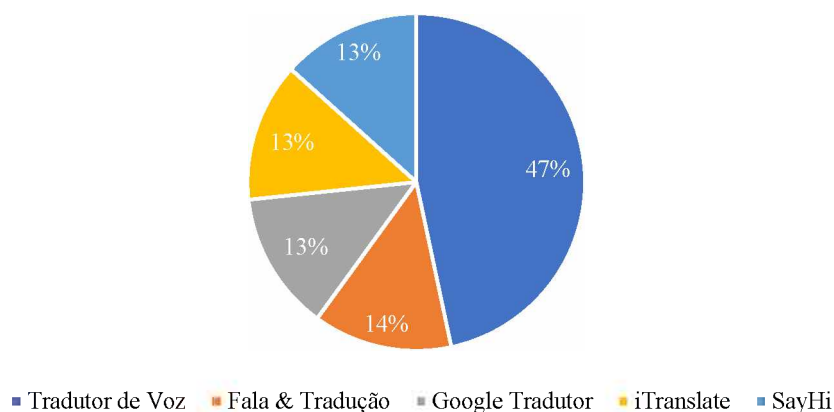
As percepções sobre a interface revelam que os usuários esperam que os aplicativos de interpretação automática possuam interfaces intuitivas, com os ícones e elementos visuais bem organizados de forma a facilitar a experiência de utilização.

6.10 Percepções sobre o modo *offline*

O modo *offline* permite que os usuários utilizem os recursos dos aplicativos sem a necessidade de uma conexão de dados. Nos cinco aplicativos, o modo *offline* requer a disponibilidade de armazenamento de dados no dispositivo móvel e a quantidade de espaço necessária depende do número de línguas que se deseja armazenar. Nenhuma das línguas requer um espaço de armazenamento superior a um gigabyte.

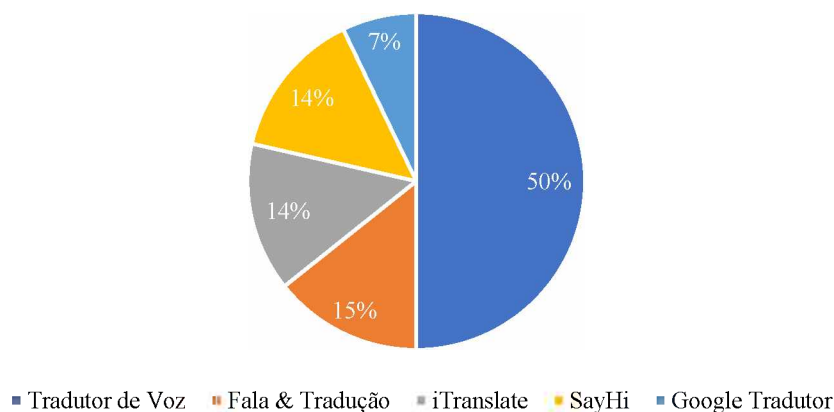
As percepções sobre o modo *offline* (PSMO) correspondem a 4% das percepções identificadas no conjunto de comentários analisado. O Tradutor de Voz é o aplicativo com o maior número de percepções do tipo PSMO e o SayHi é o aplicativo com o menor número de percepções desse tipo, como mostra a Figura 43.

Figura 43 - Distribuição da categoria PSMO por aplicativo



Fonte: Elaborado pelo autor.

Todas as percepções do tipo PSMO estão vinculadas a percepções negativas sobre a indisponibilidade ou não funcionamento do modo *offline*. Quando comparado aos demais recursos e comentários analisados, percebe-se que o modo *offline* em determinados aplicativos é um dos principais elementos que gera frustração nos usuários. O aplicativo que mais recebeu comentários negativos a respeito do modo *offline* foi o Tradutor de Voz e o que menos recebeu comentários desse tipo foi o Google Tradutor, conforme mostra a Figura 44.

Figura 44 – Distribuição das percepções negativas sobre o modo *offline* por aplicativo

Fonte: Elaborado pelo autor.

A análise das percepções sobre o modo *offline* revela que os usuários esperam que os aplicativos de interpretação automática sejam capazes de realizar traduções e interpretações automáticas até mesmo quando não há conexões de dados estáveis e de qualidade. Além disso, os usuários esperam que um grande número de línguas e pares

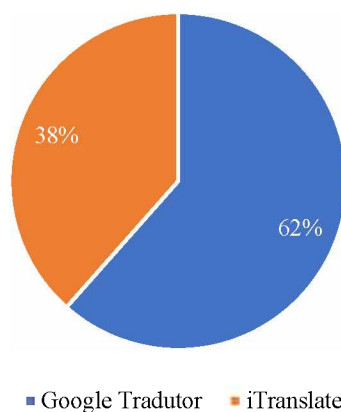
linguísticos possam ser disponibilizados no modo *offline* sem que sejam necessárias grandes quantidades de espaço na memória dos dispositivos.

6.11 Percepções sobre a tradução de texto em imagens

A tradução de texto em imagens é um recurso que possibilita aos usuários traduzir textos sem ter que digitá-los na interface do aplicativo. O recurso faz uso da câmera do dispositivo móvel e não exige que o usuário capture a imagem, bastando apontar a câmera para o texto que se deseja traduzir e manter o dispositivo imóvel por alguns instantes. Nos comentários analisados, os usuários referem-se à tradução de texto em imagens por meio das expressões “câmera de tradução”, “conversor através de fotos”, “tradução de texto por foto” e “tradução por câmera”.

