

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO GRANDE DO SUL
FACULDADE DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

**MÁQUINAS DE TRADUÇÃO APLICADA À
COMUNICAÇÃO EM TEMPO REAL PARA
DESENVOLVIMENTO DISTRIBUÍDO DE
SOFTWARE**

TIAGO DA SILVEIRA DUARTE

Dissertação apresentada como
requisito parcial à obtenção do
grau de Mestre em Ciência da
Computação na Pontifícia
Universidade Católica do Rio
Grande do Sul.

Orientador: Prof. Dr. Rafael Prikladnicki

Porto Alegre

2014

FICHA CATALOGRÁFICA

D812m Duarte, Tiago da Silveira
Máquinas de tradução aplicada à comunicação em tempo real
para desenvolvimento distribuído de software / Tiago da Silveira
Duarte. – Porto Alegre, 2014.
117 p.

Diss. (Mestrado) – Fac. de Informática, PUCRS.
Orientador: Prof. Dr. Rafael Prikladnicki.

1. Informática. 2. Reconhecimento De Voz (Informática).
3. Tradução Automática. 4. Engenharia de Software.
I. Prikladnicki, Rafael. II. Título.

CDD 006.454

**Ficha Catalográfica elaborada pelo
Setor de Tratamento da Informação da BC-PUCRS**

TERMO DE APRESENTAÇÃO

ME 12



Pontifícia Universidade Católica do Rio Grande do Sul
FACULDADE DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

TERMO DE APRESENTAÇÃO DE DISSERTAÇÃO DE MESTRADO

Dissertação intitulada "Máquinas de Tradução Aplicada à Comunicação em Tempo Real para Desenvolvimento Distribuído de Software" apresentada por Tiago da Silveira Duarte como parte dos requisitos para obtenção do grau de Mestre em Ciência da Computação, aprovada em 18/03/2014 pela Comissão Examinadora:

Prof. Dr. Rafael Prikladnicki -
Orientador

PPGCC/PUCRS

Profa. Dra. Renata Vieira -

PPGCC/PUCRS

Profa. Dra. Tayana Uchôa Conte -

UFAM

Homologada em 20/11/2014, conforme Ata No. 023 pela Comissão Coordenadora.

Prof. Dr. Luiz Gustavo Leão Fernandes
Coordenador.

PUCRS

Campus Central

Av. Ipiranga, 6681 - P32- sala 507 - CEP: 90619-900

Fone: (51) 3320-3611 - Fax (51) 3320-3621

E-mail: ppgcc@pucrs.br

www.pucrs.br/facin/pos

AGRADECIMENTOS

Primeiramente, agradeço a minha noiva Fernanda que sempre foi muito compreensiva, e entendia a importância dos inúmeros fins de semana e madrugadas acordados que possibilitaram a conclusão desta dissertação. Agradeço aos meus familiares pelo apoio e incentivo que foram fundamentais nestes últimos dois anos.

A professora Dra. Renata Vieira, agradeço pela participação como revisora dos trabalhos apresentados nestes dois anos. Obrigado pelas críticas, sugestões e pela disponibilidade.

A professora Dra. Tayana Conte, agradeço pela disponibilidade de viajar até Porto Alegre para participar da banca avaliadora deste trabalho.

Ao meu orientador, Prof. Dr. Rafael Prikladnicki, agradeço pela oportunidade, pela disponibilidade, e pela cuidadosa orientação e direcionamento da pesquisa que foram fundamentais para sua conclusão.

A HP Brazil R&D, pela indicação a bolsa de pesquisa que permitiu a realização deste trabalho.

Por fim, agradeço a PUCRS pela oportunidade e excelente infraestrutura oferecida.

A todos, muito obrigado.

MÁQUINAS DE TRADUÇÃO APLICADA À COMUNICAÇÃO EM TEMPO REAL PARA DESENVOLVIMENTO DISTRIBUÍDO DE SOFTWARE

RESUMO

O aumento da globalização possibilita que mais empresas busquem parcerias para execução de projetos em outros países. O domínio do idioma inglês, por exemplo, é considerado um dos principais fatores para o sucesso de alguns países como Irlanda e Índia em desenvolvimento global de software. Entretanto, a consultoria A.T Kearney estima que o Brasil tem apenas 5,4% da população fluente em inglês, e isto acaba se tornando um desafio para o desenvolvimento de software com equipes distribuídas geograficamente com integrantes que não dominam o mesmo idioma. Com o objetivo de reduzir as barreiras impostas pela diferença de idioma, estudos têm sido conduzidos para avaliar de que forma máquinas de tradução podem ser utilizadas em um contexto de Desenvolvimento Distribuído de Software, a fim de minimizar ou transpor tais barreiras. O tema desta pesquisa está inserido neste contexto e teve como objetivo estudar como máquinas de tradução podem ser utilizadas em projetos de desenvolvimento distribuído de software. Inicialmente foram realizados estudos na área de reconhecimento de voz e máquina de tradução, visando identificar tecnologias de tradução e reconhecimento de voz disponíveis. As tecnologias foram estudadas e foi então proposto um estudo de simulação com uma destas tecnologias, para avaliar se ela poderia ser utilizada para compor um tradutor de voz para o idioma português brasileiro. Por fim, os resultados da simulação levantaram indícios de que as tecnologias do Google para reconhecimento de voz no idioma português e para tradução do português para o inglês possuem níveis adequados de precisão e poderiam ser utilizadas no desenvolvimento de uma máquina de tradução de voz para o idioma português.

Palavras chaves: Máquinas de tradução em tempo real, Tradução de voz, Tradução de texto, Reconhecimento de voz.

MACHINE TRANSLATION APPLIED TO REAL TIME GLOBAL SOFTWARE

ABSTRACT

The globalization allows companies to seek partnership for global project development in other countries. The domain of spoken English is considered one of the main reasons for the success of countries such as Ireland and India in global software development. However, A.T Kearney consultancy estimates that Brazil has only 5,4% of population fluent in English, and this is an important challenge that makes difficult the growth of geographically distributed software development in the country. In order to reduce the barriers posed by language differences, studies have been conducted to assess how machine translation can be used in the context of global software development in order to minimize or overcome such barriers. The goal of this research was in this context and aimed to study how speech translation can be used in global software development projects. In this research we first executed studies related to speech recognition and machine translation areas and identified available technologies for translation and speech recognition. After having identified these technologies, a simulation study was planned and executed in order to evaluate one of the identified technologies. The purpose of this study was to understand the feasibility of the available technology for speech recognition and translation for Brazilian Portuguese. The findings of the simulation indicate that Google technologies for speech recognition and translation for Brazilian Portuguese have adequate accuracy and could be used to create a complete machine translation in this language.

Keywords: real time machine translation, speech translation, text translation, speech recognition.

LISTA DE FIGURAS

Figura 1 – Foco da pesquisa em desenvolvimento offshore	24
Figura 2 - Desenho de pesquisa [CON07]	27
Figura 3 - Arquitetura de máquina de tradução de voz	35
Figura 4 - Gartner 2012: Tecnologias emergentes.....	41
Figura 5 - Publicações por ano	46
Figura 6 - Publicações por tecnologia	47
Figura 7 - Publicações no idioma português	47
Figura 8 – Google Web Speech API: Avaliação reconhecimento - Média e Mediana	63
Figura 9 – Google Web Speech API: Avaliação reconhecimento por vocabulário - Media e Mediana	64
Figura 10 – Google Translator: Avaliação da tradução - Média e Mediana	66
Figura 11 – Google Translator: Avaliação da tradução por vocabulário - Media e Mediana.....	66
Figura 12 - Google: Comparativo reconhecimento e tradução.....	67
Figura 13 - Comparativo reconhecimento e tradução para vocabulário simples.....	68
Figura 14 - Comparativo reconhecimento e tradução para vocabulário complexo ...	68

LISTA DE TABELAS

Tabela 1 - Aplicações de tecnologias de reconhecimento de voz.....	48
Tabela 2 - Tecnologias de reconhecimento de voz.....	49
Tabela 3 - Tecnologias de reconhecimento de voz para o português brasileiro	51
Tabela 4 - Instalação realizada para o Microsoft Speech Platform for Servers.....	54
Tabela 5 - IBM ViaVoice: Tempo de reconhecimento.....	54
Tabela 6 - IBM ViaVoice: Tempo de reconhecimento por tipo de vocabulário.....	55
Tabela 7 - IBM ViaVoice: Qualidade do reconhecimento.....	55
Tabela 8: IBM ViaVoice: Qualidade do reconhecimento - Aceitável x Não Aceitável	55
Tabela 9 - IBM ViaVoice: Qualidade do reconhecimento por vocabulário – Aceitável x Não Aceitável.....	56
Tabela 10 - IBM ViaVoice: Avaliação geral – Média e Mediana.....	56
Tabela 11 - Avaliação geral por vocabulário - Média e Mediana	56
Tabela 12 – Google Web Speech API: Qualidade do reconhecimento.....	62
Tabela 13 – Google Web Speech API: Qualidade do reconhecimento – Aceitável X Não Aceitável	62
Tabela 14 – Google Web Speech API: Qualidade do reconhecimento por vocabulário – Aceitável X Não Aceitável	63
Tabela 15 – Google Translator: Qualidade da tradução	64
Tabela 16 – Google Translator: Qualidade da tradução – Aceitável X Não Aceitável	65
Tabela 17 – Google Translator: Qualidade da tradução por vocabulário – Aceitável X Não Aceitável	65

LISTA DE ABREVIATURAS E SIGLAS

MT – Máquina de tradução

DDS – Desenvolvimento Distribuído de Software

ASR – *Automatic Speech Recognition*

TTS – *Text to Speech*

PUCRS – Pontifícia Universidade Católica do Rio Grande do Sul

SUMÁRIO

1	INTRODUÇÃO	21
1.1	Justificativa	22
1.2	Objetivos.....	23
1.3	Contexto	23
1.4	Organização do volume	24
2	MÉTODO DE PESQUISA	25
3	REFERENCIAL TEÓRICO.....	29
3.1	Desenvolvimento distribuído de software	29
3.2	Máquinas de tradução	32
3.2.1	Abordagem baseada em regras.....	33
3.2.2	Abordagem baseada em corpus	33
3.3	Tradução de voz	34
3.4	Estudos Relacionados	36
3.5	Reconhecimento de voz: Uma Revisão Sistemática de Literatura	38
3.5.1	Motivação e contexto	39
3.5.2	Protocolo da revisão sistemática da literatura.....	41
3.5.3	Resultados	45
4	AVALIAÇÃO DAS TECNOLOGIAS PARA O PORTUGUÊS.....	53
4.1	Microsoft Speech Platform for Servers	53
4.2	IBM ViaVoice	54
4.2.1	Análise dos resultados	54
4.3	Coruja	57
4.4	Avaliação geral das tecnologias	58

4.5	Outras tecnologias	59
5	AVALIAÇÃO DE TECNOLOGIA – SIMULAÇÃO	60
5.1	Planejamento.....	60
5.2	Análise dos resultados.....	61
5.3	Qualidade do reconhecimento de voz	62
5.4	Qualidade da tradução.....	64
5.5	Análise da qualidade geral.....	67
6	CONCLUSÕES	69
6.1	Contribuições da pesquisa.....	69
6.2	Limitações desta pesquisa.....	70
6.3	Considerações finais.....	71
	REFERÊNCIAS BIBLIOGRÁFICAS	73
	ANEXO 1: Artigos resultantes da revisão sistemática	86
	ANEXO 2 - IBM ViaVoice: Protocolo para avaliação da qualidade do reconhecimento de voz	98
	ANEXO 3 – IBM ViaVoice: Resultados da avaliação do reconhecimento de voz.....	102
	ANEXO 4 - Google: Protocolo para avaliação das tecnologias de reconhecimento e tradução.....	104
	ANEXO 5 – Google: Resultado da simulação	109

1 INTRODUÇÃO

A globalização tem causado significativas mudanças nas empresas de desenvolvimento de software. Como resultado, o Desenvolvimento Distribuído de Software (DDS) tem se consolidado como uma forma de desenvolvimento de software, impulsionado, entre outros fatores, pela disponibilidade e custo dos recursos humanos ao redor do mundo [SMI13]. DDS é caracterizado quando membros das equipes estão tipicamente em países diferentes (geralmente em continentes diferentes), porém colaborando em um mesmo projeto [LAN08]. Dessa forma, cada vez mais projetos estão sendo desenvolvidos em ambientes geograficamente distribuídos, caracterizando assim o DDS [HER07].

O desenvolvimento global de software é limitado em países com falta de profissionais proficientes em inglês, o idioma mais utilizado neste tipo de projeto [CAL12]. Análise de requisitos, entre outras atividades de Engenharia de Software exigem comunicação intensa e, portanto, são diretamente impactadas por dificuldades de idioma em projetos globais de software [DAM03], [DAM07], [NUS00].

Conforme relatório da consultoria A.T. Kearney escrito em 2011, a proximidade da América Latina com o mercado consumidor americano serve como elo de ligação entre as regiões. Ao mesmo tempo em que a população hispânica cresce nos Estados Unidos, o número de pessoas fluentes em inglês também continua crescendo na América Latina. Essa região apresenta um grande conjunto de habilidades importantes [KEA11]. O Brasil se sobressai no mercado de TI e é uma potencial fonte para empresas encontrarem desenvolvedores de softwares. A indústria de TI do Brasil é grande e está crescendo. A A.T. Kearney estima que o setor emprega 1,7 milhões de pessoas, incluindo programadores, analistas de sistemas e gerentes [CAL12]. O setor cresce 6,5% por ano em média desde 2005, porém a grande maioria das empresas de TI está focada em clientes nacionais e não exporta mão de obra. Dentre os que exportam, empresas dos Estados Unidos estão entre os principais clientes, e representam mais de 80% da demanda, seguida pela América Latina e Europa [BRA10, PRI13].

O domínio do idioma inglês, por exemplo, é considerado um dos principais fatores para o sucesso de alguns países como Irlanda, Filipinas e Índia em desenvolvimento global de software [CAR01]. Entretanto, a A.T. Kearney estima que o Brasil tem apenas 10,2

milhões pessoas fluentes em inglês, ou 5,4% da população [CAL12]. Já Chile e Índia, potenciais concorrentes do Brasil no mercado global de TI, possuem respectivamente 34,7% e 8,2% (90,6 milhões no caso da Índia) da população fluente em inglês, o que é consideravelmente melhor [CAL12]. Em 2009, um estudo publicado pela KPMG indicou que uma das desvantagens dos países latino-americanos é a falta de profissionais fluentes em Inglês [KPM09]. Neste sentido, a falta de domínio do idioma inglês acaba se tornando um desafio no Brasil para o desenvolvimento de software com equipes distribuídas geograficamente com integrantes que não dominam um mesmo idioma. Com o objetivo de minimizar ou transpor as barreiras impostas pela diferença de idioma, estudos têm sido conduzidos para avaliar de que forma máquinas de tradução podem ser utilizadas em um contexto de DDS [CAL10, CAL11, CAL12].

Uma máquina de tradução (MT) pode ser definida como o uso de um computador para traduzir texto de uma linguagem natural, linguagem origem, em outra linguagem, linguagem destino [JUR08]. Em estudos recentes, máquinas de tradução são acopladas a ferramentas de comunicação de mensagem instantânea (chats) de forma a permitir comunicação em tempo real com tradução automática. Esse tipo de aplicativo permite que membros de equipes de projeto que falam idiomas diferentes possam se comunicar em sua língua nativa [CAL10]. Esses estudos avaliaram a efetividade do uso dessas máquinas de tradução de texto em projetos globais de desenvolvimento de software e tem demonstrado que a tradução automática de texto pode ser utilizada sem interrupção do fluxo da conversa [CAL11].

Neste contexto, a comunicação por voz, mediada por máquinas de tradução, surge como uma oportunidade a ser explorada para permitir a interação entre membros de equipe globais de desenvolvimento de software. O tema desta pesquisa se encontra neste contexto e tem como objetivo estudar como máquinas de tradução de voz podem ser utilizadas em projetos distribuídos de desenvolvimento de software.

1.1 Justificativa

A pesquisa atual visa estudar a viabilidade de uma solução alternativa para o problema da falta de profissionais fluentes em idiomas como o inglês em países que possuem potencial para atuação no mercado global de Tecnologia da Informação (TI), como o Brasil. Essa solução permitiria, através do uso integrado de máquinas de tradução

e tecnologias de reconhecimento de voz, que os membros da equipe conversassem no seu idioma nativo, sem a necessidade de utilizar um idioma comum.

No entanto, tal solução não eliminaria a necessidade de capacitação profissional em idiomas como o inglês, apenas permitiria uma forma alternativa de comunicação diante das dificuldades existentes na formação de profissionais, visando ampliar a oportunidade de atuação dos profissionais destes países. Para países emergentes como o Brasil, a terceirização de seus serviços em um contexto global representa uma vantagem para a economia nacional, pois é possível exportar serviços de seus profissionais, garantindo que a mão-de-obra especializada não saia de seu país de origem.

Dessa forma, se o uso de máquinas de tradução de voz em projetos globais de software provar ser viável, então é possível com isso gerar novas oportunidades de negócios para países com forte atuação no mercado de TI e baixo número de profissionais proficientes em idiomas como o inglês.

1.2 Objetivos

O objetivo geral desta pesquisa é estudar como as máquinas de tradução de voz podem ser utilizadas em projetos de desenvolvimento distribuído de software. Os objetivos secundários desta pesquisa são listados abaixo:

- Estudar as tecnologias que compõem uma máquina de tradução de voz.
- Identificar as principais pesquisas e contribuições na área de máquinas de tradução aplicada a projetos de desenvolvimento de software.
- Avaliar a viabilidade de se utilizar máquinas de tradução de voz em projetos de DDS para o português brasileiro.

Sintetizando, a questão de pesquisa definida para este trabalho é: “Como máquinas de tradução de voz podem ser utilizadas em projetos de desenvolvimento distribuído de software?”