As percepções sobre a tradução de texto em imagens (PSTTI) correspondem a 3% das percepções identificadas no conjunto de comentários analisado. Quando os dados desta análise foram coletados, a tradução de texto em imagens era divulgada nas *webpages* do Google Tradutor, iTranslate e Tradutor de Voz, mas somente os dois primeiros aplicativos ofereciam o recurso em suas interfaces. A maioria das percepções desse tipo foram atribuídas ao Google Tradutor, como mostra a Figura 45.

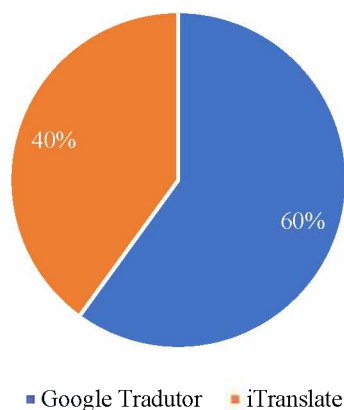
Figura 45 - Distribuição da categoria PSTTI por aplicativo



Fonte: Elaborado pelo autor.

A maioria das percepções do tipo PSTTI estão vinculadas a percepções negativas sobre inconsistências nas traduções e não funcionamento do recurso de tradução de texto em imagens. O Google Tradutor recebeu 60% das percepções negativas vinculadas às percepções do tipo PSTTI e o iTranslate 40%, conforme mostra a Figura 46.

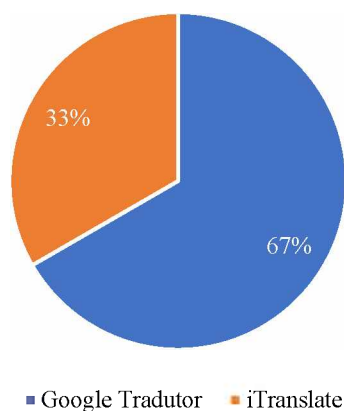
Figura 46 – Distribuição das percepções negativas sobre a tradução de texto em imagens por aplicativo



Fonte: Elaborado pelo autor.

As percepções positivas sobre a tradução de texto em imagens estão vinculadas ao fato de o recurso, de fato, ser capaz de reconhecer textos em imagens e não à qualidade das traduções. O aplicativo que mais recebeu comentários positivos relacionados à tradução de texto em imagens foi o Google Tradutor, conforme mostra a Figura 47.

Figura 47 – Distribuição das percepções positivas sobre a tradução de texto em imagens por aplicativo



Fonte: Elaborado pelo autor.

A análise das percepções do tipo PSTTI sugere que os usuários esperam que o serviço de tradução de texto em imagens produza traduções de qualidade independentemente da qualidade da imagem a ser processada ou das configurações da câmera. Além disso, os usuários esperam que este serviço esteja disponível em todos os aplicativos pagos, pois o Google Tradutor oferece esse serviço sem nenhum tipo de cobrança.

Considerações Finais

O objetivo geral do presente estudo foi identificar as percepções dos usuários brasileiros de serviços de IA oferecidos por cinco aplicativos móveis da App Store. Para atingir tal objetivo, defini os critérios para a seleção dos cinco aplicativos móveis dos quais foram extraídos os comentários dos usuários; tratei das principais características dos serviços oferecidos pelos cinco aplicativos; coletei os comentários fidedignos compartilhados pública e anonimamente por usuários brasileiros desse tipo de tecnologia; analisei os comentários a fim de identificar as percepções neles veiculadas; e classifiquei as percepções a fim de tratar brevemente das expectativas dos usuários.

Os resultados corroboram as duas hipóteses levantadas. De fato, vários comentários veiculam expectativas e percepções contraditórias e, por mais que o usuário as separe em períodos, o fato de essas expectativas e percepções comporem um mesmo comentário pode levar o observador a ter dificuldades com a interpretação delas. Com efeito, as expectativas e percepções dos usuários apontam para a necessidade de os aplicativos melhorarem os serviços de IA por eles oferecidos, principalmente no que se refere à tradução de textos em imagens oferecida pelo Google Tradutor e a funcionalidade de tradução *offline* oferecida pelo aplicativo Tradutor de Voz.

Ao compararem a maioria dos aplicativos ao Google Tradutor, os usuários dão a entender que desconhecem o fato de que o módulo de tradução automática no cerne desses aplicativos muitas vezes foi desenvolvido com os mesmos dados utilizados para o desenvolvimento do Google Tradutor. Além disso, os usuários dos cinco aplicativos parecem exigir perfeição tanto dos humanos quanto das máquinas, sem se sensibilizarem de que o homem, a ciência e a tecnologia convergem e têm suas limitações próprias.

Por fim, os dados do presente estudo mostram que as expectativas e percepções dos usuários merecem atenção por parte tanto dos desenvolvedores de aplicativos de IA quanto dos intérpretes humanos, pois os parâmetros que devem ser levados em consideração no julgamento da qualidade de traduções e interpretações automáticas nem sempre estão ao alcance daquele usuário que nunca consumiu traduções e interpretações.

Referências

- AMANO, Shin-ya; NOGAMI, Hiroyasu; MIIKE, Seiji. A step towards Telecommunication with Machine Interpreter. In: **Second International Conference on Theoretical and Methodological Issues in Machine Translation of Natural Languages**. 1988. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.521.5536&rep=rep1&type=pdf>>. Acesso em: 11 out. 2021.
- APPLE. Ratings, Reviews, and Responses. **Apple**, 2020. Disponível em: <<https://developer.apple.com/app-store/ratings-and-reviews/>>. Acesso em: 20 out. 2021.
- ARROJO, Rosemary. Os “estudos da tradução” como área de pesquisa independente: dilemas e ilusões de uma disciplina em (des)construção. **DELTA: Documentação de Estudos em Linguística Teórica e Aplicada**, v. 14, p. 423-454, 1998. Disponível em: <<https://www.scielo.br/j/delta/a/qfsD3KLYkcsCcmzzfBmkmsM/?lang=pt&format=html>>. Acesso em: 31 out. 2021. <https://doi.org/10.1590/S0102-44501998000200007>.
- AVAZU. **Global Internet Industry Research - Brazil**. 2003. Disponível em: <[http://avazuinc.com/wordpress/wp-content/download/en/Global Internet Industry Research - Brazil Part II.pdf](http://avazuinc.com/wordpress/wp-content/download/en/Global%20Internet%20Industry%20Research%20-%20Brazil%20Part%20II.pdf)>.
- BAHL, L. *et al.* Further results on the recognition of a continuously read natural corpus. In: **ICASSP'80. IEEE International Conference on Acoustics, Speech, and Signal Processing**. IEEE, 1980. p. 872-875. Disponível em: <<https://ieeexplore.ieee.org/abstract/document/1170862/>>. Acesso em: 11 out. 2021. <https://doi.org/10.1109/ICASSP.1980.1170862>.
- BAIGORRI-JALÓN, Jesús. Conference interpreting: From modern times to space technology. **Interpreting**, v. 4, n. 1, p. 29-40, 1999. Disponível em: <<https://www.jbe-platform.com/content/journals/10.1075/intp.4.1.05bai>>. Acesso em: 11 out. 2021. <https://doi.org/10.1075/intp.4.1.05bai>.
- BERBER-IRABIEN, Diana *et al.* **Information and communication technologies in conference interpreting**. Universitat Rovira i Virgili, 2010. Disponível em: <<https://www.tdx.cat/handle/10803/8775#page=1>>. Acesso em: 11 out. 2021.
- BRAUN, Sabine. Technology and interpreting. In: **The Routledge Handbook of translation and technology**. Routledge, 2019. p. 271-288. Disponível em: <<https://www.taylorfrancis.com/chapters/edit/10.4324/9781315311258-16/technology-interpreting-sabine-braun>>. Acesso em: 11 out. 2021. <https://doi.org/10.4324/9781315311258-16>.
- BREINLINGER, J.; HAGIN, A.; WRIGHT, J. The Problems with 5-Star Rating Systems, and How to Fix Them. **Harvard Business Review**, 2019. Disponível em: <<https://hbr.org/2019/07/the-problems-with-5-star-rating-systems-and-how-to-fix-them>>. Acesso em: 11 out. 2021.
- BÜHLER, Hildegund. Linguistic (semantic) and extra-linguistic (pragmatic) criteria for the evaluation of conference interpretation and interpreters. **Multilingua**, v. 5, n. 4, p. 231-235, 1986. Disponível em: <<https://www.degruyter.com/document/doi/10.1515/mult.1986.5.4.231/html>>. Acesso em: 11 out. 2021. <https://doi.org/10.1515/mult.1986.5.4.231>.
- CHESTERMAN, Andrew. Bridge concepts in translation sociology. **Benjamins translation library**, v. 74, p. 171, 2007. Disponível em: <<https://books.google.com.br/books?hl=pt-BR&lr=&id=wXg6AAAAQBAJ&oi=fnd&pg=PA171&dq=Bridge+concepts+in+translation+so>>.

ciology&ots=K7wY-G8qtj&sig=OgJa9YkYcX7FbKwLv9e5oym2eIY>. Acesso em: 11 out. 2021. <https://doi.org/10.1075/btl.74>.

CHEUNG, Andrew KF. Non-native accents and simultaneous interpreting quality perceptions. **Interpreting**, v. 15, n. 1, p. 25-47, 2013. Disponível em: <<https://www.jbe-platform.com/content/journals/10.1075/intp.15.1.02che>>. Acesso em: 11 out. 2021. <https://doi.org/10.1075/intp.15.1.02che>.

COLLADOS-AÍS, Á.; GARCÍA-BECERRA, O. Quality Criteria. In: POCHHACKER, Franz (Ed.). **Routledge encyclopedia of interpreting studies**. Routledge, 2015. Disponível em: <<https://books.google.com.br/books?hl=pt-BR&lr=&id=01JyDAAAQBAJ&oi=fnd&pg=PP1&dq=Quality+Criteria+Routledge+encyclopedia+of+interpreting+studies&ots=awjZLwH9cp&sig=F5Pe6BbFHLg3i0ZYBWY9CHB-Y8A>>. Acesso em: 11 out. 2021. <https://doi.org/10.4324/9781315678467>.

COLLADOS-AÍS, Á. **La evaluación de la calidad en interpretación simultánea: La importancia de la comunicación no verbal**. Granada: Comares, 1998.

COLLADOS AÍS, Á. *et al.* **La Evaluación de la Calidad en Interpretación Simultánea: Parámetros de Incidencia**. Granada: Comares, 2007. Disponível em: <https://www.researchgate.net/profile/Olalla-Garcia-Becerra/publication/285600238_La_evaluacion_de_la_calidad_en_interpretacion_simultanea_parametros_de_incidencia/links/5809f1f608ae45e02c0d60af/La-evaluacion-de-la-calidad-en-interpretacion-simultanea-parametros-de-incidencia.pdf>. Acesso em: 11 out. 2021.