1.3 Contexto

Essa pesquisa é parte de um projeto em parceria entre pesquisadores da Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS) e da Universidade de Bari na Itália, financiado parcialmente pela FAPERGS. O foco deste projeto está em realizar pesquisas

científicas que estudam a viabilidade do uso de máquinas de tradução em projetos de DDS. Em relação ao tipo de projeto de DDS, é necessário considerar que a forma de distribuição das equipes pode variar entre empresas. Conforme a Figura 1, *Outsourcing* e *Insourcing* representam formas de relacionamento entre empresas; *Inshore* e *Offshore* representando distribuições geográficas. O foco desta pesquisa está no desenvolvimento distribuído de software realizado por equipes de países diferentes que não possuam um idioma comum conforme destacado na figura abaixo.

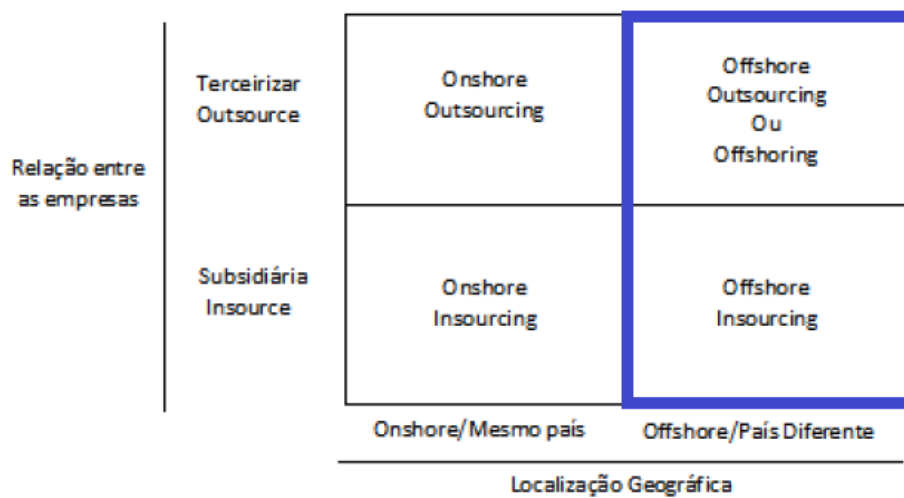


Figura 1 – Foco da pesquisa em desenvolvimento offshore

Neste sentido, a pesquisa propõe estudar como máquinas de tradução de voz podem ser utilizadas para permitir a execução de um projeto global de software sem a necessidade do domínio de um idioma único. Tais máquinas de tradução, normalmente, são compostas de três componentes básicos: reconhecimento de voz, tradução e síntese de voz [WAI08]. Devido à limitação de tempo e recursos, são escopo desta pesquisa os componentes de reconhecimento de voz e tradução, ficando de fora a síntese de voz.

1.4 Organização do volume

Este volume está organizado em 6 seções. Na próxima seção é apresentado o método de pesquisa empregado. A Seção 3 apresenta o referencial teórico. A Seção 4 apresenta uma revisão de literatura realizada envolvendo tradução de voz e seus componentes. A Seção 4 apresenta avaliações realizadas sobre as tecnologias encontradas. A Seção 5 apresenta os resultados de uma simulação realizada envolvendo reconhecimento de voz e tradução de texto. E por fim, a Seção 6 apresenta a conclusão e considerações finais.

2 MÉTODO DE PESQUISA

Tendo por base o objetivo original definido para esta pesquisa, de estudar como máquinas de tradução de voz podem ser utilizadas em projetos de DDS, foi proposto o desenvolvimento de um estudo empírico. O uso de estudos empíricos para estudar o desenvolvimento de software em situações realistas permite a avaliação de tecnologias maduras e identificação de problemas presentes em tecnologias menos maduras. Através destes estudos é possível avaliar a efetividade das ferramentas de desenvolvimento e métodos propostos em vários ambientes e situações diferentes [SHU01].

A introdução de tecnologias de software recém-definidas no contexto industrial pode trazer consequências indesejáveis caso a tecnologia não possua um grau adequado de maturidade [MAF06]. Diante disso, foi utilizada a metodologia descrita em [MAF06] como base para o desenvolvimento desta pesquisa (Figura 2). Essa metodologia descreve como conduzir estudos de novas tecnologias desde a fase conceitual até a introdução dessas tecnologias na indústria.

Cada uma das etapas desta metodologia é detalhada a seguir [MAF06, SHU01]:

- Estudos Secundários: essa etapa consiste na condução de estudos secundários (tais como revisão sistemática da literatura) com objetivo de buscar evidências primárias na área em estudo. Os procedimentos a serem seguidos durante a condução da revisão sistemática podem amadurecer o problema e reduzir drasticamente riscos associados à definição de uma tecnologia. Como consequência, oportunidades de melhoria pertinentes à área de pesquisa podem ser também identificadas.
- Proposta Inicial: essa etapa consiste na produção, com base no conhecimento adquirido e nas evidências identificadas através da condução da revisão sistemática, de uma proposta inicial da tecnologia.
- Estudo de Viabilidade: essa etapa verifica, através de estudos empíricos e experimentais, a viabilidade do processo ou tecnologia sendo analisada. O maior objetivo do pesquisador é gerar novas hipóteses sobre o novo processo e sua utilidade, não sendo possível obter o controle total das variáveis relacionadas ao estudo. Pesquisas utilizando estudantes são aplicáveis nesta etapa. Embora seus resultados não possam ser diretamente aplicados à indústria, estudos com alunos

podem permitir que novos conceitos sejam testados antes de utilizar os profissionais da indústria.

- **Estudo de Observação:** essa etapa avalia todos os passos, em detalhes, que constituem o processo de aplicação da nova tecnologia de forma a garantir que cada passo é efetivo e é executado na ordem correta. A forma mais efetiva de avaliar cada passo do processo é através de estudos observacionais. Nestes estudos o participante realiza uma tarefa específica enquanto observado pelo pesquisador, o objetivo é coletar dados sobre como as tarefas são executadas por cada indivíduo e quais os tipos de problemas enfrentados.
- **Estudo de Caso com Ciclo de Vida Real:** essa etapa consiste de um caso de estudo avaliando o processo ou tecnologia em um ciclo de vida real de desenvolvimento de software. Neste momento, é avaliado como o novo processo interage com outros aspectos do ciclo de desenvolvimento de software. Essas interações podem levar a problemas que não foram identificados nas fases anteriores onde o processo estava sendo analisado de forma isolada.
- **Estudo de Caso na Indústria:** uma vez que nas etapas anteriores o processo já foi adaptado para ser executado em um ciclo de vida real e demonstrou ser eficaz, nesta fase o processo ou tecnologia é avaliado em um ambiente de indústria. Nesta etapa é realizado um estudo de caso para investigar como o novo processo reage em um ambiente industrial e quais são os problemas não previstos desta interação. Essa etapa ficou por último, pois estudos envolvendo profissionais da indústria requerem altos investimentos, e o objetivo do pesquisador deve ser avaliar a tecnologia tanto quanto possível antes de solicitar que parceiros da indústria invistam seu tempo e dinheiro.

É importante ressaltar que a metodologia descrita anteriormente é iterativa, e que cada etapa pode ser executada múltiplas vezes adaptando o processo ou tecnologia até que os resultados esperados sejam atingidos.

Neste contexto, a pesquisa aqui proposta teve seu escopo definido para os passos iniciais, conforme destacado na Figura 2. O objetivo foi realizar um estudo de viabilidade do uso de máquinas de tradução de voz em projetos globais de desenvolvimento de software, a partir da condução de estudos secundários e da proposta do uso de tradução de voz no contexto de projetos de DDS globais.

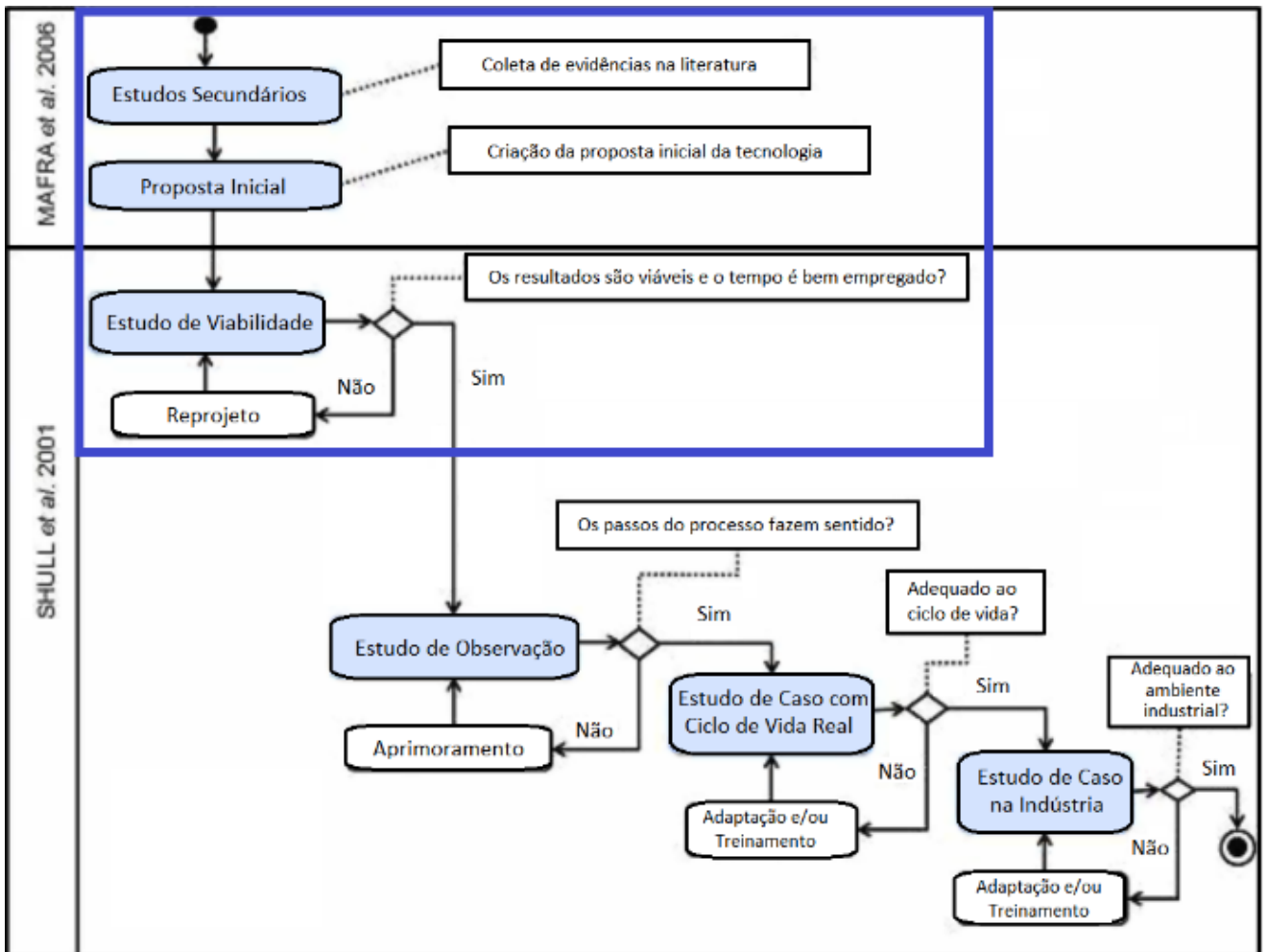


Figura 2 - Desenho de pesquisa [CON07]

A seguir, temos o detalhamento das atividades que foram realizadas como parte dessa pesquisa.

1) Realização de estudos secundários com objetivo de conhecer a área sendo estudada, amadurecer o entendimento sobre o problema de pesquisa e conhecer as tecnologias disponíveis:

- a. Investigação sobre a existência de estudos envolvendo tradução automática de voz para o idioma português brasileiro aplicado em um contexto de desenvolvimento distribuído de software ou engenharia de software.
- b. Investigação sobre a existência de tecnologias de tradução automática de voz para o português brasileiro. Como não foram encontrados estudos nem tecnologia de tradução de voz madura para o português brasileiro, optou-se por investigar e avaliar as tecnologias relacionadas com reconhecimento de voz e tradução de texto de forma geral.

- c. Revisão sistemática de literatura sobre tecnologias de reconhecimento de voz. A tecnologia de reconhecimento voz é componente importante para a construção de um tradutor de voz, e por isso optou-se pelo uso da revisão sistemática de literatura para identificar contribuições e tecnologias existentes nesta área.
- 2) Investigação exploratória (pesquisa ad hoc) na internet com objetivo de encontrar tecnologias comerciais ou que ainda não tivessem sido exploradas no meio acadêmico para reconhecimento de voz do português brasileiro. Essa pesquisa foi importante para garantir que foram identificadas tecnologias existentes tanto no meio acadêmico como tecnologias disponíveis em ferramentas comerciais.
- 3) Elaboração de proposta inicial com base no conhecimento adquirido:
 - a. Proposta de um estudo de simulação para avaliar se tecnologias de reconhecimento de voz e tradução do Google poderiam ser utilizadas para compor um tradutor de voz para o idioma português brasileiro. As frases avaliadas como parte dessa simulação incluíram frases e expressões do dia a dia de um projeto de desenvolvimento de software.
- 4) Execução de estudo de viabilidade com objetivo de avaliar proposta inicial:
 - a. Execução de experimento que simula uma máquina de tradução de voz para o idioma português brasileiro. O experimento usou o Google Web Speech API para reconhecimento de voz e Google Translator para tradução.
- 5) Avaliação da necessidade de reprojeto com base nos resultados do estudo de viabilidade:
 - a. O estudo de viabilidade indicou que as tecnologias do Google apresentam desempenho adequado e poderiam ser utilizadas como componentes no desenvolvimento de um tradutor de voz em tempo real para o português brasileiro. Neste caso não foi necessária alteração na proposta inicial.

3 REFERENCIAL TEÓRICO

Esta seção apresenta o referencial teórico referente aos principais conceitos que envolvem essa pesquisa.

3.1 Desenvolvimento distribuído de software

As últimas décadas têm testemunhado uma tendência constante e irreversível em direção à globalização de negócios, especialmente os relacionados à alta tecnologia e ao desenvolvimento de software. Forças econômicas estão transformando mercados nacionais em mercados globais gerando novas formas de competição e cooperação que vão além dos limites nacionais [HER01, HER07]. O Brasil, por exemplo, assim como outros países emergentes como China e Índia estão aumentando significativamente sua importância como competidores globais. Nos últimos anos, a expansão da indústria nacional e o crescimento econômico estão alavancando a econômica brasileira no mercado global [WEI11]. Esta mudança está tendo um profundo impacto não só na comercialização e distribuição, mas também na forma como os produtos são concebidos, construídos, testados e entregues aos clientes [HER01, HER07].

A globalização tem causado significativas mudanças nas empresas de desenvolvimento de software. Como resultado, o DDS tem se consolidado como uma forma de desenvolvimento de software, impulsionado, entre outros fatores pela disponibilidade e custo dos recursos humanos ao redor do mundo [SMI13]. Dessa forma, o desenvolvimento de software se torna uma atividade que envolve múltiplos locais de trabalho, regiões e culturas, enfim, globalmente distribuída [PRI03].

O desenvolvimento distribuído de software (DDS) surgiu nos anos 90 quando as empresas começaram a desenvolver software com times distribuídos [LAN08]. O DDS é caracterizado sempre que um ou mais membros de equipes envolvidas no projeto estiverem fisicamente distantes dos demais [AUD07].

A forma de distribuição de equipes em DDS varia de acordo com as empresas e projetos a serem desenvolvidos. Como mostrou a Figura 1, podemos classificar o DDS utilizando os conceitos a seguir:

- *Outsourcing* e *Insourcing* (formas de relacionamento entre empresas);
Onshore e *Offshore* (distribuições geográficas).

- *Outsourcing*: empresas contratam os serviços de outra empresa. Também conhecido como terceirização [PRI09]. Essa forma é considerada a mais rápida e fácil de implementar [AUD07, BIN07].
- *Insourcing*: empresas criam seus próprios centros de desenvolvimento. Esta forma permite maior controle sobre os negócios da empresa, maior flexibilidade e menor custo no longo prazo. Essa forma é considerada mais complexa de implementar [AUD07].
- *Offshore*: empresa envia seus projetos ou serviços para empresas terceirizadas ou subsidiárias localizadas em um país diferente da matriz [PIL06, AUD07].
- *Onshore*: empresa envia seus projetos ou serviços para empresas terceirizadas ou subsidiárias localizadas no mesmo país [PRI09].

Países componentes do grupo econômico BRIC (Brasil, Rússia, Índia e China) estão cada vez mais se tornando uma alternativa atraente contratação de *Offshore outsourcing*, pois apresentam baixos custos comparados com Estados Unidos e países da Europa.

Quando a distância física entre os integrantes dos times de um projeto inclui mais de um país, tem-se o Desenvolvimento Global de Software [LAN08]. Alguns dos principais fatores que motivam o desenvolvimento global de software são:

- Necessidade de manter um conjunto global de recursos que podem ser compartilhados globalmente não importando o local onde estão fisicamente, tornando os custos mais competitivos [HER01].
- A vantagem da proximidade do mercado, incluindo conhecimento dos clientes e condições locais [HER01, LAN08 KNO07].
- A necessidade de formar rapidamente empresas e equipes virtuais para explorar oportunidades de negócios [HER01].
- Grande pressão para colocar novos produtos mais rapidamente no mercado através do uso de equipes distribuídas em diferentes fusos horários permitindo desenvolvimento 24/7 [HER01]. Abordagens como essa permitem reduzir o tempo dos projetos [LAN08, DAM06, PRI08].

- Necessidade de reduzir custos através da terceirização da produção para países onde o custo de produção seja mais baixo [LAN08, DAM06, PRI08, AUD07, MAR09, KNO07].