CORTELLESA, V.; MARCO, A. DI; INVERARDI, P. What Is Software Performance? *In: Model-Based Software Performance Analysis*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011. p. 1-7. Disponível em: <https://link.springer.com/chapter/10.1007/978-3-642-13621-4_1>. Acesso em: 11 out. 2021. https://doi.org/10.1007/978-3-642-13621-4_1.

COSTELLO, S. How Many Apps are in the App Store? **Lifewire**, 2020. Disponível em: <<https://www.lifewire.com/how-many-apps-in-app-store-2000252>>. Acesso em: 11 out. 2021.

DOOLEY, John; ZUKOWSKI, John. **Software development and professional practice**. Berkeley: Apress, 2011. Disponível em: <<https://link.springer.com/book/10.1007/978-1-4302-3802-7>>. Acesso em: 11 out. 2021. <https://doi.org/10.1007/978-1-4302-3802-7>.

VAN DOORN, Jenny *et al.* Customer engagement behavior: Theoretical foundations and research directions. **Journal of service research**, v. 13, n. 3, p. 253-266, 2010. Disponível em: <<https://journals.sagepub.com/doi/abs/10.1177/1094670510375599>>. Acesso em: 11 out. 2021. <https://doi.org/10.1177/1094670510375599>.

DUNLEAVY, Patrick. **Authoring a PhD: How to plan, draft, write and finish a doctoral thesis or dissertation**. Macmillan International Higher Education, 2003. Disponível em: <https://books.google.com.br/books?hl=pt-BR&lr=&id=_gTcCwAAQBAJ&oi=fnd&pg=PP1&dq=Authoring+a+PhD:+How+to+plan,+draft,+write+and+finish+a+doctoral+thesis+or+dissertation&ots=XuI3CB062N&sig=wm1-RQXdTvczSK5Y0ctsKGbx8U0>. Acesso em: 11 out. 2021.

ECK, Matthias *et al.* Jibbig: Speech-to-speech translation on mobile devices. In: **2010 IEEE Spoken Language Technology Workshop**. IEEE, 2010. p. 165-166. Disponível em: <<https://ieeexplore.ieee.org/abstract/document/5700843/>>. Acesso em: 11 out. 2021. <https://doi.org/10.1109/SLT.2010.5700843>.

ENGLANDER, Karen. The rise of English as the language of science. In: **Writing and Publishing Science Research Papers in English**. Springer, Dordrecht, 2014. p. 3-4. Disponível em: <https://link.springer.com/chapter/10.1007/978-94-007-7714-9_1>. Acesso em: 11 out. 2021. https://doi.org/10.1007/978-94-007-7714-9_1

ESQUEDA, Marileide Dias; FREITAS, Flávio de Sousa. A intenção discursiva nos sistemas de interpretação automática. **Domínios de Linguagem**, v. 13, n. 2, p. 511-551, 2019. Disponível em: <<http://www.seer.ufu.br/index.php/dominiosdelinguagem/article/view/45057>>. Acesso em: 11 out. 2021. <https://doi.org/10.14393/DL38-v13n2a2019-4>.

EXPECTATIVA. In: MICHAELIS, Dicionário Brasileiro da Língua Portuguesa. São Paulo: Editora Melhoramentos Ltda, 2015. Disponível em: <<https://michaelis.uol.com.br/moderno-portugues/busca/portugues-brasileiro/expectativa/>>. Acesso em: 20 out. 2021. (Entrada em dicionário).

FERREIRA JUNIOR, Fernando G. Uma interlíngua conexcionista. **Trabalhos em Linguística Aplicada**, v. 46, p. 219-231, 2007. Disponível em: <<https://www.scielo.br/j/tla/a/6k8hy5sxxRYS7xvqYTfjkZw/abstract/?lang=pt&format=html>>. Acesso em: 31 out. 2021. <https://doi.org/10.1590/S0103-18132007000200006>.

FLEISCHMANN, Marvin *et al.* The role of software updates in information systems continuance—An experimental study from a user perspective. **Decision Support Systems**, v. 83, p. 83-96, 2016. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0167923616000026>>. Acesso em: 13 out. 2021. <https://doi.org/10.1016/j.dss.2015.12.010>.

FREITAS, F. DE S. **O Estado da arte da interpretação automática: do pós-guerra aos apps de tradução automática de fala**. Uberlândia, Universidade Federal de Uberlândia, 2016. Monografia.

FREITAS, F.; ESQUEDA, M. Interpretação automática ou tradução automática de fala: conceitos, definições e arquitetura de software. **Tradterm**, v. 29, p. 104-145, 13 jul. 2017. Disponível em: <<https://www.revistas.usp.br/tradterm/article/view/134416>>. Acesso em: 13 out. 2021. <https://doi.org/10.11606/issn.2317-9511.v29i0p104-145>.

FREITAS, F. S.; ESQUEDA, M. D. **Tradução e Interpretação Automáticas: Origens**. Curitiba: Editora CRV, 2020. <https://doi.org/10.24824/978854443937.1>.

GARCEZ, Eliane Maria Stuart; RADOS, Gregório J. Varvakis. Necessidades e expectativas dos usuários na educação a distância: estudo preliminar junto ao Programa de Pós-Graduação em Engenharia de Produção da Universidade Federal de Santa Catarina. **Ciência da Informação**, v. 31, p. 13-26, 2002. Disponível em: <<https://www.scielo.br/j/ci/a/XDXFmWQpBQTctdHhmDLzydH/abstract/?lang=pt>>. Acesso em: 13 out. 2021. <https://doi.org/10.1590/S0100-19652002000100003>.

GEHRKE, Manfred; SCHMIDBAUER, Otto. German-Japanese Speech Translation in CSTAR. In: **Grundlagen und Anwendungen der Künstlichen Intelligenz**. Springer, Berlin, Heidelberg, 1993. p. 338-349. Disponível em: <https://link.springer.com/chapter/10.1007/978-3-642-78545-0_32>. Acesso em: 13 out. 2021. https://doi.org/10.1007/978-3-642-78545-0_32.

GILE, Daniel; ESQUEDA, Marileide Dias; FREITAS, Flávio de Sousa. Da comunicação à qualidade em interpretação e tradução. **Letras & Letras**, v. 35, n. 2, p. 219-245, 2019. Disponível em: <<http://www.seer.ufu.br/index.php/letraseletras/article/view/51596>>. Acesso em: 13 out. 2021. <https://doi.org/10.14393/LL63-v35n2-2019-14>.

GLOBE NEWSWIRE. **The speech to speech translation market is expected to record a growth at a CAGR of 9.4%, over the forecast period (2020 - 2025)**. Disponível em: <<https://www.globenewswire.com/news-release/2020/04/09/2014564/0/en/The-speech-to-speech-translation-market-is-expected-to-record-a-growth-at-a-CAGR-of-9-4-over-the-forecast-period-2020-2025.html>>. Acesso em: 15 out. 2021.

GORLA, Alessandra *et al.* Checking app behavior against app descriptions. In: **Proceedings of the 36th international conference on software engineering**, 2014. p. 1025-1035. Disponível em: <<https://dl.acm.org/doi/abs/10.1145/2568225.2568276>>. Acesso em: 13 out. 2021. <https://doi.org/10.1145/2568225.2568276>.

GOEL, R. Are you asking for customer feedback? **UX Collective**, 2018. Disponível em: <<https://uxdesign.cc/are-you-asking-for-customer-feedback-3bd1597ce5d8>>. Acesso em: 13 out. 2021.

GRBIĆ, N. Quality. In: POCHHACKER, Franz (Ed.). **Routledge encyclopedia of interpreting studies**. Routledge, 2015. Disponível em: <<https://books.google.com.br/books?hl=pt-BR&lr=&id=01JyDAAAQBAJ&oi=fnd&pg=PP1&dq=Quality+Criteria+Routledge+encyclopedia+of+interpreting+studies&ots=awjZLwH9cp&sig=F5Pe6BbFHLg3i0ZYBWY9CHB-Y8A>>. Acesso em: 11 out. 2021. <https://doi.org/10.4324/9781315678467>.

GRÖNROOS, C. **Marketing: gerenciamento e serviços: a competição por serviços na hora da verdade**. Rio de Janeiro: Campus, 1993.

HAHN, Walther v; AMTRUP, Jan W. Speech-to-speech translation: The project verbmobil. In: **Proceedings of SPECOM**, 1996. p. 51-56. Disponível em: <<https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.49.9403&rep=rep1&type=pdf>>. Acesso em: 13 out. 2021.

HALE, S. **Community interpreting**. Springer, 2007. Disponível em: <<https://books.google.com.br/books?hl=pt-BR&lr=&id=O5CDDAAAQBAJ&oi=fnd&pg=PP1&dq=Community+interpreting&ots=vTetPsumg2&sig=T25D9rjsetN145UDASX6LtANgtA>>. Acesso em: 13 out. 2021. <https://doi.org/10.1057/9780230593442>.

IKEDA, Yoshio; ISHII, Yohei. Recognition of two psychological conditions of a single cow by her voice. **Computers and Electronics in Agriculture**, v. 62, n. 1, p. 67-72, 2008. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0168169907001901>>. Acesso em: 13 out. 2021. <https://doi.org/10.1016/j.compag.2007.08.012>.

JOHNSTON, R., CLARCK, G. **Administração de Operações de Serviços**. São Paulo: Atlas, 2002. 562p.

KITANO, Hiroaki. Phi DM-Dialog: an experimental speech-to-speech dialog translation system. **Computer**, v. 24, n. 6, p. 36-50, 1991. Disponível em: <<https://ieeexplore.ieee.org/abstract/document/86837/>>. Acesso em: 13 out. 2021. <https://doi.org/10.1109/2.86837>.

KITANO, Hiroaki. The ΦDMDIALOG System. In: **Speech-to-Speech Translation**. Springer, Boston, MA, 1994. p. 47-113. Disponível em: <https://link.springer.com/chapter/10.1007/978-1-4615-2732-9_4>. Acesso em: 13 out. 2021. https://doi.org/10.1007/978-1-4615-2732-9_4.

KONDO, Masaomi *et al.* Intercultural communication, negotiation and interpreting. **Benjamins Translation Library**, v. 23, p. 149-166, 1997. Disponível em: <<https://books.google.com.br/books?hl=pt->

BR&lr=&id=zQVy80kxjMC&oi=fnd&pg=PA149&dq=Intercultural+communication,+negotiation+and+interpreting&ots=RViT3BDw0w&sig=QqICf6ptClkaTA3rUxWvNQAj9x4>. Acesso em: 13 out. 2021. <https://doi.org/10.1075/btl.23.10kon>.