Embora o desenvolvimento global de software traga inúmeras vantagens para as empresas, ele também introduz uma série de novos desafios à tarefa de desenvolver software. Os principais desafios são:

- Distância e diferenças de fuso horário: grande distância física entre os membros da equipe assim como diferenças de fuso horário podem levar a problemas de comunicação, coordenação e gerenciamento do projeto trazendo uma série de inconvenientes [HER01, WU12]. Por exemplo, diferenças de fuso horário podem impedir a realização de uma ligação rápida para esclarecer uma dúvida sobre um determinado algoritmo. Já a distribuição da equipe em regiões distantes impede que os membros da equipe tenham a oportunidade de conversar informalmente em corredores e almoços diminuindo a proximidade dos membros da equipe [MOC01, WU12].
- Diferenças culturais como tradições, costumes e comportamentos: em times globais problemas relacionados à cultura emergem muito rapidamente. Culturas diferentes pensam, comunicam-se e agem de formas distintas [BIN07]. Em determinado país algo pode ser considerado trivial e corriqueiro, mas em outro local pode ser algo completamente incomum ou uma ofensa.
- Baixa confiança: pode haver falta de confiança e falta de vontade de comunicar abertamente entre membros de regiões diferentes. Isto pode ser causado por preocupações relacionadas à estabilidade no emprego, ou seja, pode existir a percepção de que ao facilitar o trabalho do colega de outra região, o profissional está ao mesmo tempo tornando-se substituível e, por fim, colocando seu emprego em risco [MOC01].
- Diferença de idioma nativo: língua falada é uma componente importante relacionado à distância cultural nacional. Muitas empresas hesitam ou expressam preocupação sobre estabelecer parcerias ou colaborar com nações onde nível de fluência em inglês é fraco. O domínio do inglês é considerado umas das principais razões para o sucesso de projetos globais de TI de países como Filipinas e Singapura [CAR01].

As empresas americanas têm investido na elaboração e disponibilização de cursos de inglês para aqueles profissionais estrangeiros que não possuem a fluência necessária com o objetivo de tentar melhorar a comunicação dentro da empresa [CAR01]. Além disso, estudos envolvendo máquinas de tradução têm sido realizados de forma a tentar obter soluções alternativas para o problema da diferença de idioma em desenvolvimento distribuído de software [CAL10].

3.2 Máquinas de tradução

Uma máquina de tradução (MT) pode ser definida como o uso de um computador para traduzir texto de uma linguagem natural (*text-to-text translation*), linguagem origem, em outra linguagem, linguagem destino [JUR08]. Durante a década de 1950 o interesse em máquinas de tradução foi alimentado por visões de alta velocidade e alta qualidade de tradução de textos arbitrários, especialmente aqueles de interesse militar que muito frequentemente financiaram os projetos de MT. Nessa época acreditava-se estar bem próximo de construir uma máquina de tradução perfeita [SLO84]. Já na década de 1960, a desilusão crescia conforme a complexidade dos problemas linguísticos se tornava mais aparente e neste momento perceberam que o problema da tradução não era tão simples como se pensava. [SLO84].

Nas últimas décadas, os estudos de máquinas de tradução e tecnologias relacionadas tem se tornado um dos principais temas das pesquisas atuais. O crescente interesse e empolgação relacionados a esta área podem ser explicados pela união de dois fatores: grande aumento da capacidade de processamento dos computadores; e necessidade crescente e valorização de tecnologia que permita comunicação entre pessoas de idiomas diferentes em um mundo globalizado [WAI08].

O problema da máquina de tradução é complicado, principalmente porque a tradução, por si só envolve uma quantidade enorme de conhecimento humano que deve ser codificado de forma a poder ser processado por uma máquina. Línguas naturais são altamente ambíguas; duas línguas nem sempre expressam o mesmo conteúdo do mesmo modo [ARN03, CAL10].

Embora existam abordagens híbridas, máquinas de tradução podem ser classificadas em duas categorias principais de acordo com o modelo linguístico sendo

utilizado: baseada em regras e baseada em *corpus*. A seguir, vemos a definição de cada um desses conceitos.

3.2.1 Abordagem baseada em regras

A abordagem de tradução baseada em regras (*rule-based*) utiliza informações sobre a linguagem origem e destino expressas na forma de regras, explicitamente codificadas por especialistas humanos, para traduzir textos. Essas regras cobrem semântica, morfologia e sintaxe de cada linguagem. Essa abordagem depende fortemente de conhecimento da linguagem assim como de dicionários bilíngues [ARN03].

As maiores vantagens desta abordagem incluem: (a) traduções mais precisas, isto é, traduções mais fiéis ao significado do texto original; (b) a capacidade de codificar explicitamente conhecimentos linguísticos para que ambos seres humanos e sistemas computacionais possam entender; (c) facilidade de diagnosticar e corrigir erros de tradução como regras erradas em dicionários [CAL10].

No entanto, a tradução baseada em regras tem suas desvantagens, destacando especialmente o considerável esforço humano requerido para desenvolver os recursos linguísticos que permitam esse tipo de tradução, como por exemplo, regras morfológicas, dicionários bilíngues, regras de transferência lexicais e estruturais.

3.2.2 Abordagem baseada em corpus

A abordagem de tradução baseada em corpus (*corpus-based*) utiliza grandes coleções de pares de textos como fonte de conhecimento a partir do qual o sistema aprende a realizar traduções. Cada par consiste no texto na linguagem origem e no equivalente já traduzido na linguagem destino [OCH04, CAL10].

Esse tipo de abordagem tende a produzir traduções mais fluentes do que os sistemas baseados em regras, que ao invés costumam produzir traduções mais “mecânicas”. No entanto, tal abordagem requer grandes quantidades de textos paralelos (na ordem de dezenas de milhões de palavras) para atingir razoável qualidade da tradução [OCH04]. Comparado ao *rule-based*, a abordagem *corpus-based* é particularmente atraente para os pesquisadores, porque os sistemas podem ser treinados automaticamente, sem qualquer intervenção humana direta.

O tradutor do Google é um exemplo de sistema estatístico de máquina de tradução que segue uma abordagem baseada em corpus. Na verdade, o sistema não aplica as

regras gramaticais, uma vez que seus algoritmos são com base na análise estatística diferentemente da técnica tradicional baseada em regras.

Tradutores estatísticos como o do Google aplicam técnicas estatísticas de aprendizagem para a construção de modelos de linguagem e de tradução a partir de um grande número de textos, tanto textos na língua destino como exemplos de textos na língua origem já traduzidos manualmente para a língua destino [CAL10].

3.3 Tradução de voz

A tecnologia de tradução de voz (*speech to speech translation*) tenta romper as barreiras da linguagem e permitir que pessoas com diferentes idiomas nativos possam conversar cada um utilizando sua própria linguagem nativa [PAU10]. Governos, empresas comerciais, organizações acadêmicas e humanitárias todas enfrentam a internacionalização e globalização em uma escala sem precedentes, todos dependem de informação e da capacidade de interagir e comunicar [WAI08].

A importância dessa tecnologia está aumentando significativamente devido ao grande número de oportunidades para comunicação entre pessoas de diferentes idiomas, tanto profissionalmente como em viagens de turismo [PAU10].

Maior integração requer natural e eficaz comunicação internacional entre diferentes idiomas. É verdade que existem línguas comuns de comunicação (Inglês, Espanhol, Mandarim), mas as habilidades de linguagem variam e muitas vezes impedem integração real e oportunidades iguais para todos.

Para a comunidade de pesquisa, tradução de voz também apresenta problemas novos e fascinantes que parecem ser solucionáveis, principalmente, devido ao maior poder computacional, a aparentemente ilimitada informação disponível na internet e as promissoras novas técnicas de máquina de tradução [WAI08].

Conforme a Figura 3, esses sistemas combinam tecnologias de reconhecimento automático de voz (ASR), máquina de tradução (MT) e síntese de voz (TTS) para respectivamente compreender o que foi dito, traduzir para uma língua destino e, por fim, falar o texto traduzido na língua destino [STA10].

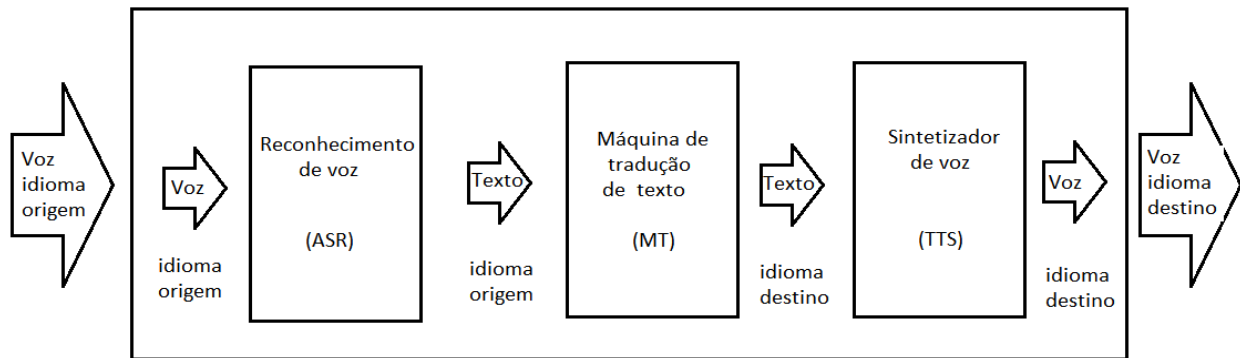


Figura 3 - Arquitetura de máquina de tradução de voz

Recentemente, um progresso considerável foi feito em sistemas de tradução de voz (S2S) [WAI08]. Os primeiros tradutores de voz foram criados no final de 1980 [WAI91, MOR93, ROE92]. Eles demonstraram que a tradução de voz era possível e atraíram grande atenção. Alguns exemplos de sistemas de tradução de voz da época são JANUS [WAI91] e ASURA [MOR93]. Esses sistemas não permitiam diálogo livre e requeriam que o usuário utilizasse frases seguindo uma determinada estrutura, além de terem um vocabulário restrito.

As restrições impostas pelos tradutores de voz iniciais eram muito grandes, pois seres humanos geralmente não conseguem restringir o estilo de falar a uma forma específica ou mesmo lembrar de utilizar um mesmo conjunto de palavras quando interagindo com o sistema. Entretanto, seres humanos conseguem se limitar a utilizar um domínio ou vocabulário específico para resolver algumas tarefas desde que o sistema possua a precisão e velocidade necessária. Alguns exemplos de tarefas que podem fazer uso destes tradutores de voz específicos são: reserva de hotel, aluguel de carros, emergências médicas, registro em hotel, missões policiais ou militares [WAI08].

Neste sentido, novas tecnologias foram desenvolvidas para tratar o problema do estilo de falar, permitindo que os usuários falassem espontaneamente. Os primeiros sistemas de tradução de voz “espontânea” foram desenvolvidos no início da década de 1990. Alguns exemplos destes sistemas de tradução são: C-STAR (Consortium for Speech Translation Advanced Research) [WAI08], Verbmobil [BUB96], Nespole (Negotiating through Spoken Language in Ecommerce) [WAI08] e Transtac (Translation System for Tactical Use) [WAI08].

Embora existisse um grande número de aplicações para máquinas de tradução de voz de domínio específico, ainda existia um grande número de situações que não podiam

ser atendidas por esse tipo de máquina de tradução. Alguns exemplos de situações que requerem máquinas de tradução de domínio ilimitado são: tradução de notícias de rádio, discursos parlamentares, palestras acadêmicas, conversas de telefone, e reuniões em geral.

Nos anos 2000, várias pesquisas iniciaram para tentar desenvolver máquinas de tradução de voz que fossem aplicáveis a qualquer domínio como os citados anteriormente. Vemos alguns exemplos dessas pesquisas abaixo:

- TC-STAR (Technology and Corpora for Speech to Speech Translation): o foco deste projeto é permitir a automática transcrição e tradução de discursos parlamentares [HAM07].
- GALE (Global Autonomous Language Exploitation): o foco deste projeto é a tradução de notícias de rádio de diversas linguagens para o inglês [COH07].

De forma a possibilitar tradução livre de domínio específico, normalmente são utilizadas máquinas de tradução estatísticas baseada em corpus paralelos, pois essa abordagem permite que os sistemas aprendam sozinhos [WAI08].

Em resumo, as máquinas de tradução de voz progrediram de sistemas altamente limitados para sistemas capazes de traduzir discurso espontâneo sobre qualquer tópico.

3.4 Estudos Relacionados

Até o momento não foram identificados estudos relacionados envolvendo máquinas de tradução de voz e DDS. Entretanto, diversos estudos têm sido realizados com o objetivo de verificar se máquinas de tradução podem ser utilizadas para permitir a comunicação entre pessoas que não dominam um idioma comum [YAS11, YAN11, CAL10, HAU10].

Em [YAS11], um fórum de discussão de internet é utilizado para realizar um experimento entre alunos. Este fórum possui integração com máquinas de tradução de texto, e automaticamente traduz o conteúdo das mensagens postadas para o idioma do usuário que está lendo permitindo assim a comunicação de pessoas utilizando sua língua nativa. O fórum foi utilizado para realizar um experimento entre 80 estudantes envolvendo a língua inglesa e japonesa. Esse experimento demonstrou que embora os participantes tivessem mencionado dificuldades de entender as mensagens traduzidas, eles conseguiram terminar com sucesso as tarefas propostas no experimento. Diferentemente da expectativa inicial de que a qualidade dos resultados da tradução automática

determinaria o sucesso da comunicação, os dados indicam que a comunicação efetiva se baseia mais na capacidade dos participantes de estabelecerem um terreno comum (*common ground*) na conversação.

Em [YAN11, CAL10, HAU10], é demonstrado que é possível integrar máquinas de tradução de texto a ferramentas de comunicação instantânea para permitir comunicação em tempo real entre indivíduos que possuem diferentes idiomas nativos.

Seguindo a linha de estudos relacionados a máquinas de tradução integradas a ferramentas de comunicação instantâneas, no estudo [HAU10] foi realizado um experimento que consiste de um jogo envolvendo participantes com diferentes idiomas nativos interagindo através de um chat com tradução automática. Este experimento dividiu as mensagens do experimento em duas categorias descritas abaixo.

- sócio emocionais: essas mensagens são usadas para manter o relacionamento entre os participantes do experimento. Podem ser positivas como elogios e agradecimentos, ou negativas indicando frustração e pouca disposição para colaborar. Uma grande quantidade de mensagens sócio emocionais positivas pode ser indicativo de produção de trabalho de alta qualidade pelo grupo.
- relacionadas a tarefas: essas mensagens são relacionadas as interações entre os participantes referentes a execução de tarefas do experimento como, por exemplo, realizar ou responder perguntas.

Esse experimento indicou que máquinas de tradução não somente aumentam a qualidade e desempenho geral da comunicação, mas como também aumentam o número de mensagens sócio emocionais positivas [HAU10].

Os estudos [CAL10, CAL11, CAL12] investigam se a adoção de máquinas de tradução também integradas a ferramentas de comunicação instantânea podem beneficiar projetos globais de desenvolvimento de software. A pesquisa conduz uma série de experimentos que consistem na execução de atividades como reuniões de priorização e seleção de requisitos entre participantes do Brasil e da Itália fazendo uso de chats com e sem tradução automática. O foco em reuniões de discussão de requisitos deve-se ao fato dessas reuniões requererem intensa comunicação.

A pesquisa ainda está em desenvolvimento, e por isso não existem ainda resultados conclusivos. Até o momento, a pesquisa demonstrou que a tradução não interrompe o fluxo da conversa e é bem aceita pelos participantes do experimento.

3.5 Reconhecimento de voz: Uma Revisão Sistemática de Literatura

Uma revisão sistemática de literatura é uma forma de identificar, avaliar e interpretar a pesquisa disponível para uma determinada questão de pesquisa, área ou fenômeno de interesse [KIT07]. Estudos individuais que contribuem para uma revisão sistemática são chamados de estudos primários.

As razões mais comuns para a execução de uma revisão sistemática são [KIT07]:

- Agrupar ou resumir evidências existentes relacionadas a um determinado tratamento ou tecnologia.
- Identificar lacunas em uma presente pesquisa de forma a ser capaz de sugerir novas áreas de estudo ou investigação.
- Proporcionar uma base de conhecimento para justificar novas atividades de pesquisa.

Para essa pesquisa, a razão para a execução de uma revisão sistemática de literatura se baseia na necessidade de agrupar e resumir evidências existentes relacionadas a tecnologias de reconhecimento de voz, caracterizando assim uma revisão sistemática do tipo mapeamento sistemático.

É comum iniciar uma pesquisa com um algum tipo de revisão de literatura. Entretanto, se não for realizada de forma completa e rigorosa acredita-se que terá pouco valor científico [KIT07]. Uma revisão sistemática é realizada de acordo com uma estratégia pré-definida. A revisão sistemática requer consideravelmente mais esforço que uma revisão tradicional e sua maior vantagem é prover informações sobre os efeitos de um determinado fenômeno em um grande conjunto de ambientes e métodos. Se o estudo fornece resultados consistentes, a revisão sistemática pode prover evidências de que o fenômeno é robusto e transferível. Se os estudos geram resultados inconsistentes, então a origem da variação pode ser estudada. Algumas características que diferem a revisão sistemática da convencional são [KIT07]:

- Definem um protocolo de revisão que especifica a questão de pesquisa a ser respondida e os métodos que serão utilizados para realizar a revisão.

- São baseadas em uma estratégia de busca que busca detectar o maior número de literatura relevante possível.
- Documenta a estratégia de busca para que os leitores possam acessar e avaliar quanto ao rigor utilizado e completude.
- Deixa claro critérios de inclusão e exclusão para avaliar cada potencial estudo primário.
- Especifica a informação a ser obtida de cada estudo primário incluindo critério de qualidade pelos quais são avaliados cada estudo primário.
- É pré-requisito para análise quantitativa.

As fases e tarefas de uma revisão sistemática podem ser resumidas conforme a seguir [KIT07]:

- Planejamento da revisão: identifica necessidade, desenvolve protocolo de revisão.
- Condução da revisão: identifica a pesquisa, seleciona os estudos de acordo com critérios definidos, avalia a qualidade dos estudos, extrai informações e sintetiza os resultados encontrados.
- Reporte da revisão: reporta publicamente a revisão sistemática realizada.

3.5.1 Motivação e contexto

Uma máquina de tradução de voz normalmente é composta de três componentes: (1) reconhecimento de voz, (2) tradução de texto e (3) síntese de voz [WAI08]. Devido à limitação de tempo, a síntese de voz não faz parte do escopo da pesquisa de máquinas de tradução de voz. Para projetar um tradutor de voz é necessário ter disponível cada componente isoladamente.