KUREMATSU, Akira *et al.* Language processing in connection with speech translation at ATR Interpreting Telephony Research Laboratories. **Speech communication**, v. 10, n. 1, p. 1-9, 1991. Disponível em: <<https://www.sciencedirect.com/science/article/pii/016763939190023M>>. Acesso em: 13 out. 2021. [https://doi.org/10.1016/0167-6393\(91\)90023-M](https://doi.org/10.1016/0167-6393(91)90023-M).

KURZ, Ingrid. Conference interpreting: Quality in the ears of the user. **Meta: journal des traducteurs/Meta: Translators' Journal**, v. 46, n. 2, p. 394-409, 2001. Disponível em: <<https://www.erudit.org/en/journals/meta/1900-v1-n1-meta159/003364ar/abstract/>>. Acesso em: 15 out. 2021. <https://doi.org/10.7202/003364ar>.

KURZ, I. Conference interpretation: Expectations of different user groups. **The Interpreters' Newsletter**, v. 5, p. 13-21, 1993. Disponível em: <<https://www.openstarts.units.it/bitstream/10077/4908/1/KurzIN5.pdf>>. Acesso em: 15 out. 2021.

KURZ, I; PÖCHHACKER, F. Quality in TV interpreting. **Translatio-Nouvelles de la FIT-FIT Newsletter**, v. 15, n. 3, p. 4, 1995.

LEE, T. Speech Translation. In: CHAN, Sin-wai (Ed.). **Routledge encyclopedia of translation technology**. Routledge, 2015. Disponível em: <https://books.google.com.br/books?hl=pt-BR&lr=&id=S0FWBQAAQBAJ&oi=fnd&pg=PP1&dq=Routledge+Encyclopedia+of+Translation+Technology&ots=fmyx0_FHbP&sig=A-cewnVK2fE_541-VVohM7YojVs>. Acesso em: 15 out. 2021. <https://doi.org/10.4324/9781315749129>.

LEVIN, Lori S. *et al.* An interlingua based on domain actions for machine translation of task-oriented dialogues. In: **ICSLP**. 1998. p. 1155-1158. Disponível em: <<https://isl.anthropomatik.kit.edu/pdf/Levin1998.pdf>>. Acesso em: 15 out. 2021.

MASANO, A. C. R. **Expectativas e percepções do mercado bancário de pessoas físicas de alta renda no município de São Paulo**. 2006. Tese de Doutorado. Universidade de São Paulo. Disponível em: <<https://www.teses.usp.br/teses/disponiveis/12/12139/tde-22052007-144154/en.php>>. Acesso em: 15 out. 2021. <https://doi.org/10.11606/D.12.2007.tde-22052007-144154>.

MATSUDA, Shigeki *et al.* Multilingual speech-to-speech translation system: VoiceTra. In: **2013 IEEE 14th International Conference on Mobile Data Management**. IEEE, 2013. p. 229-233. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6569095>>. Acesso em: 15 out. 2021. <https://doi.org/10.1109/MDM.2013.99>.

MAIER-HEIN, Lena *et al.* Session independent non-audible speech recognition using surface electromyography. In: **IEEE Workshop on Automatic Speech Recognition and Understanding**. IEEE, 2005. p. 331-336. Disponível em: <<https://ieeexplore.ieee.org/abstract/document/1566521/>>. Acesso em: 15 out. 2021. <https://doi.org/10.1109/ASRU.2005.1566521>.

MOCH, S.; GATES, M. **The Researcher Experience in Qualitative Research**. California: SAGE Publications, Inc., 2000. Disponível em: <https://books.google.com.br/books?hl=pt-BR&lr=&id=yi85DQAAQBAJ&oi=fnd&pg=PP1&dq=The+Researcher+Experience+in+Qualitative+Research&ots=pU06feo2L3&sig=deItwVWsvq-5k22vPAj_-sCVAVs>. Acesso em: 15 out. 2021. <https://dx.doi.org/10.4135/9781483348995>.

MORIMOTO, Tsuyoshi *et al.* ATR's speech translation system: ASURA. In: **Third European Conference on Speech Communication and Technology**. 1993. Disponível em: <https://www.isca-speech.org/archive_v0/archive_papers/eurospeech_1993/e93_1291.pdf>. Acesso em: 15 out. 2021.

MOSER, Peter. Expectations of users of conference interpretation. **Interpreting**, v. 1, n. 2, p. 145-178, 1996. Disponível em: <<https://www.jbe-platform.com/content/journals/10.1075/intp.1.2.01mos>>. Acesso em: 15 out. 2021. <https://doi.org/10.1075/intp.1.2.01mos>.

NAKAMURA, Satoshi. **Overcoming the language barrier with speech translation technology**. NISTEP Science & Technology Foresight Center, 2009. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.472.1019&rep=rep1&type=pdf>>. Acesso em: 15 out. 2021.

NGUYEN, David T. Evaluating impact of storage on smartphone energy efficiency. In: **Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication**. 2013. p. 319-324. Disponível em: <<https://dl.acm.org/doi/abs/10.1145/2494091.2501083>>. Acesso em: 15 out. 2021. <https://doi.org/10.1145/2494091.2501083>.

OLIVER, Richard L. A cognitive model of the antecedents and consequences of satisfaction decisions. **Journal of marketing research**, v. 17, n. 4, p. 460-469, 1980. Disponível em: <<https://journals.sagepub.com/doi/abs/10.1177/002224378001700405>>. Acesso em: 15 out. 2021. <https://doi.org/10.1177/002224378001700405>.

PERCEPÇÃO. In: MICHAELIS, Dicionário Brasileiro da Língua Portuguesa. São Paulo: Editora Melhoramentos Ltda, 2015. Disponível em: <<https://michaelis.uol.com.br/moderno-portugues/busca/portugues-brasileiro/percepção/>>. Acesso em: 20 out. 2021. (Entrada em dicionário).

PÖCHHACKER, Franz. **Introducing interpreting studies**. Routledge, 2016. <https://doi.org/10.4324/9781315649573>.

PÖCHHACKER, Franz. **User Expectations**. In: PÖCHHACKER, Franz (Ed.). **Routledge encyclopedia of interpreting studies**. Routledge, 2015. Disponível em: <https://books.google.com.br/books?hl=pt-BR&lr=&id=01JyDAAAQBAJ&oi=fnd&pg=PP1&dq=Routledge+Encyclopedia+of+Interpreting+Studies&ots=awjZOzH4at&sig=-2S4xihSBhjM_GsAhdiaiumDoCo>. Acesso em: 15 out. 2021. <https://doi.org/10.4324/9781315678467>.

PRODANOV, Cleber Cristiano; DE FREITAS, Ernani Cesar. **Metodologia do trabalho científico: métodos e técnicas da pesquisa e do trabalho acadêmico-2ª Edição**. Editora Feevale, 2013.

RASMUSSEN, B. Os números não mentem: Android ou iOS, qual é o melhor? **Canaltech**, 2019. Disponível em: <<https://canaltech.com.br/produtos/os-numeros-nao-mentem-android-ou-ios-qual-e-o-melhor-7657/>>. Acesso em: 15 out. 2021.

RAWLINSON, K. **Google tradutor “turns interpreter” with voice function**. BBC News Technology. 2015. Disponível em: <<https://www.bbc.com/news/technology-30812277>>. Acesso em: 15 out. 2021.

RAYNER, Manny *et al.* Spoken language translation with mid-90's technology: A case study. In: **Third European Conference on Speech Communication and Technology**. 1993. Disponível

em: <https://www.isca-speech.org/archive_v0/archive_papers/eurospeech_1993/e93_1299.pdf>. Acesso em: 15 out. 2021.

RAYNER, Manny; BOUILLON, Pierrette; CARTER, David. Using Corpora to Develop Limited-Domain Speech Translation Systems. **Proc. Translating and the Computer**, v. 17, 1995. Disponível em: <<https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.51.7647&rep=rep1&type=pdf>>. Acesso em: 20 out. 2021.

REEVES, Nigel BR. Translating and interpreting as cultural intermediation—some theoretical issues reconsidered. **Translation and Interpreting: Bridging East and West**, p. 33-50, 1994. Disponível em: <<https://books.google.com.br/books?hl=pt-BR&lr=&id=ISAHkO3jZIEC&oi=fnd&pg=PA33&dq=Translating+and+interpreting+as+cultural+intermediation%E2%80%94some+theoretical+issues+reconsidered&ots=1G0pZPVXsI&sig=sHiT8DNf8qADiGdR3wtnoWZk860>>. Acesso em: 20 out. 2021.

RICCARDI, Alessandra. On the evolution of interpreting strategies in simultaneous interpreting. **Meta: Journal des traducteurs/Meta: Translators' Journal**, v. 50, n. 2, p. 753-767, 2005. Disponível em: <<https://www.erudit.org/en/journals/meta/1900-v1-n1-meta881/011016ar/abstract/>>. Acesso em: 20 out. 2021. <https://doi.org/10.7202/011016ar>.

RICCI, L. A Militarized PDA Voice-to-Voice Phrase Translator. In: **Handheld & Wireless Solutions Journal**, v.2, Winter 2002. p.54-55. Disponível em: <http://projectile.sv.cmu.edu/research/public/talks/speechTranslation/otherPaper/ADS-Militarized_PDA-Story.pdf>. Acesso em: 20 out. 2021.

ROJO, R. H. R. Fazer Linguística Aplicada em perspectiva sócio-histórica: privação sofrida e leveza de pensamento. In: MOITA-LOPES, Luiz Paulo da (Org.). **Por uma linguística interdisciplinar**. São Paulo: Parábola Editorial, 2006. p. 224-274.

SALDANHA, Gabriela; O'BRIEN, Sharon. **Research methodologies in translation studies**. Routledge, 2014. <https://doi.org/10.4324/9781315760100>.

SOLOMON, Délcio Vieira. **Como fazer uma monografia: elementos de metodologia do trabalho científico**. Belo Horizonte: Interlivros, 1978.

SELIGMAN, Mark; WAIBEL, Alex. 12 Advances in Speech-to-Speech Translation Technologies. **Advances in Empirical Translation Studies: Developing Translation Resources and Technologies**, p. 217, 2019. Disponível em: <https://books.google.com.br/books?hl=pt-BR&lr=&id=BIOWDwAAQBAJ&oi=fnd&pg=PA217&dq=Advances+in+Speech-to-Speech+Translation+Technologies&ots=VsUcZBP5_i&sig=Sg9Nt2CzlpwgqhOmD3dwKbMzqyQ>. Acesso em: 20 out. 2021. <https://doi.org/10.1017/9781108525695.012>.

SENDA, S. Multilingual Speech Translation. **New Breeze**, p. 9–12, abr. 2018. Disponível em: <https://www.ituaj.jp/wp-content/uploads/2018/04/nb30-2_web-04Special-MultilingualTranslation.pdf>. Acesso em: 20 out. 2021.