A tradução de texto no idioma português é possível através de sistemas atualmente disponíveis no mercado e que oferecem bom desempenho como o tradutor estatístico do Google já utilizado em experimentos de tradução de texto [CAL11, CAL12].

A pesquisa envolvendo reconhecimento de voz evoluiu bastante nos últimos anos e está movendo dos laboratórios para aplicações de mercado [MEN12].

A Figura 4 representa graficamente a maturidade, adoção e principais tendências em tecnologia do ano de 2012, essa figura também é conhecida como *hype cycle* e foi desenvolvido pelo instituto de pesquisas Gartner. De acordo com o *hype cycle*, uma tecnologia pode ser classificada em 5 fases diferentes no decorrer do tempo [GAR13]:

- *Technology trigger*: a primeira fase do hyper cycle representa um avanço significativo referente a uma tecnologia e que desperta interesse no tema. Recentes histórias sobre provas de conceitos surgem na mídia. Geralmente nenhum produto utilizável existe.
- *Peak of inflated expectations*: na segunda fase existe uma exagerada publicidade, entusiasmo e expectativas não realísticas sobre a tecnologia em questão. Podem existir algumas aplicações de sucesso da tecnologia, mas essas são muito raras.
- *Trough of disillusionment*: tecnologias entram nessa fase por que falham em atender as expectativas e rapidamente se tornam fora de moda, conseqüentemente a imprensa abandona o tema ou tecnologia.
- *Slope of enlightenment*: mais exemplos de como a tecnologia pode beneficiar as empresas começam a surgir e se tornam amplamente entendidos por todos. Segunda e terceira geração de produtos começa a ser desenvolvida. Mais empresas financiam projetos pilotos.
- *Plateau of productivity*: adoção geral da tecnologia começa a acontecer. Tecnologia encontra ampla aplicabilidade e relevância no mercado justificando investimentos na mesma.

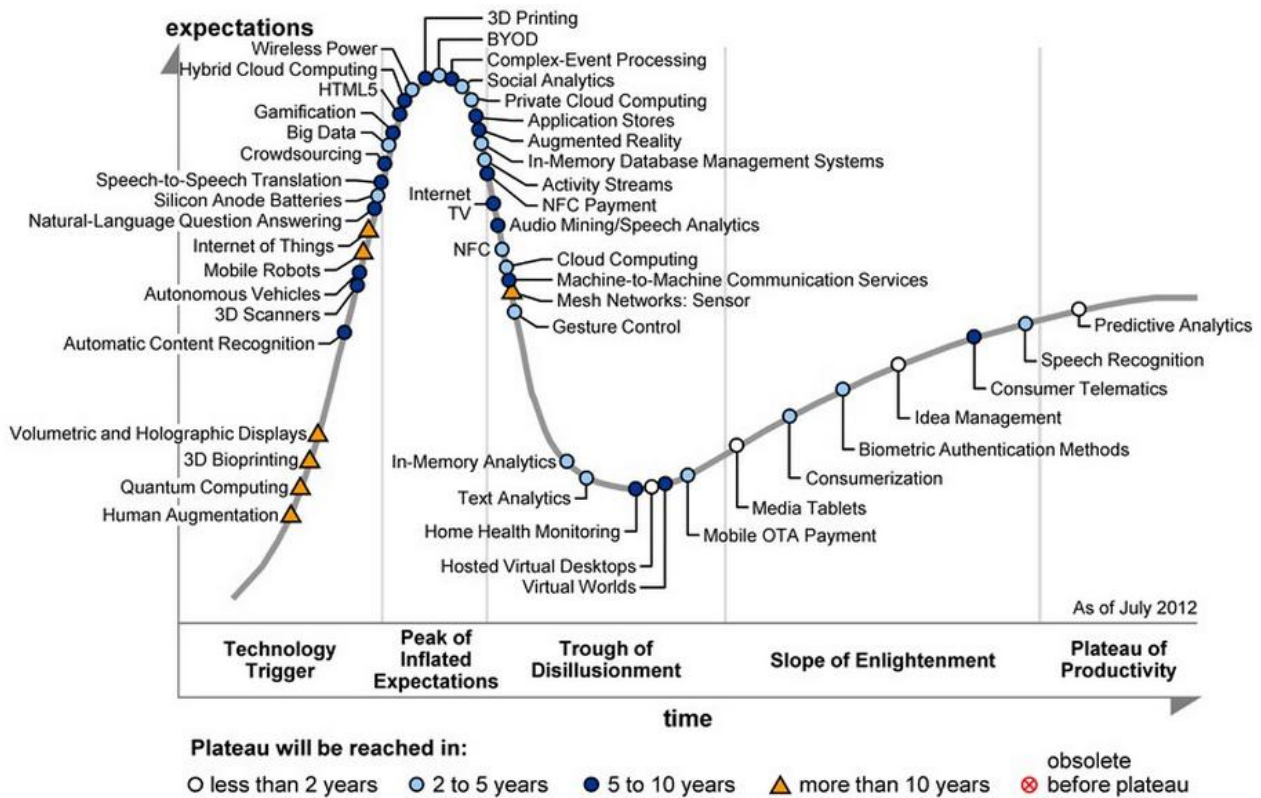


Figura 4 - Gartner 2012: Tecnologias emergentes

Como pode ser visto na Figura 4, a tecnologia de reconhecimento de voz ou *speech recognition* em inglês se encontra no final da fase “*Slope of enlightenment*” e é classificada por uma círculo azul claro indicando que a tecnologia deve atingir o “*Plateau of productivity*” entre os próximos 2 e 5 anos. Logo, é possível que nos próximos anos tenha-se o uso intenso de tecnologia de reconhecimento de voz.

Assim, foi realizada uma busca por referencial teórico com o propósito de caracterizar o estado da arte referente a tecnologias de reconhecimento de voz através da realização de um estudo secundário, a revisão sistemática da literatura.

3.5.2 Protocolo da revisão sistemática da literatura

Objetivo

Realizar uma revisão sistemática da literatura, do tipo mapeamento sistemático, com o propósito de identificar quais os principais estudos e tecnologias de reconhecimento de voz disponíveis atualmente.

Questão de pesquisa

A revisão tem a pretensão de responder as seguintes questões:

- Quais são os principais estudos e tecnologias de reconhecimento automático de voz disponíveis atualmente?

Estratégia de busca

A primeira tarefa na identificação da literatura presente na área é a construção de uma *string* de busca. Desta forma, a *string* de busca definida foi:

("voice recognition" OR " voice translation" OR " voice recognizer" OR " voice recognizers" OR " voice to text" OR " speech recognition" OR " speech translation" OR " speech recognizer" OR " speech recognizers" OR "speech to text" OR " speech to speech" OR "speech-to-speech")

Bases de dados

Inicialmente foi definido que a pesquisa seria executada em diversas bases de dados. Esse protocolo inicial definia que as seguintes bases de dados seriam utilizadas:

- ACM Digital Library
- IEEEExplore
- Willey
- Science Direct – Elsevier
- Springer Link

O primeiro teste da *string* foi realizado na base de dados IEEE e foram retornados 18.959 registros. O motivo para tantos registros serem retornados foi que a *string* de busca está pesquisando não apenas as *keywords* do artigo, mas também o conteúdo existente no resumo. Além disso, algumas das palavras sendo utilizadas aparecem muitas vezes em resumos de artigos mesmo que esse não seja o foco do mesmo. Neste caso, optou-se por restringir a pesquisa apenas por *keywords* através da utilização de comandos específicos.

Como a questão de pesquisa definida era bastante abrangente, resultando em uma enorme quantidade de dados a serem analisados, optou-se por utilizar uma base de dados que indexa publicações de diversas bases de dados. A base de dados SCOPUS permite

pesquisar em um lugar apenas IEEE, ACM, Elsevier, entre outras, e atualmente possui 49 milhões de publicações [SCO13].

A pesquisa na base de dados SCOPUS foi restrita a área “Computer Science” para evitar que artigos de outras áreas que não são do interesse dessa revisão fossem retornados. Foram incluídos também como palavras chaves “*Speech translation*” e “*Speech to speech*” na *string* de busca, pois se acreditava que esses artigos poderiam descrever tecnologias importantes sobre reconhecimento de fala, resultando assim na seguinte *string* de busca:

(“voice recognition” OR “ voice translation” OR “ voice recognizer” OR “ voice recognizers” OR “ voice to text” OR “ speech recognition” OR “ speech translation” OR “ speech recognizer” OR “ speech recognizers” OR “speech to text” OR “ speech to speech” OR “speech-to-speech” OR “speech translation” OR “speech to speech”).

Seleção de estudos

Para permitir a seleção sistematizada de artigos foram definidos nesta seção alguns critérios de inclusão e exclusão de estudos. O objetivo desses critérios é selecionar apenas estudos que empregam tecnologias de reconhecimento de voz para o idioma português, inglês ou espanhol. Dessa forma, removendo um grande número de estudos voltados para os idiomas asiáticos.

A seleção dos estudos foi realizada em duas fases respeitando os critérios de inclusão e exclusão definidos a seguir. Na primeira fase cada artigo era analisado apenas por seu título e resumo. Na segunda fase, os artigos selecionados na fase anterior foram lidos na íntegra e um breve resumo foi criado para cada um dos artigos identificando as principais contribuições. De acordo com Kitchenham, a extração de dados dos artigos deve sempre que possível ser realizada por dois ou mais pesquisadores de forma independente. Essa abordagem permite que pesquisadores discutam resultados divergentes em busca de um consenso. Por limitações de tempo, a extração de dados e análise dos resultados foi feita apenas por um pesquisador.

Idioma dos estudos

O idioma dos estudos deve ser o inglês ou português. A escolha pelo inglês deve-se a grande quantidade de estudos neste idioma.

Critérios para seleção de estudos

Os artigos devem atender os critérios definidos abaixo para serem incluídos na revisão:

- Artigos de conferência e periódicos;
- Estudos primários ou secundários;
- Artigos na área de ciência da computação.
- Artigos disponíveis para download através do convênio da PUCRS com outras instituições.

A decisão por restringir por artigos referentes a área de Ciência da Computação se deve a existência de inúmeras pesquisas em áreas relacionadas a saúde e que utilizam também como palavras chave termos como “*speech recognition*” ou “*voice recognition*”, porém, vinculados a artigos que tratam de problemas relacionados a distúrbios da fala.

Critérios de exclusão de estudos

Para definir um conjunto adequado de critérios de exclusão para essa revisão optou-se por executar algumas consultas prévias a base de dados e analisar seus resultados. Nessas pesquisas preliminares foram encontrados diversos artigos com foco muito diferente do que essa revisão se propõe a esclarecer. Sendo assim, a partir dessas pesquisas preliminares definimos um conjunto de critérios que identificam artigos que não estão relacionados à questão de pesquisa da revisão e, portanto, podem ser removidos dessa revisão. Esses critérios são listados abaixo e qualquer artigo que se enquadre em algum deles será removido da revisão.

- Artigos que não sejam da área de computação ou comunicação.
- Artigos técnicos que focam apenas em melhorar ou alterar algoritmos relacionados a reconhecimento de voz. Alguns exemplos de artigos que se enquadram nesses critérios são aqueles que tratam especificamente de *Hidden markov models* (HMM), *Dynamic time warping* (DTW) ou *Neural networks*.
- Artigos que focam em redução de ruídos.
- Estudos que não sejam em sua totalidade no idioma inglês ou português.
- Artigos curtos.

- Artigos que tratam de distúrbios de fala ou leitura labial.
- Artigos que focam em sistemas de reconhecimento de voz específicos para línguas como: malaio, persa, chinês, japonês, turco, africano, russo, árabe.
- Artigos cujo foco central não é reconhecimento de voz, mas tradução de voz.

Avaliação de qualidade

Os tipos de artigos retornados eram bastante diversos dificultando o estabelecimento de um critério de qualidade geral. Assim, optou-se por analisar todos os arquivos retornados visto que o objetivo era avaliar quais e como estão sendo empregadas as tecnologias de reconhecimento de voz.

Extração dos dados

Após a leitura detalhada de cada artigo foi elaborado um breve resumo de cada um respondendo as seguintes perguntas que são relevantes para a questão de pesquisa:

- Para que o sistema de reconhecimento de voz é utilizado?
- O que é interpretado? Comandos simples ou discursos espontâneos complexos?
- Quais os idiomas reconhecidos pelo sistema?
- Qual o toolkit ou framework utilizado para treinar o modelo de reconhecimento?
- Sistema está disponível livremente para uso?
- De qual projeto o estudo faz parte?

Os dados extraídos foram armazenados em planilha eletrônica.

3.5.3 Resultados

A partir da busca na base Scopus, foram retornados um total de 63 artigos, já aplicando os critérios de inclusão e exclusão definidos anteriormente. A lista de todos os artigos resultantes da execução do protocolo de revisão sistemática é apresentada no ANEXO 1.

A Figura 5 apresenta os artigos por ano de publicação, indicando que o número de estudos envolvendo reconhecimento de voz vem aumentando. O ano de 2013 mostra poucos artigos, pois esta revisão foi realizada no início do ano.

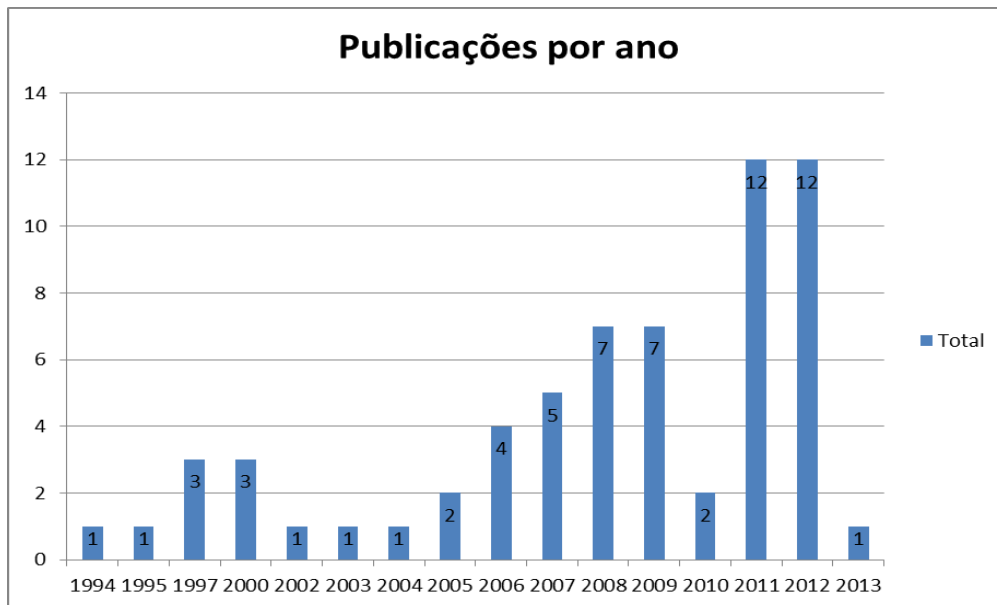


Figura 5 - Publicações por ano

As tecnologias de reconhecimento de voz Microsoft Speech SDK, Sphinx e HTK são as mais utilizadas nos estudos revisados (Figura 6). Essas três tecnologias não tratam apenas de sistemas de reconhecimento de fala, mas sim de plataformas completas para desenvolvimento de aplicações com suporte a reconhecimento de voz.

É importante observar que na Figura 6 foram omitidos em torno de 30 artigos que ou não identificavam explicitamente o sistema de reconhecimento utilizado ou desenvolviam um sistema proprietário/específico para uma determinada função.

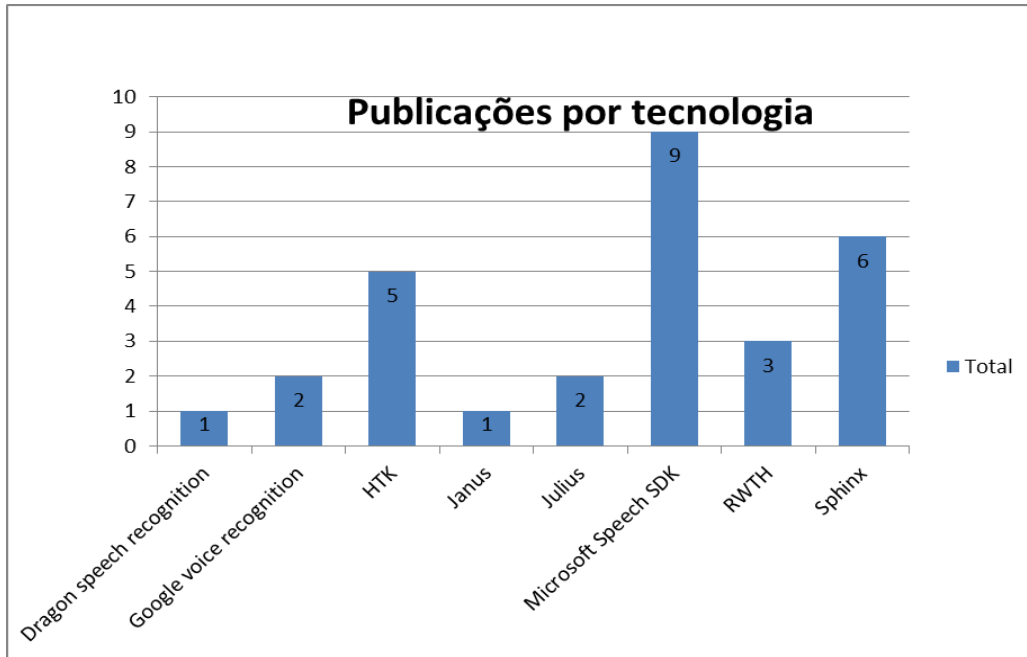


Figura 6 - Publicações por tecnologia

A Figura 7 indica que o número de estudos relacionado a reconhecimento automático de fala para o idioma português ainda é bastante reduzido.

Publicações sobre reconhecimento de fala para o português

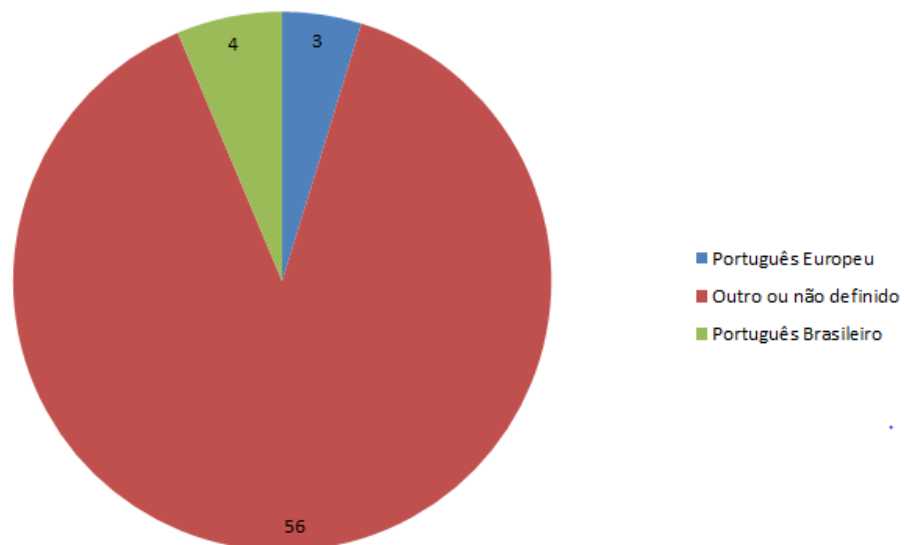


Figura 7 - Publicações no idioma português

A Tabela 1 mostra um resumo das principais aplicações para tecnologias de reconhecimento de voz de acordo com os resultados da revisão sistemática. Conforme a tabela é possível observar um grande número de estudos envolvendo o uso de reconhecimento de voz na área de telecomunicações.

Tabela 1 - Aplicações de tecnologias de reconhecimento de voz

Aplicação	Descrição	Artigos que citam a tecnologia
Automóveis	<ul style="list-style-type: none"> Automação de acesso a dispositivos de automóveis. 	[GIA05], [MIN03], [MIN04]
Telecomunicação	<ul style="list-style-type: none"> Atendimento ao cliente automático. Reconhecimento de voz embutido em ferramenta de mensagem instantânea. Geração automática de legendas em tempo real para programas de televisão. Indexação e transcrição de conteúdos de vídeos e programas de televisão. 	[OLI12], [ATA09], [PER97], [VYS95], [GOR97], [CAM06], [ORT09], [NGU02], [NET08], [NMV11], [PER12]
Sistemas de busca e requisição de informações	<ul style="list-style-type: none"> Pesquisa em bases de dados por voz. Requisições e consultas a sistemas por voz. 	[BAC08], [YU12], [SET94]
Comunicação entre humano e computador	<ul style="list-style-type: none"> Integração de comandos de voz a sistemas operacionais. Integração de comandos de voz a aplicações. Integração de reconhecimento de voz a aplicações de dispositivos móveis como celular. 	[YU12], [LAU97], [CHA11], [COH08]
Educação	<ul style="list-style-type: none"> Auxílio à aprendizagem de idiomas. Transcrição de palestras e aulas em ambiente acadêmico. 	[BAI07], [TRA06], [YU12].
Automação da casa	<ul style="list-style-type: none"> Automação de casa permitindo controle de equipamentos e eletrodomésticos por voz. 	[YU12], [MAH09]
Sistemas de radiologia	<ul style="list-style-type: none"> Transcrição de relatórios radiológicos. 	[CHA11], [PEZ08]

A Tabela 2 mostra um resumo das principais tecnologias de reconhecimento de voz encontradas a partir da revisão sistemática. O Microsoft Speech API é utilizado no maior número de estudos, pois oferece bom desempenho para reconhecimento de comandos de voz. Nesse tipo de aplicação uma gramática é definida e restringe quais são os comandos aceitáveis e qual a estrutura do comando. Esse tipo de reconhecimento é bastante importante para aplicações como interação com robôs, controle do sistema operacional,

controle de aplicativos, atendimento automático de clientes, e qualquer outra aplicação que pode se beneficiar de interação através de um conjunto reduzido de palavras.

Já o Sphinx e HTK vão muito mais além e permitem a criação por completo de novos sistemas de reconhecimento de voz tanto para processamento de comandos quanto para processamento de ditado. O fator limitante que, algumas vezes, impede a criação de reconhecedores de voz para certas línguas é a falta de recursos linguísticos para treinar os modelos acústicos e de linguagem. Porém, havendo recursos esses frameworks são capazes de criar sistemas de reconhecimento bastante poderosos.

Tabela 2 - Tecnologias de reconhecimento de voz

Tecnologia	Descrição	Artigos que citam a tecnologia (considerando apenas o que foi retornado pela revisão sistemática)
Microsoft Speech API (SAPI)	<ul style="list-style-type: none"> • Interface COM nativa do sistema operacional Windows. • Permite reconhecimento de fala, processamento de comandos ditos e síntese de voz. • Mais utilizado por desenvolvedores C++. • É considerada uma interface para desenvolvedores avançados devido a sua complexidade. • Distribuído em todas as versões do Windows desde o Windows 98. • Idiomas suportados: inglês americano, inglês britânico, espanhol, francês, alemão, chinês simplificado e tradicional. 	[ODE07], [VAN12], [CAN12], [YUA11], [MAH09], [LAH05], [PRA13], [HAS11], [YUA11]
Microsoft .NET System.Speech	<ul style="list-style-type: none"> • Permite reconhecimento de fala e síntese de voz. • Projetado para permitir acesso fácil e intuitivo a recursos de processamento de voz do sistema operacional. • Utilizada por desenvolvedores .NET. • Suporta os mesmos idiomas que a Microsoft Speech API (API) visto que ambos acessam recursos do sistema operacional. 	[ODE07]
Microsoft Unified Communications API	<ul style="list-style-type: none"> • Permite a criação de aplicações servidoras que incorporam reconhecimento de voz junto com outras tecnologias como voz sobre IP, mensagens instantâneas, teleconferência, chamadas de voz ou vídeo. • Possui forte integração com o Microsoft Lync. • Permite o desenvolvimento de aplicações que funcionam como uma camada intermediária de acesso ao MS Lync. • Suporta mais de 20 diferentes tipos de 	Nenhuma referência direta encontrada nesta revisão sistemática. Porém, inúmeras referências à tecnologia Microsoft Speech API levaram a consulta no site da Microsoft e ao conhecimento sobre essa tecnologia.

	<p>idiomas.</p> <ul style="list-style-type: none"> • Permite processamento de ditado apenas para um pequeno conjunto de idiomas. 	<p>Informações sobre essa tecnologia foram retiradas diretamente do site da Microsoft.</p>
Microsoft Speech Platform for Server	<ul style="list-style-type: none"> • Permite desenvolver aplicações relacionadas à telefonia, aplicações que requerem interação via voz ou suporte a reconhecimento de voz em aplicações servidoras. • Permite o reconhecimento e síntese de voz de 26 idiomas diferentes incluindo o português brasileiro. • Para o português brasileiro o reconhecimento é limitado a comandos definidos em gramática específica. 	<p>Nenhuma referência direta encontrada nesta revisão sistemática. Porém, inúmeras referências à tecnologia Microsoft Speech API levaram a consulta no site da Microsoft e ao conhecimento sobre essa tecnologia.</p> <p>Informações sobre essa tecnologia foram retiradas diretamente do site da Microsoft.</p>
Sphinx 4	<ul style="list-style-type: none"> • Moderno sistema de reconhecimento de voz baseado em cadeias de Markov desenvolvido como open-source em Java. • Tecnologia independente de idioma que permite construir sistema para reconhecer qualquer linguagem. • Permite a criação de modelos acústicos e de linguagem. • Possui componentes dedicados a realizar tarefas específicas relacionadas a reconhecimento de voz. • Inclui uma variedade de atuais técnicas de reconhecimento. Permite o desenvolvimento de novas pesquisas na área. • Idiomas suportados por padrão: inglês, chinês, francês, espanhol, alemão e russo. 	<p>[TSO11], [FRE00], [RAU06], [KNI10]</p>
HTK - Hidden Markov Model Toolkit	<ul style="list-style-type: none"> • Plataforma para criação e manipulação de modelos baseados em cadeias de Markov. • Permite análise da fala, treinamento de HMM, testes e análises de resultados. • Disponibiliza uma série de módulos e ferramentas desenvolvidas na linguagem C. • Desenvolvido para pesquisa na área de reconhecimento automático de fala. • Usado por muitos anos em aplicações comerciais e estudos acadêmicos. 	<p>[NET11], [AME00], [GIA05]</p>
Julius	<ul style="list-style-type: none"> • Decodificador de alto desempenho para reconhecimento de fala contínuo e de grande vocabulário. • Criado para desenvolvedores e pesquisas na área de reconhecimento de voz. • Permite executar decodificação em tempo quase real na maioria dos computadores • Principais algoritmos de decodificação estão completamente implementados. • Projetado de forma modular permitindo independência em relação a modelos estruturais específicos e suportando vários tipos de HHMs. 	<p>[OLI12], [SIL10], [PER12]</p>

	<ul style="list-style-type: none"> • Permite modelos criados em outras ferramentas como HTK, CMU-CAM SLM toolkit. 	
Java Speech API	<ul style="list-style-type: none"> • Especificação de interface de programação para reconhecimento e síntese de voz na linguagem de programação JAVA. • Disponível gratuitamente. • É apenas uma especificação. Qualquer um pode implementá-la. • Existem bem poucas implementação para JSAPI. • FreeTTS é um implementação conhecida para síntese de voz. • IBM Speech For Java é uma implementação conhecida para reconhecimento de voz que requer instalado a aplicação comercial IBM ViaVoice. 	[NET11]* Citada brevemente. Maior parte das informações foi retirada do site da Microsoft.

Por fim, a Tabela 3 apresenta um resumo das principais tecnologias de reconhecimento de voz disponíveis para português brasileiro.

Tabela 3 - Tecnologias de reconhecimento de voz para o português brasileiro

Tecnologia com suporte a português brasileiro	Descrição	Artigos que citam a tecnologia
IBM ViaVoice	<ul style="list-style-type: none"> • Permite o reconhecimento de voz para o idioma português livre de gramática específica. • O usuário necessita treinar o sistema. • Produto comercial descontinuado na versão 9. • Alguns estudos ainda utilizam esse sistema para comparar o desempenho de novas ferramentas e protótipos sendo desenvolvidos para o reconhecimento de voz no português. • Sistema funciona apenas no Windows XP. • Quase nenhuma documentação está disponível visto que o produto foi descontinuado. • Não foram encontradas bibliotecas disponíveis para permitir a integração entre o IBM ViaVoice 	[NET11]

	9 e novas aplicações que possam fazer uso de reconhecimento de voz.	
Coruja	<ul style="list-style-type: none"> • Projetado para permitir reconhecimento de ditado e comandos para o português brasileiro. • Não requer treinamento do usuário. • Aplicação ainda em desenvolvimento. • Criado pelo grupo FalaBrasil. • Atualmente na versão 0.2. • De acordo com um dos pesquisadores do grupo, a aplicação ainda não apresenta resultados suficientes para uso em aplicações que necessitam reconhecimento espontâneo livre de gramática específica. • Utiliza diversos modelos acústicos e de linguagem criados pelo grupo. 	<p>Nenhuma referência direta encontrada nesta revisão sistemática. Porém, o artigo [NET11] fazia referência a estudos desenvolvidos pelo grupo de pesquisas FalaBrasil. Visita ao site do grupo levou ao conhecimento desta tecnologia.</p> <p>Informações sobre essa tecnologia foram retiradas diretamente do site do FalaBrasil e de contatos com o grupo de pesquisa.</p>
Microsoft Speech Platform for Servers	<ul style="list-style-type: none"> • Permite reconhecimento de comandos de gramática específica para o português brasileiro. • Esse framework é voltado para o desenvolvimento de aplicações de atendimento automático de ligações de usuários com suporte a reconhecimento de voz. • Apresenta bom desempenho para reconhecimento de comandos. 	<p>Nenhuma referência direta encontrada nesta revisão sistemática. Porém, inúmeras referências à tecnologia Microsoft Speech API levaram a consulta no site da Microsoft e ao conhecimento sobre essa tecnologia.</p>

De acordo com a revisão, apenas o IBM ViaVoice e o Coruja permitem o reconhecimento de voz para o português brasileiro. No entanto, o IBM ViaVoice é um produto comercial que foi descontinuado e não oferece suporte ou possibilidade de integração de suas funcionalidades com novas aplicações sendo desenvolvidas. Em contrapartida, o Coruja é um próspero e recente sistema de reconhecimento de voz também para o idioma português brasileiro. De acordo com os pesquisadores do grupo, o Coruja hoje apresenta um desempenho bom para reconhecimento de comandos de gramática específica, porém não é robusto ainda para executar o reconhecimento de ditado com desempenho aceitável. No Capítulo 4 serão avaliadas algumas das tecnologias encontradas para o idioma português brasileiro.

4 AVALIAÇÃO DAS TECNOLOGIAS PARA O PORTUGUÊS

Esse capítulo avalia as principais tecnologias identificadas durante a revisão sistemática da literatura que suportam o reconhecimento de voz no idioma português brasileiro. Essa avaliação tem a intenção de verificar o desempenho a fim de avaliar a possibilidade de utilizar essas tecnologias de reconhecimento de voz como componentes para a construção de um tradutor de voz para o português brasileiro. Nesta seção são avaliadas as seguintes tecnologias: Microsoft Speech Platform For Server, IBM ViaVoice e o Coruja.

4.1 Microsoft Speech Platform for Servers

O Microsoft Speech Platform for Servers foi a primeira tecnologia a ser avaliada, visto que de acordo com o site da Microsoft a tecnologia suportaria o reconhecimento de voz para 26 idiomas incluindo o português brasileiro. Como o suporte a ditados existia para todos os idiomas suportados pelo Microsoft Speech API também se acreditava que o mesmo era válido para o MS Speech Platform. Essa tecnologia seria o que a Microsoft teria de mais promissor e moderno.

Como é pré-requisito da tecnologia a instalação em um sistema operacional servidor, foi instalado o Windows Server 2008 em uma máquina virtual, e todos os componentes necessários que compõem o Microsoft Speech Platform for Servers. O ambiente de desenvolvimento .NET também foi instalado para permitir o desenvolvimento da aplicação de reconhecimento voz.

Após o início do desenvolvimento da aplicação se identificou que era possível apenas implementar o reconhecimento de comandos mas não o de ditado livre para o português brasileiro. Após investigações em fóruns especializados isto se confirmou. Mais tarde, foi possível encontrar uma pequena nota dentro da documentação especificando que o suporte de ditado não é implementado para o português brasileiro assim como para muitos outros idiomas. Por este motivo, foi descartado o uso do Microsoft Speech Platform for Servers. A Tabela 4 a seguir indica as aplicações e bibliotecas utilizadas na instalação e avaliação da tecnologia Microsoft Speech Platform for Servers.

Tabela 4 - Instalação realizada para o Microsoft Speech Platform for Servers

Instalação	Descrição
Windows Server 2008	Pré-requisito para o uso do Speech Platform.
SpeechPlatformRuntime x64, MicrosoftSpeechPlatformSDK x64	Instalação do Microsoft Speech Platform for Servers.
MSSpeech_SR_pt-BR_TELE	Adiciona suporte ao português brasileiro.
MSSpeech_SR_en-US_TELE	Adiciona suporte ao inglês americano.
Microsoft Visual Studio 2010	Ambiente de desenvolvimento.

4.2 IBM ViaVoice

O IBM ViaVoice foi a segunda tecnologia a ser avaliada. Embora descontinuado, ainda existiam versões disponíveis da aplicação na internet. Foram realizadas tentativas de instalação no Windows 8 e 7, porém sem sucesso. Foi necessário a instalação do ViaVoice em uma máquina virtual rodando sistema operacional Windows XP. O IBM ViaVoice foi avaliado de acordo com um protocolo que pode ser consultado no ANEXO 2 . O resultado da execução deste protocolo é apresentado no ANEXO 3.

4.2.1 Análise dos resultados

O tempo de reconhecimento é em torno de 1.7 segundos desconsiderando o tempo de pronuncia da frase conforme mostrado na Tabela 5. Frases contendo estrutura mais complexa e maiores levam mais tempo para serem reconhecidas conforme indicado pela Tabela 6.

Tabela 5 - IBM ViaVoice: Tempo de reconhecimento

Média de tempo do reconhecimento (segundos)
1,714285714

Tabela 6 - IBM ViaVoice: Tempo de reconhecimento por tipo de vocabulário

Vocabulário	Média (segundos)
Complexo	1,882352941
Simple	1,454545455
Total	1,714285714

A Tabela 7 mostra que 46% das frases foram reconhecidas com máxima precisão. Se for considerado que os níveis de precisão “completamente preciso” e “bastante preciso” são considerados aceitáveis, neste caso então 64% das frases reconhecidas foram classificadas como aceitáveis conforme Tabela 8.

Tabela 7 - IBM ViaVoice: Qualidade do reconhecimento

Avaliação	Quantidade	Percentual
Completamente preciso	13	46,43%
Bastante preciso	5	17,86%
Um pouco impreciso	4	14,29%
Completamente impreciso	6	21,43%
Total	28	100,00%

Tabela 8: IBM ViaVoice: Qualidade do reconhecimento - Aceitável x Não Aceitável

Avaliação secundária	Quantidade	Percentual
Aceitável	18	64,29%
Não aceitável	10	35,71%
Total	28	100,00%

A Tabela 9, no entanto, mostra que a precisão do reconhecimento cai consideravelmente dependendo da complexidade das frases:

- 81% das frases classificadas como simples tiveram precisão de reconhecimento considerados aceitáveis;
- Por outro lado, apenas 53% das frases classificadas como complexas tiveram precisão de reconhecimento considerados aceitáveis.