SOUZA, Adriano Mendonça; GRIEBELER, Deizi; GODOY, Leoni Pentiado. Qualidade na prestação de serviços fisioterápicos: estudo de caso sobre expectativas e percepções de clientes. **Production**, v. 17, p. 435-453, 2007. Disponível em: <<https://www.scielo.br/j/prod/a/FHWpYWyZNQWxH4G5zLrvWpP/abstract/?lang=pt>>. Acesso em: 20 out. 2021. <https://doi.org/10.1590/S0103-65132007000300004>.

SPERBER, Matthias; NIEHUES, Jan; WAIBEL, Alex. Toward robust neural machine translation for noisy input sequences. In: **International Workshop on Spoken Language Translation (IWSLT)**. 2017. p. 18. Disponível em: <<https://workshop2017.iwslt.org/downloads/P04-Paper.pdf>>. Acesso em: 20 out. 2021.

STENTIFORD, F. W. M.; STEER, M. G. A speech driven language translation system. In: **European Conference on Speech Technology**. 1987.

TOMITA, Masaru; TOMABECHI, Hideto; SAITO, Hiroaki. Speech Trans: An Experimental Real-Time Speech-To-Speech Translation System. **Estudo de Linguagem**, 1990. Disponível em: <<https://pdfs.semanticscholar.org/7269/a87ada4c21947ba8d098e916400abee475cb.pdf>>. Acesso em: 20 out. 2021.

TOMASSINI, Rodrigo; AQUINO, Rodrigo; CARVALHO, Frederico A.. Expectativas e percepções em seqüências de serviço. **RAE eletrônica**, v. 7, 2008. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1676-56482008000200004&lng=en&nrm=iso>. Acesso em: 20 out. 2021. <https://doi.org/10.1590/S1676-56482008000200004>

WAHLSTER, Wolfgang. **Verbmobil: Translation of face-to-face dialogs**. 1993. Relatório. Disponível em: <https://publikationen.sulb.uni-saarland.de/bitstream/20.500.11880/24979/1/RR_93_34.pdf>. Acesso em: 20 out. 2021.

WAHLSTER, Wolfgang. Mobile speech-to-speech translation of spontaneous dialogs: an overview of the final Verbmobil system. **Verbmobil: Foundations of speech-to-speech translation**, p. 3-21, 2000. Disponível em: <https://link.springer.com/chapter/10.1007/978-3-662-04230-4_1>. Acesso em: 20 out. 2021. https://doi.org/10.1007/978-3-662-04230-4_1.

WAHLSTER, Wolfgang (Ed.). **Verbmobil: foundations of speech-to-speech translation**. Springer Science & Business Media, 2013.

WAIBEL, Alex *et al.* JANUS: a speech-to-speech translation system using connectionist and symbolic processing strategies. In: **Acoustics, Speech, and Signal Processing, IEEE International Conference on**. IEEE Computer Society, 1991. p. 793-796. Disponível em: <<https://www.computer.org/csdl/proceedings-article/icassp/1991/0003793/12OmNsbGvHJ>>. Acesso em: 20 out. 2021. <https://doi.ieeeecomputersociety.org/10.1109/ICASSP.1991.150456>.

WAIBEL, Alex; FUGEN, Christian. Spoken language translation. **IEEE Signal Processing Magazine**, v. 25, n. 3, p. 70-79, 2008. Disponível em: <<https://ieeexplore.ieee.org/abstract/document/4490203/>>. Acesso em: 20 out. 2021. <https://doi.org/10.1109/ICASSP.1991.150456>.

WAIBEL, Alex. **Portable object identification and translation system**. U.S. Patent Application n. 10/090,559, 4 set. 2003. Disponível em: <<https://patents.google.com/patent/US20030164819A1/en>>. Acesso em: 20 out. 2021. <https://doi.org/10.1109/MSP.2008.918415>.

WILLIAMS, Jenny; CHESTERMAN, Andrew. **The map: a beginner's guide to doing research in translation studies**. Routledge, 2014. <https://doi.org/10.4324/9781315760513>.

YANG, Jie *et al.* **Towards automatic sign translation**. CARNEGIE-MELLON UNIV PITTSBURGH PA, 2001. Disponível em: <<https://apps.dtic.mil/sti/citations/ADA458664>>. Acesso em: 20 out. 2021.

ZHANG, Jing *et al.* A PDA-based sign translator. In: **Proceedings. Fourth IEEE International Conference on Multimodal Interfaces**. IEEE, 2002. p. 217-222. Disponível em: <<https://ieeexplore.ieee.org/abstract/document/1166996/>>. Acesso em: 20 out. 2021. <https://doi.org/10.1109/ICMI.2002.1166996>.

ZEITHAML, V. A. *et al.* **Delivering quality service: balancing customer perceptions and expectations**. New York: The Free Press, 1990. <https://doi.org/10.5860/choice.28-0390>.

ZHENG, Jing *et al.* Implementing SRI's Pashto speech-to-speech translation system on a smart phone. In: **IEEE Spoken Language Technology Workshop**. IEEE, 2010. p. 133-138. Disponível em: <<https://ieeexplore.ieee.org/abstract/document/5700835/>>. Acesso em: 20 out. 2021. <https://doi.org/10.1109/SLT.2010.5700835>.

ZHOU, Bowen *et al.* A hand-held speech-to-speech translation system. In: **IEEE Workshop on Automatic Speech Recognition and Understanding (IEEE Cat. No. 03EX721)**. IEEE, 2003. p. 664-669. Disponível em: <<http://ieeexplore.ieee.org/document/1318519/>>. Acesso em: 20 out. 2021. <https://doi.org/10.1109/ASRU.2003.1318519>.