Tabela 9 - IBM ViaVoice: Qualidade do reconhecimento por vocabulário – Aceitável x Não Aceitável

Aceitação por vocabulário	Quantidade	Percentual
Complexo	17	60,71%
Aceitável	9	52,94%
Não aceitável	8	47,06%
Simples	11	39,29%
Aceitável	9	81,82%
Não aceitável	2	18,18%
Total	28	100,00%

Conforme indicado na Tabela 10, em média o reconhecimento do IBM ViaVoice é avaliado como 2,89 (mediana é 3) com maior parte do reconhecimento produzido (64,29%) caindo nas categorias “completamente preciso” ou “bastante preciso”.

Tabela 10 - IBM ViaVoice: Avaliação geral – Média e Mediana

Média	Mediana
2,892857	3

A Tabela 11 mostra a média e mediana se for considerado o tipo de vocabulário. Neste caso, nota-se que o reconhecimento produzido é inferior quando o vocabulário é complexo. Sendo apenas 52,94% classificado como “completamente preciso” ou “bastante preciso”

Tabela 11 - Avaliação geral por vocabulário - Média e Mediana

Vocabulário	Média	Mediana
Simples	3,272727	4
Complexo	2,647059	3

Além disso, o IBM ViaVoice é um produto descontinuado e que não permite a integração com outras aplicações via interfaces de programação (API).

Em resumo, o sistema apresenta desempenho insatisfatório para discurso espontâneo e contendo vocabulário complexo, especialmente no quesito tempo de processamento, visto que o ViaVoice não permite reconhecimento de voz em tempo real ou próximo disto. Após ditar uma frase inteira ainda é necessário aguardar alguns segundos para ver a frase sendo transcrita na tela.

Na avaliação realizada foram descartadas algumas tentativas de reconhecimento de certas frases, pois o participante claramente não expressou corretamente algumas palavras levando ao reconhecimento totalmente incorreto das frases. No experimento realizado ficou clara a importância das palavras serem claramente pronunciadas, do contrário, o reconhecimento não é realizado de forma correta.

O resultado abaixo do esperado do ViaVoice para reconhecimento espontâneo, especialmente para frase complexas, associado a lentidão do processamento do que foi dito não permite que o ViaVoice seja utilizado como reconhecedor de voz para o desenvolvimento de um sistema de tradução de voz automática que tem como objetivo ser utilizado para comunicação em tempo real.

4.3 Coruja

Outro sistema de reconhecimento de voz avaliado foi o Coruja. Esse sistema está sendo desenvolvido pelo grupo de pesquisas FalaBrasil e suportaria reconhecimento de comandos definidos por gramática específica e também reconhecimento de ditado para o português brasileiro. Atualmente, o Coruja está ainda em fase de desenvolvimento.

No site no grupo de pesquisas FalaBrasil [FAL14] está disponível um projeto completo que utiliza o Coruja para permitir controle de voz a arquivos de apresentação powerpoint. O objetivo dessa aplicação é permitir que usuários abram um arquivo powerpoint e controlem por voz as principais funções como colocar em modo de apresentação, avançar, voltar e sair de modo de apresentação.

Para teste dessa aplicação foi necessária a instalação do Microsoft Visual Studio 2010. O projeto PPTController foi baixado do site FalaBrasil junto com uma versão completa do Coruja. Algumas configurações foram necessárias no projeto para apontar corretamente para o modelo acústico. Depois disto, foi possível realizar os primeiros testes utilizando o reconhecimento de comandos do Coruja.

O projeto PPTController tinha definido na sua gramática em torno de 6 comandos. O reconhecimento de voz é livre de usuário e não requer treinamento individual de cada usuário. Nos testes realizados utilizando a aplicação PPTController o coruja apresentou um desempenho muito inferior ao esperado

errando múltiplas vezes apesar de a gramática definida ser bem pequena. Em alguns momentos, simplesmente nada era reconhecido. O tempo de reconhecimento de cada comando foi superior se comparado aos testes realizados com o ViaVoice.

Após esse teste inicial, iniciou-se a participação no fórum de discussão do grupo de pesquisas FalaBrasil com o objetivo de obter um outro projeto pronto que tivesse já implementado reconhecimento de voz com suporte a ditado. Após algumas semanas sem respostas, foi possível fazer contato direto com os membros do grupo de pesquisa.

Em conversa direta com os pesquisadores do grupo FalaBrasil, foi obtido um vídeo que mostrava o Coruja sendo utilizado para reconhecimento de ditado no idioma português. O vídeo demonstrava o que mais tarde foi afirmado pelos pesquisadores do grupo: o coruja ainda não era robusto o suficiente para permitir bom reconhecimento de ditado. O reconhecimento de ditado apresentava alto número de erros no reconhecimento, além de ser mais lento que o IBM ViaVoice. A partir disto, ficou claro que o Coruja ainda não pode ser utilizado como mecanismo de reconhecimento de voz para o desenvolvimento de uma máquina de tradução de voz.

4.4 Avaliação geral das tecnologias

Em resumo, como parte desta revisão sistemática não foram encontradas tecnologias que permitam o reconhecimento de ditado para o português brasileiro com desempenho suficiente que permita a sua integração em um sistema de tradução de voz.

A dificuldade em encontrar tecnologias para o reconhecimento do português se deve a escassez de recursos acústicos e linguísticos para esse idioma. Para o português brasileiro ainda faltam bases de dados de áudio compartilhadas, visto que apenas alguns esforços individuais de centros de pesquisa são encontrados [DAS12]. A disponibilidade de bases de dados de áudio é determinante para o desenvolvimento de sistemas de reconhecimento de fala. O compartilhamento de bases de dados de áudio entre pesquisadores europeus e dos Estados Unidos é a razão do maior progresso obtido em seus respectivos países nas últimas décadas. Nesse caso, para existir exemplos

suficientes de forma a permitir métodos estatísticos trabalharem apropriadamente, as bases de dados de áudio precisam ser representativas e, portanto, difíceis de construir tanto em tempo quanto em investimentos financeiros [DAS12].

4.5 Outras tecnologias

Após não ter sido encontrada tecnologia de reconhecimento de voz com desempenho adequado para reconhecimento do português brasileiro durante a revisão sistemática da literatura, optou-se pela realização de uma pesquisa livre (ad hoc) e não estruturada na ferramenta de busca do Google visando encontrar tecnologias que suportassem o português brasileiro e que não tenham sido retornadas pela revisão sistemática. O objetivo desta pesquisa foi complementar o levantamento realizado anteriormente incluindo também tecnologias comerciais que ainda não foram avaliadas em estudos acadêmicos.

Como resultado dessa pesquisa foi encontrada uma nova tecnologia de reconhecimento de voz. Essa tecnologia consiste em uma nova API de reconhecimento de voz (*Web Speech API*) desenvolvida pelo Google e que permite aos desenvolvedores integrarem reconhecimento de voz em suas aplicações web [CHR01]. A tecnologia é bastante recente tendo sido lançada em fevereiro de 2013 junto da versão 25 no navegador Chrome.

Essa nova API suporta o reconhecimento de fala contínuo em diversos idiomas, incluindo o português brasileiro. Entretanto, tal tecnologia possui restrições visto que pode ser utilizada apenas em aplicações web dentro do navegador Chrome [CHR02]. Essa tecnologia é bastante promissora e estudos adicionais são necessários para avaliar seu desempenho e possibilidade de integração a um sistema de tradução de voz. Neste sentido, o próximo capítulo visa avaliar e explorar essa nova tecnologia.

5 AVALIAÇÃO DE TECNOLOGIA – SIMULAÇÃO

Para a construção de um tradutor de voz é necessário encontrar componentes para reconhecimento de voz, tradução de texto e síntese. A tecnologia do Google Web Speech API tem potencial para ser utilizada como componente de reconhecimento de voz, porém ainda é recente para utilizar essa tecnologia na construção de um protótipo de tradução de voz, visto que até o momento não foi encontrado nenhum estudo avaliando o desempenho da tecnologia. O Google também disponibiliza uma tecnologia para tradução de texto chamada de Google Translator. Essa tecnologia já vem sendo utilizada em estudos acadêmicos e produz bons resultados [CAL10, CAL11, CAL12].

Antes de realizar a construção de um protótipo de tradutor de voz envolvendo essas duas tecnologias é importante a execução de uma simulação que permita avaliar o desempenho dessas tecnologias. Em [CAL10] é apresentada uma simulação realizada para avaliar duas tecnologias diferentes de tradução de texto. Devido a similaridades entre os estudos, a simulação realizada como parte desse trabalho é baseada em [CAL10], porém neste caso são avaliados a qualidade do reconhecimento de voz e da tradução ao invés de comparar duas tecnologias de tradução. As próximas seções mostram o planejamento dessa simulação e os resultados encontrados.

5.1 Planejamento

Essa seção apresenta o plano sobre como será realizada essa simulação. Os objetivos da simulação são listados a seguir:

- 1) Avaliar a qualidade do reconhecimento de voz da tecnologia Google Web Speech API para o idioma português brasileiro.
- 2) Avaliar a qualidade da tradução para o inglês produzida pelo Google Translator se recebido como entrada o texto, em português brasileiro, resultante do reconhecimento de voz realizado pelo Google Web Speech API.

A avaliação da qualidade de tradução ou reconhecimento de voz é uma tarefa subjetiva e são muitas as discordâncias em relação a metodologias, especialmente no caso da tradução [CAL10].

No entanto, avaliação é essencial. Nesta simulação, cada participante é também responsável por avaliar o resultado do reconhecimento ou tradução. Em nossa simulação cada avaliador da tecnologia de reconhecimento de voz, analisa o quão adequado é o resultado da transcrição em relação ao texto que foi originalmente dito. O sistema de avaliação adotado é a escala de 4 pontos Likert. Essa escala foi utilizada em [CAL10] e parece apropriada para nosso objetivo devido aos seguintes motivos: (a) não possui muitas opções de valores; (b) pode ser aplicada visto que as descrições são claras; (c) não possui um valor intermediário. A não existência de valores intermediários proporciona melhores resultados, pois não é possível atribuir um valor neutro a avaliação.

A simulação foi realizada com a participação de quatro estudantes do curso de Mestrado em Ciência da Computação. Como parte do experimento, os participantes foram solicitados a realizar a leitura de 20 frases no idioma português. Cada frase foi transcrita automaticamente pela tecnologia de reconhecimento de voz do Google. Após a transcrição, cada frase foi automaticamente traduzida para o idioma inglês pela tecnologia de tradução do Google. Após a execução da transcrição e da tradução, cada participante foi solicitado a avaliar o nível de precisão da transcrição e tradução realizadas, permitindo assim avaliar o nível de precisão dessas tecnologias. O protocolo detalhado da simulação é apresentado no ANEXO 4 e o resultado da execução é apresentado no ANEXO 5.

5.2 Análise dos resultados

A simulação realizada pode ser dividida em duas etapas: reconhecimento da voz e tradução do texto. Sendo assim, a análise dos resultados também será realizada de forma separada.

5.3 Qualidade do reconhecimento de voz

O tempo de simulação não foi coletado durante a simulação devido a limitações técnicas da API, porém observou-se que o tempo de reconhecimento do Web Speech API é próximo do real. Conforme os participantes falam a frase, a tecnologia já reconhece quase que instantaneamente.

A Tabela 12 mostra que 40% das frases foram reconhecidas com máxima precisão e 17,5% foram considerados completamente imprecisos não sendo possível extrair qualquer informação a respeito do que foi reconhecido.

Similarmente ao estudo realizado em [CAL10], será considerado nível de precisão aceitável se a qualidade do reconhecimento for classificada como “completamente preciso” ou “bastante preciso”. De acordo com essa classificação, então 65% das frases reconhecidas tiveram nível de qualidade considerado aceitável conforme mostrado na Tabela 13.

Tabela 12 – Google Web Speech API: Qualidade do reconhecimento

Avaliação	Quantidade de frases reconhecidas	Percentual
Completamente preciso	32	40,00%
Bastante preciso	20	25,00%
Um pouco imprecise	14	17,50%
Completamente imprecise	14	17,50%
Total	80	100,00%

Tabela 13 – Google Web Speech API: Qualidade do reconhecimento – Aceitável X Não Aceitável

Avaliação secundária	Quantidade de frases reconhecidas	Percentual
Aceitável	52	65,00%
Não aceitável	28	35,00%
Total	80	100,00%

Similarmente aos resultados do estudo realizado com o IBM ViaVoice, vemos na Tabela 14 que a precisão do reconhecimento cai consideravelmente dependendo da complexidade das frases:

- 75% das frases classificadas como simples tiveram precisão de reconhecimento considerados aceitável;

- Por outro lado, apenas 55% das frases classificadas como complexas tiveram precisão de reconhecimento considerados aceitável.

Tabela 14 – Google Web Speech API: Qualidade do reconhecimento por vocabulário – Aceitável X Não Aceitável

Aceitação por vocabulário	Quantidade de frases reconhecidas	Percentual
Complexo	40	50,00%
Aceitável	22	55,00%
Não aceitável	18	45,00%
Simples	40	50,00%
Aceitável	30	75,00%
Não aceitável	10	25,00%
Total	80	100,00%

Conforme indicado na Figura 8, em média o reconhecimento do Google Web Speech API é avaliado como 2,85 (mediana é 3) com maior parte do reconhecimento produzido (65%) caindo nas categorias “completamente preciso” ou “bastante preciso”.

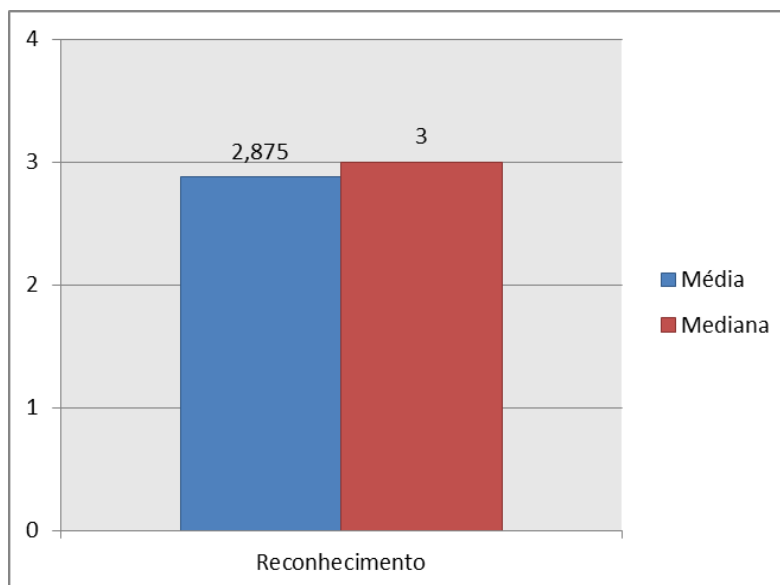


Figura 8 – Google Web Speech API: Avaliação reconhecimento - Média e Mediana

A Figura 9 apresenta a média e mediana se for considerado o tipo de vocabulário.

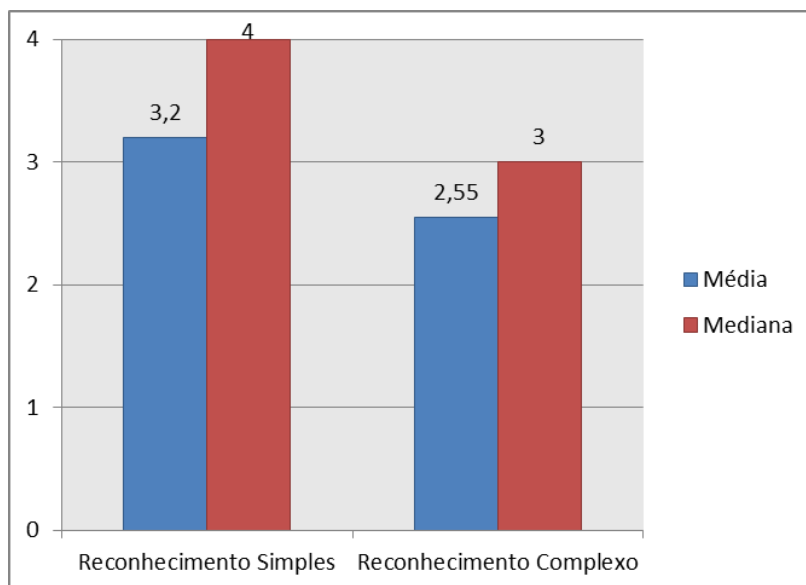


Figura 9 – Google Web Speech API: Avaliação reconhecimento por vocabulário - Média e Mediana

O reconhecimento de voz para vocabulário simples tem média e mediana situada na faixa de valores entre 3 e 4 correspondente a classificação “completamente preciso”. Entretanto, o nível de precisão diminui quando a complexidade do vocabulário aumenta. Para vocabulário complexo, a média e mediana estão situadas entre 2 e 3 que corresponde a classificação “bastante precisa”. Para vocabulário complexo apenas 55% das frases foram classificadas como “completamente preciso” ou “bastante preciso”

5.4 Qualidade da tradução

O tempo de simulação não foi coletado durante a simulação, porém observou-se que o tempo de tradução do Google Translator é próximo do real. A Tabela 15 Tabela 12 apresenta que 32,5% das frases foram traduzidas com máxima precisão e 21,25% foram consideradas completamente imprecisas não sendo possível extrair qualquer informação a respeito do que foi traduzido.

Similarmente ao estudo realizado em [CAL10], será considerado nível de precisão aceitável se a qualidade da tradução for classificada como “completamente precisa” ou “bastante precisa”. De acordo com essa classificação, então 58,75% das frases traduzidas tiveram nível de qualidade considerado aceitável conforme mostrado na Tabela 16.

Tabela 15 – Google Translator: Qualidade da tradução

Avaliação	Quantidade de traduces	Percentual
-----------	------------------------	------------

Completamente precisa	26	32,50%
Bastante precisa	21	26,25%
Um pouco imprecisa	16	20,00%
Completamente imprecisa	17	21,25%
Total	80	100,00%

Tabela 16 – Google Translator: Qualidade da tradução – Aceitável X Não Aceitável

Avaliação secundária	Quantidade de traduces	Percentual
Aceitável	47	58,75%
Não aceitável	33	41,25%
Total	80	100,00%

A Tabela 17 mostra que a precisão da tradução cai dependendo da complexidade das frases. Como nesta simulação é utilizado como entrada para a tradução, os valores de saída do reconhecimento de voz, esse comportamento é provavelmente explicado pela pior qualidade do reconhecimento de voz para vocabulário complexo.

- 62,5% das frases classificadas como simples tiveram precisão de tradução considerada aceitável;
- Por outro lado, apenas 55% das frases classificadas como complexas tiveram precisão de tradução considerada aceitável.

Tabela 17 – Google Translator: Qualidade da tradução por vocabulário – Aceitável X Não Aceitável

Aceitação por complexidade	Quantidade de traduções	Percentual
complexo	40	50,00%
Aceitável	22	55,00%
Não aceitável	18	45,00%
Simple	40	50,00%
Aceitável	25	62,50%
Não aceitável	15	37,50%
Total	80	100,00%

Conforme indicado na Figura 10, em média a tradução do Google Translator é avaliada como 2,7 (mediana é 3) com maior parte da tradução produzida (58,75%) caindo nas categorias “completamente precisa” ou “bastante precisa”.

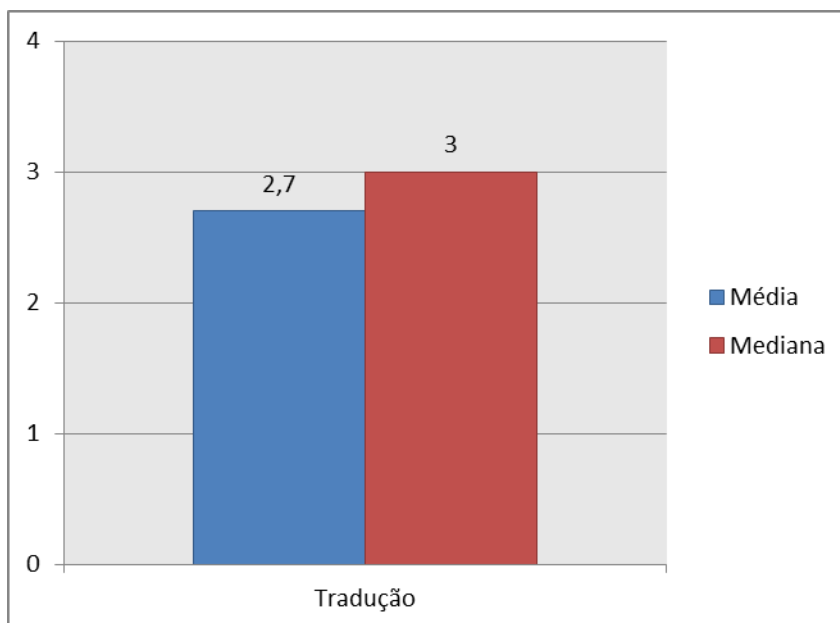


Figura 10 – Google Translator: Avaliação da tradução - Média e Mediana

A Figura 11 apresenta a média e mediana se for considerado o tipo de vocabulário.

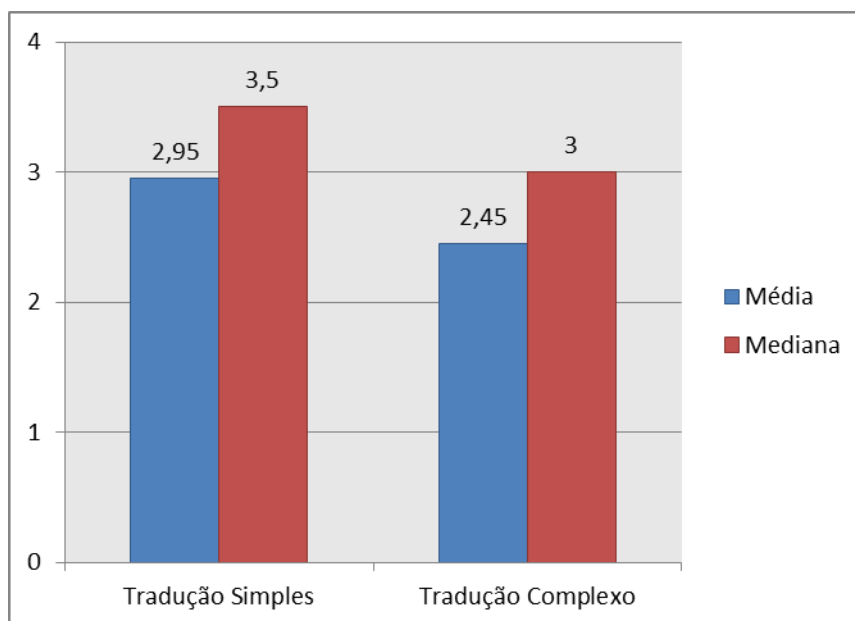


Figura 11 – Google Translator: Avaliação da tradução por vocabulário - Média e Mediana

A tradução para vocabulário simples tem média e mediana situada na faixa de valores entre 2 e 3, quase dentro da faixa entre 3 e 4 caracterizando que a precisão da tradução para frases simples se encontra entre as classificações “bastante preciso” e “completamente preciso”. O nível de precisão diminui quando

a complexidade do vocabulário aumenta. Para vocabulário complexo, a média e mediana estão situadas entre 2 e 3 que corresponde a classificação “bastante precisa”.

5.5 Análise da qualidade geral

A Figura 12 compara as informações de média e mediana obtidas durante a avaliação de qualidade de reconhecimento e tradução. A simulação demonstrou que ambas as tecnologias mesmo quando usadas de forma integrada produzem resultados “bastante precisos” (faixa entre 2 e 3), o que é uma precisão considerada adequada.

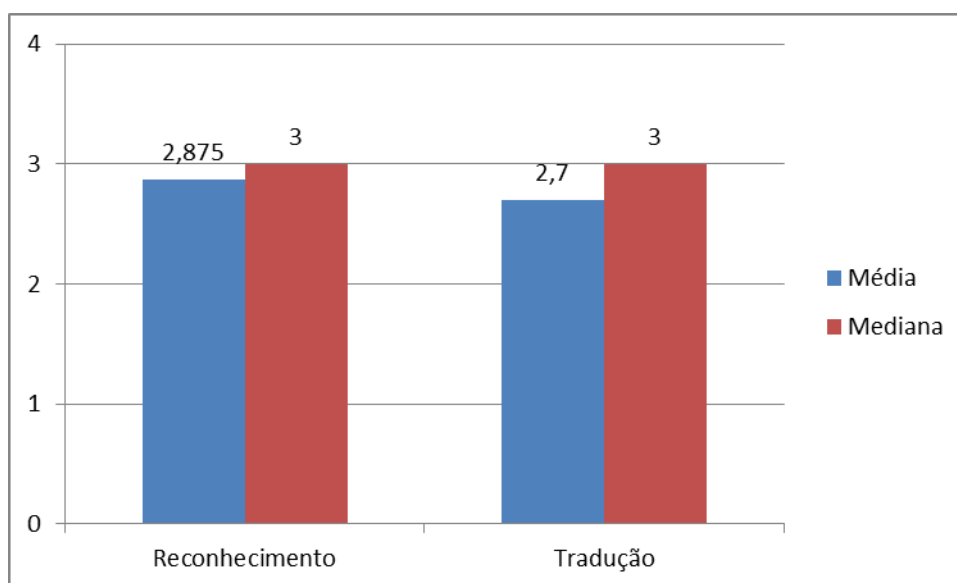


Figura 12 - Google: Comparativo reconhecimento e tradução

A Figura 13 demonstra que a qualidade do reconhecimento e tradução é alta quando vocabulário classificado como simples é processado. A Figura 14 demonstra que o desempenho das tecnologias diminui quando o vocabulário é complexo, porém mesmo assim ainda apresenta precisão classificada como aceitável permanecendo na mesma faixa de classificação entre 2 e 3.

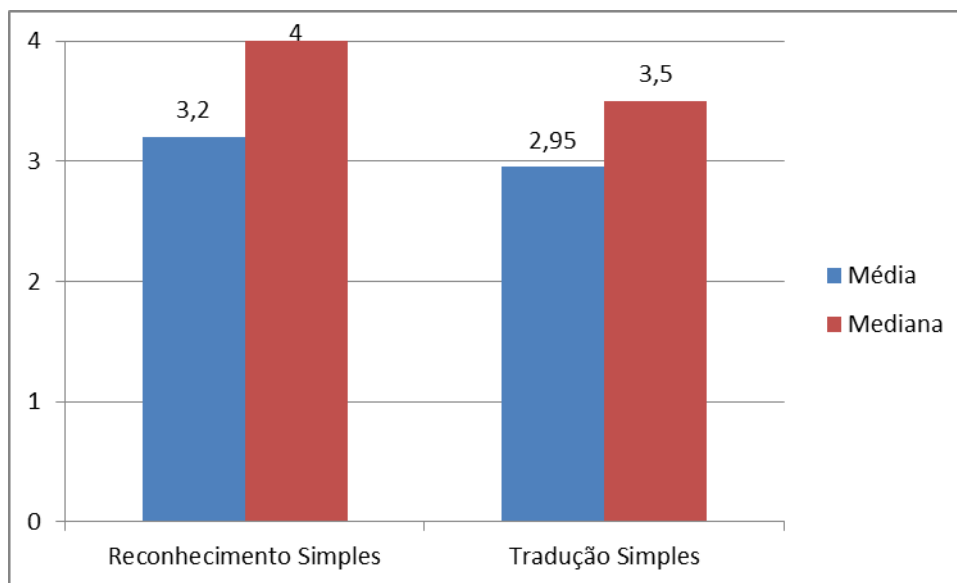


Figura 13 - Comparativo reconhecimento e tradução para vocabulário simples

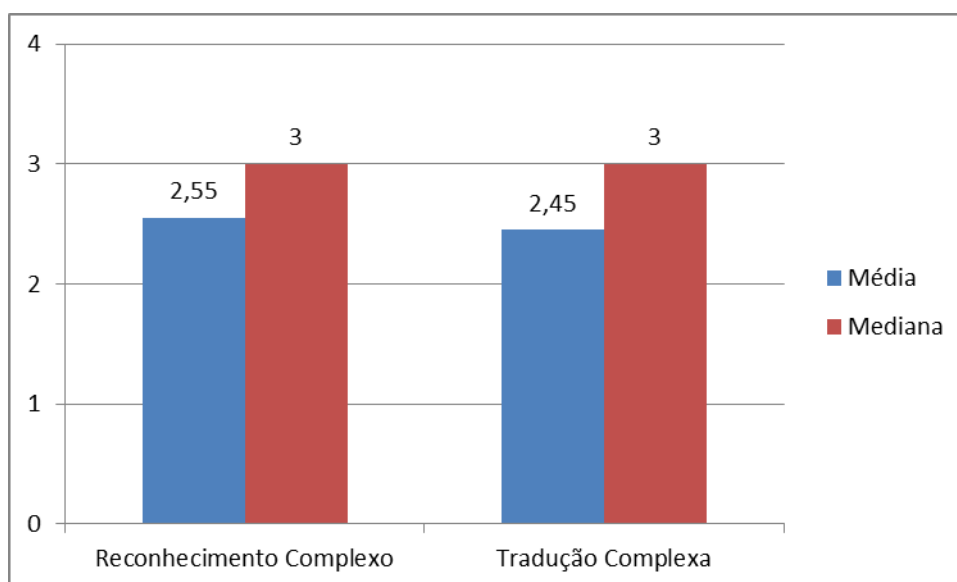


Figura 14 - Comparativo reconhecimento e tradução para vocabulário complexo

De acordo com os dados apresentados, temos indícios de que as tecnologias do Google para reconhecimento de voz no idioma português e para tradução do português para o inglês possuem níveis adequados de precisão e poderiam ser utilizadas no desenvolvimento de uma máquina de tradução de voz para o idioma português.

6 CONCLUSÕES

Essa pesquisa é parte de um projeto em parceria entre pesquisadores da Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS) e da Universidade de Bari na Itália, financiado pela Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS). O foco deste projeto está em realizar pesquisas científicas que estudam a viabilidade do uso de máquinas de tradução em projetos de DDS, em situações onde os membros da equipe não dominam um idioma comum. O objetivo desta pesquisa foi o de contribuir com este projeto, investigando especificamente como máquinas de tradução de voz poderiam ser utilizadas em projetos de DDS.

6.1 Contribuições da pesquisa

A pesquisa iniciou com uma revisão de literatura (estudos secundários) com o objetivo de entender quais eram as tecnologias que envolvem tradução de voz e se essas tecnologias permitiam uso com o português brasileiro. Foi avaliada também a existência de estudos envolvendo máquinas de tradução de voz e DDS.

O levantamento realizado não encontrou pesquisas envolvendo máquinas de tradução de voz e DDS. Além disso, constatou-se que existem poucas tecnologias disponíveis para reconhecimento de voz no idioma português brasileiro, e isso se deve a escassez de material para treinar os modelos acústicos e linguísticos nesse idioma [DAS12].

Para reconhecimento de voz no idioma português brasileiro foram encontradas três tecnologias e que foram avaliadas nesta pesquisa (avaliação para elaboração de proposta inicial): (1) Coruja, (2) IBM ViaVoice e (3) Google Web Speech API. O coruja e o IBM ViaVoice demonstraram não ser adequados para o uso em um sistema de tradução de voz. Já o Google Web Speech API, embora seja uma tecnologia bastante recente tendo sido lançada em fevereiro de 2013, produziu resultados considerados aceitáveis pelos avaliadores da tecnologia durante a simulação realizada.

Essa simulação (estudo de viabilidade) levantou indícios de que a tecnologia de reconhecimento de voz e tradução do Google apresenta bom

desempenho para o idioma português brasileiro e poderia ser utilizada para o desenvolvimento de uma máquina de tradução de voz que permitisse a comunicação em tempo real entre equipes de desenvolvimento de software distribuídas que não possuem idioma em comum.

A questão de pesquisa definida originalmente visava identificar como máquinas de tradução de voz poderiam ser utilizadas em projetos de desenvolvimento distribuído de software. Devido a escassez de pesquisas envolvendo máquinas de tradução para o português brasileiro e DDS, optou-se por focar em avaliar a viabilidade da criação de uma máquina de tradução para esse idioma através da união das tecnologias de reconhecimento de voz e tradução existentes não sendo possível responder por completo a questão de pesquisa originalmente proposta.

6.2 Limitações desta pesquisa

As simulações realizadas possuem uma série de limitações e seria necessária a realização de estudos futuros para confirmar a viabilidade do uso dessas novas tecnologias na composição de uma máquina de tradução de voz com precisão suficiente para permitir adequada comunicação por voz. Algumas destas limitações são listadas a seguir.

- 1) Limitações relacionadas ao experimento de simulação de máquinas de tradução de voz utilizando tecnologia Google:
 - O experimento contou com apenas quatro participantes. Um estudo com um número maior de participantes poderia agregar maior solidez aos resultados.
 - Cada participante foi responsável por avaliar o reconhecimento e tradução das frases que ditou. Isto poderia acrescentar algum viés no momento da avaliação dos resultados.
 - Cada participante foi responsável por avaliar a tradução das frases que ditou. A diferença de nível de inglês dos participantes pode ter gerado discrepâncias na avaliação da precisão das traduções.
 - As frases utilizadas no experimento foram classificadas pelo autor previamente como simples ou complexas. Essa classificação se deu com

base no tamanho das frases e vocabulário utilizado. Porém, não foi estabelecido um critério bem definido para diferenciar simples de complexo, podendo gerar divergências nos resultados que utilizaram tais classificações.

- A simulação envolveu leitura de frases, não ocorrendo conversa entre uma ou mais pessoas e isto pode ter facilitado a obtenção de melhores índices de precisão.

2) Limitações relacionadas ao mapeamento sistemático realizado para reconhecimento de voz:

- Estudos foram selecionados utilizando apenas a base de dados SCOPUS. Embora a cobertura da SCOPUS seja bastante abrangente é possível que estudos importantes não tenham sido retornados.
- A questão de pesquisa definida é bastante abrangente o que resultou em um grande número de registros retornados que tiveram que ser manualmente analisados. A análise de um grande conjunto de dados possibilita uma maior chance de erro na aplicação do protocolo de pesquisa.
- A revisão sistemática do tipo mapeamento sistemático foi executada apenas por um pesquisador. A inexistência de mais de um revisor pode ocasionar em erros na classificação e seleção dos artigos, além de vieses nos resultados.
- A revisão sistemática contempla bases de dados contendo apenas pesquisas e estudos acadêmicos. Neste caso, é possível que existam outras tecnologias importantes no mercado para reconhecimento de voz que não tenham sido citadas nesta revisão.

6.3 Considerações finais

A pesquisa desenvolvida neste trabalho identificou alguns dos principais desafios e dificuldades para o desenvolvimento de uma máquina de tradução. Espera-se que os resultados obtidos contribuam para estimular a pesquisa nas

áreas de reconhecimento e tradução de voz envolvendo o português brasileiro tanto na área acadêmica quanto na indústria, no contexto de DDS.

As tecnologias e ferramentas para reconhecimento de voz encontradas durante a pesquisa foram publicadas na revista IEEE Software em uma coluna relacionada a tecnologias de software [DUA14]. Essa pesquisa também foi apresentada no SEIF Brazil Workshop 2013, realizado no Rio de Janeiro e promovido pela Microsoft Research.

Sendo assim, visto que os estudos envolvendo máquinas de tradução e DDS são recentes, esses resultados poderão servir de estímulo para novos estudos neste contexto, unindo as áreas de engenharia de software e processamento de linguagem natural.

REFERÊNCIAS BIBLIOGRÁFICAS

- [AME00] Amengual J.C., Castano A., Castellanos A., Jimenez V.M., Llorens D., Marzal A., Prat F., Vilar J.M., Benedi J.M., Casacuberta F., Pastor M., Vidal E. "The EUTRANS-I spoken language translation system" , Machine Translation, Depto. de Leng. y Sist. Informaticos, Campus Riu Sec, Universitat Jaume I, 12071 CastellÃ³n de la Plana, Spain, 2000.
- [ARN03] D. Arnold, "Why translation is difficult for computers", In Computers and Translation: A translator's guide. Benjamins Translation Library, 2003.
- [ATA09] Atayero A.A., Ayo C.K., Nicholas I.-O., Ambrose A. "Implementation of 'ASR4CRM': An automated speech-enabled customer care service system" , IEEE EUROCON 2009, EUROCON 2009, 2009.
- [AUD07] Audy, J. L. N.; Prikładnicki, R. "Desenvolvimento Distribuído de Software: Desenvolvimento de software com equipes distribuídas". Rio de Janeiro: Elsevier, 2007.
- [BAC08] Bacchiani M., Beaufays F., Schalkwyk J., Schuster M., Strobe B. "Deploying GOOG-411: Early lessons in data, measurement, and testing", ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, Google, Inc., 2008.
- [BAI07] Bain K., Hines J., Lingras P., Qin Y. "Using speech recognition and intelligent search tools to enhance information accessibility" , Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), Liberated Learning Consortium, Saint Mary's University, Halifax, NS B3H 3C3, Canada; Mathematics and Computer Science Department, Saint Mary's University, Halifax, NS B3H 3C3, Canada; 3IBM China Human Ability and Accessibility Center, 25/F, IBM Tower, 2A Gong Ti Bei Road, Chaoyang District, Beijing 100027, China, 2007.
- [BIN07] Bin X., "A Service Oriented Model for Role Based Global Cooperative Software Development," Convergence Information Technology, 2007.

International Conference on , vol., no., pp.376-381, 21-23 Nov. 2007

- [BRA10] Brazil IT-BPO Book 2008-2009, published by Brasscom, Brazilian Association of Information Technology and Communication Companies, São Paulo, SP, Brazil, 2010.
- [BUB96] Bub, T.; Schwinn, J. "VERBMOBIL: the evolution of a complex large speech-to-speech translation system In Spoken Language", 1996. ICSLP 96. Proceedings., Fourth International Conference on, volume 4, pages 2371–2374 vol.4. IEEE, October 1996.
- [CAL10] Calefato, F.; Lanubile, F.; Minervini, P. "Can Real-Time Machine Translation Overcome Language Barriers in Distributed Requirements Engineering?", Proc. 5th Int'l Conference on Global Software Engineering (ICGSE'10), Princeton, NJ, USA, Aug. 23-26, pp. 257-264, 2010.
- [CAL11] Calefato, F., Lanubile, F., Prikladnicki, R., "A Controlled Experiment on the Effects of Machine Translation in Multilingual Requirements Meetings", Global Software Engineering (ICGSE), 2011 6th IEEE International Conference on, 2011.
- [CAL12] Calefato, F.; Lanubile, F.; Conte, T.; Prikladnicki, R. "Assessing the impact of real-time machine translation on requirements meetings: a replicated experiment", In Proceedings of the ACM-IEEE international symposium on Empirical software engineering and measurement, ESEM '12, pages 251–260, New York, NY, USA, 2012.
- [CAM06] Campbell J.D., Naclerio C. "Instant Messaging comprehension with non-keyboard composition" , Conference on Human Factors in Computing Systems - Proceedings, UMBC, 1000 Hilltop Circle, Baltimore, MD 21250, United States; Colgate University, Hamilton, NY 13346, United States, 2006.
- [CAN12] Canovas A., Tomas J., Lloret J., Garcia M. "Statistical speech translation system based on voice recognition optimization using multimodal sources of knowledge and characteristics vectors" , Computer Standards and Interfaces, Instituto de Investigacion para la

Gestion Integrada de Zonas Costeras, Universidad Politecnica de Valencia, Spain, 2012.

- [CAR01] Carmel, E.; Agarwal, R. Tactical approaches for alleviating distance in global software development. *Software, IEEE*, 18(2):22–29, March 2001.
- [CHA11] Chang Y.-S., Hung S.-H., Wangy N.J.C., Linz B.-S. "CSR: A cloud-assisted speech recognition service for personal mobile device" , *Proceedings of the International Conference on Parallel Processing, Department of Computer Science and Information Engineering, National Taiwan University, Taipei, 100, Taiwan; Inda Mobile Technologies, Inc., India; Department of Information Management, National Taiwan University of Science and Technology, Taipei, 100, Taiwan, 2011.*
- [CHR01] "Chrome Blog: Bringing voice recognition to the web". Disponível em: < <http://chrome.blogspot.com.br/2013/02/bringing-voice-recognition-to-web.html> >. Acesso: 20 set. 2013
- [CHR02] "Voice Driven Web Apps: Introduction to the Web Speech API". Disponível em: < <http://updates.html5rocks.com/2013/01/Voice-Driven-Web-Apps-Introduction-to-the-Web-Speech-API> >. Acesso: 20 set. 2013
- [COH07] Cohen, J. The GALE project: A description and an update. In *Automatic Speech Recognition & Understanding, 2007. ASRU. IEEE Workshop on*, page 237. IEEE, December 2007.
- [COH08] Cohen J. "Embedded speech recognition applications in mobile phones: Status, trends, and challenges" , *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, SRI International, 333 Ravenswood Road, Menlo Park, CA 94025, United States, 2008.*
- [CON07] Conte, T.; Massolar, J.; Mendes, E.; Travassos, G. H. "Web Usability Inspection Technique Based on Design Perspectives". In: *Proceedings of the 21th Brazilian Symposium on Software Engineering (SBES*

2007), v. 1, pp. 394-410, João Pessoa, Brazil. October, 2007.

- [DAM03] Damian, D.; Zowghi, D. "Requirements Engineering Challenges in Multi-Site Software Development Organizations", Requirements Engineering Journal, 8-3, 2003.
- [DAM06] Damian, D.; Moitra, D., "Guest Editors' Introduction: Global Software Development: How Far Have We Come?", Software, IEEE, vol.23, no.5, Sept.-Oct. 2006.
- [DAM07] Damian, D. "Stakeholders in Global Requirements Engineering: Lessons Learned from Practice", IEEE Software, 24-2, 2007.
- [DUA14] Duarte, T.; Prikladnicki, R.; Calefato, F.; Lanubile, F. "Speech Recognition for Voice-Based Machine Translation", IEEE Software, vol. 31, no. 1, 2014.
- [FAL14] "Reconhecimento de voz para o português brasileiro. UFPA – Fala Brasil" Disponível em: <<http://www.laps.ufpa.br/falabrasil/>>. Acesso em: 1 jan. 2014.
- [FRE00] Frederking R., Rudnicky A., Hogan C., Lenzo K. "Interactive speech translation in the DIPLOMAT project" , Machine Translation, Language Technologies Institute, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, United States, 2000.
- [GIA05] Giannakopoulos T., Tatlas N.-A., Ganchev T., Potamitis I. "A practical, real-time speech-driven home automation front-end" , IEEE Transactions on Consumer Electronics, Informatics and Telecommunications Dept., University of Athens, 15784 Athens, Greece; Electrical and Computer Engineering Dept., University of Patras, 26500 Patras, Greece; Technological Educational Institute of Crete, Department of Music Technology and Acoustics, 74100 Rethymno Crete, Greece, 2005
- [GOR97] Gorin A.L., Riccardi G., Wright J.H. "How may I help you?" , Speech Communication, AT and T Labs.-Research, Florham Park, NJ, United States, 1997.

- [HAM07] Hamon, O.; Mostefa, D.; Choukri, K. "End-to-end evaluation of a speech-to-speech translation system in TC-STAR," in Proc. MT-Summit, Copenhagen, 2007.
- [HAS11] Hassan Z., Mohamad A.R., Kalil M.R., Murah M.Z. "Evaluation of Microsoft speech recognition in controlling robot soccer" , Proceedings of the 2011 International Conference on Pattern Analysis and Intelligent Robotics, ICPAIR 2011, Electrical Engineering Department, Ungku Omar Polytechnic, Jalan Raja Musa Mahadi, 31400 Ipoh, Perak, Malaysia; Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, Bangi, Malaysia, 2011.
- [HAU10] Hautasaari, A. Machine translation effects on group interaction: an intercultural collaboration experiment. In Proceedings of the 3rd international conference on Intercultural collaboration, ICIC '10, pages 69–78, New York, NY, USA, 2010.
- [HER01] Herbsleb JD, Moitra D. 2001. Guest editor's introduction:Global software development. IEEE Software 18(2), 2001.
- [HER07] Herbsleb J. D.; "Global Software Engineering: The Future of Socio-technical Coordination", Future of Software Engineering(FOSE'07), IEEE, 2007.
- [JUR08] Jurafsky, D.; Martin, J. H. "Speech and Language Processing 2nd ed.," Prentice Hall Series in Artificial Intelligence, Prentice Hall, 2008.
- [KEA07] Kearney, A.T. "Destination Latin America: A Nearshore Alternative, Technical Report, 2007.
- [KEA11] Kearney, A.T. "Offshoring Opportunities Amid Economic Turbulence", The A.T. Kearney Global Services Location Index, 2011.
- [KIT07] Kitchenham, B.A. "Guidelines for performing Systematic Literature Reviews in Software Engineering"; 2007.
- [KNI10] Knight A.; Almeroth KC.; "Fast caption alignment for automatic indexing of audio". Int J Multimed Data Eng Manag, 2010

- [KNO07] Knob, F. F., "RiskFree4ppm: uma proposta de processo para o gerenciamento de portfólios de projetos distribuídos." Dissertação de Mestrado, PPGCC, Faculdade de Informática, PUCRS, 2007.
- [LAH05] Lahtinen S., Peltonen J. "Adding speech recognition support to UML tools" , Journal of Visual Languages and Computing, Institute of Software Systems, Tampere University of Technology, P.O. Box 553, FIN-33101 Tampere, Finland, 2005.
- [LAU97] Lau R. "WebGALAXY: Beyond point and click - A conversational interface to a browser" , Computer Networks, Spoken Language Systems Group, MIT Laboratory for Computer Science, 545 Technology Square, Cambridge, MA 02139, United States, 1997.
- [LAN08] Lane, M.; Agerfalk, P., "On the Suitability of Particular Software Development Roles to Global Software Development", Global Software Engineering, 2008. ICGSE 2008. IEEE International Conference on , vol., no., pp.3-12, 17-20 August 2008.
- [MAF06] Mafra, S.; Barcelos, R.; Travassos, G. H.; "Aplicando uma Metodologia Baseada em Evidência na Definição de Novas Tecnologias de Software". In: Proceedings of the 20th Brazilian Symposium on Software Engineering (SBES 2006), v. 1, pp. 239 – 254, Florianopolis. October 2006.
- [MAH09] Mardiana B., Hazura H., Fauziyah S., Zahariah M., Hanim A.R., Noor S.M.K. "Homes appliances controlled using speech recognition in wireless network environment" , ICCTD 2009 - 2009 International Conference on Computer Technology and Development, Faculty of Electronics and Computer Engineering, Universiti Teknikal Malaysia Melaka, Ayer Keroh, Malacca, Malaysia, 2009.
- [MAR09] Martignoni, R., "Global Sourcing of Software Development - A Review of Tools and Services," Global Software Engineering, 2009. ICGSE 2009. Fourth IEEE International Conference on , vol., no., pp.303-308, 13-16, July 2009.
- [MEN12] Meng J., Zhang J., Zhao H. "Overview of the speech recognition

technology" , Proceedings - 4th International Conference on Computational and Information Sciences, ICCIS 2012, School of Control and Computer Engineering, North China Electric Power University, Baoding, China, 2012.

- [MIN03] Minker W., Haiber U., Heisterkamp P., Scheible S. "Intelligent dialog overcomes speech technology limitations: The SENECa example" , International Conference on Intelligent User Interfaces, Proceedings IUI, DaimlerChrysler - Res./Technology, 89013 Ulm, Germany; Temic - Sprachverarbeitung GmbH, 89077 Ulm, Germany, 2003.
- [MIN04] Minker W., Haiber U., Heisterkamp P., Scheible S. "The SENECA spoken language dialogue system" , Speech Communication, University of Ulm, Department of Information Technology, Albert-Einstein-Allee 43, 89081 Ulm, Germany; DaimlerChrysler, Res. and Technology, 89013 Ulm, Germany; Temic - SDS GmbH, 89077 Ulm, Germany, 2004.
- [MOC01] Mockus, A.; Herbsleb, J. Challenges of global software development. In Software Metrics Symposium, 2001. METRICS 2001. Proceedings. Seventh International, pages 182–184. IEEE, 2001.
- [MOR93] Morimoto, T.; Takezawa, T.; Yato, F.; Sagayama, S.; Tashiro, T.; Nagata, M.; Kurematsu, A. "ATR's speech translation system: ASURA," in Proc. Eurospeech'93, Geneva, Italy, pp. 1291–1294, September 1993.
- [NET08] Neto J., Meinedo H., Viveiros M., Cassaca R., Martins C., Caseiro D. "Broadcast news subtitling system in Portuguese" , ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, L2F, Spoken Language Systems Lab., INESC-ID; Institute Superior Tecnico, Technical University of Lisbon, 2008.
- [NET11] Neto N., Patrick C., Klautau A., Trancoso I. "Free tools and resources for Brazilian Portuguese speech recognition" , Journal of the Brazilian Computer Society, Federal University of Pará, Augusto Correa, 1, Belém, Brazil; IST/INESC-ID, Alves Redol, 9, Lisbon, Portugal, 2011.

- [NGU02] Nguyen L., Matsoukas S., Davenport J., Kubala F., Schwartz R., Makhoul J. "Progress in transcription of broadcast news using Byblos" , Speech Communication, BBN Technologies, Verizon Communications Inc., Cambridge, MA 02138, United States, 2002.
- [NMV11] Neto J., Meinedo H., Viveiros M. "A media monitoring solution" , ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, Spoken Language Systems Lab., INESC-ID, Portugal; Instituto Superior Técnico, Technical University of Lisbon, Portugal; VoiceInteraction, Speech Processing Technologies, SA, Portugal, 2011.
- [NUS00] Nuseibeh, B.; Easterbrook, S. "Requirements engineering: a roadmap," Proc. Int'l Conf. on the Future of Software Engineering (ICSE '00), pp. 35-46, June 2000.
- [OCH04] Och, F. J.; Ney, H. The alignment template approach to statistical machine translation. Computational Linguistics, vol. 30, no. 4, pp. 417-449, 2004.
- [ODE07] Odell J., Mukerjee K. "Architecture, user interface, and enabling technology in Windows Vista's speech systems" , IEEE Transactions on Computers, Speech Components Group, Microsoft Corporation, One Microsoft Way, Redmond, WA 98052, United States, 2007.
- [OLI12] Oliveira A.L.C., Silva E.S., Macedo H.T., Matos L.N. "Brazilian Portuguese speech-driven answering system [Sistema de atendimento com interação de fala para o Português do Brasil]" , Proceedings of the 6th Euro American Conference on Telematics and Information Systems, EATIS 2012, Universidade Federal de Sergipe, Av. Marechal Rondon, s/n, CEP 49100-000 São Cristovão - SE, Brazil, 2012.
- [ORT09] Ortega A., Garcia J.E., Miguel A., Lleida E. "Real-time live broadcast news subtitling system for Spanish" , Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, Communications Technology Group (GTC), Aragon Institute for Engineering Research (I3A), University of Zaragoza, Spain,

2009.

- [PAL95] Palvia, P.C. A dialectic view of information systems outsourcing - pros and cons, *Information and Management* 29 265-275. 1995.
- [PER12] Pereira M.H.R., Padua F.L.C., Silva G.D., Assis G.T., Zenha T.M. "A multimedia information system to support the discourse analysis of video recordings of television programs" , Iberian Conference on Information Systems and Technologies, CISTI, Lab. de Pesquisas Interdisciplinares Em Informação Multimídia, Centro Federal de Educação Tecnológica de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil; Instituto de Ciências Exatas e Biológicas - ICEB, Universidade Federal de Ouro Preto - UFOP, Ouro Preto, Minas Gerais, Brazil, 2012.
- [PER97] Perdue R.J. "The way we were: Speech technology, platforms and applications in the 'Old' AT&T" , *Speech Communication*, Lucent Technologies, Bell Laboratories, 6200 E. Broad St., Columbus, OH 15303, United States, 1997.
- [PEZ08] Pezzullo J.A., Tung G.A., Rogg J.M., Davis L.M., Brody J.M., Mayo-Smith W.W. "Voice recognition dictation: Radiologist as transcriptionist" , *Journal of Digital Imaging*, Department of Diagnostic Imaging, Warren Alpert School of Medicine, Brown University, Providence, RI, United States; Department of Diagnostic Imaging, Rhode Island Hospital, 593 Eddy Street, Providence, RI 02903, United States, 2008.
- [PIL06] Pilatti, L. S. M., "Estrutura e Características para Análise de Ambientes de Desenvolvimento Global de Software em Organizações Offshore Insourcing." Dissertação de Mestrado, PPGCC, Faculdade de Informática, PUCRS, 2006.
- [PRA13] Prabuwo A.S., Sheikh Abdullah S.N.H., Ganason R., Shamsudin H.C. "Restricted voice recognition for robot movement control" , *International Journal of Advancements in Computing Technology*, Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor Darul Ehsan, Malaysia; Department of Electrical Engineering, Politeknik Sultan Haji Ahmad

Shah, Semambu, 25350 Kuantan, Pahang Darul Makmur, Malaysia, 2013.

- [PRI03] Prikladnicki, R.; Audy, J. L. N.; Evaristo, Roberto. Global software development in practice lessons learned. *Softw. Process: Improve. Pract.*, 8(4):267–281, October 2003.
- [PRI08] Prikladnicki, R.; Damian, D.; Audy, J., "Patterns of Evolution in the Practice of Distributed Software Development in Wholly Owned Subsidiaries: A Preliminary Capability Model," *Global Software Engineering, 2008. ICGSE 2008. IEEE International Conference on*, vol., no., pp.99-108, 17-20 August 2008
- [PRI09] Prikladnicki, R., "Padrões de Evolução na Prática de Desenvolvimento de Software em Ambientes de Internal Offshoring: Um Modelo de Capacidade." Tese de Doutorado, PPGCC, Faculdade de Informática, PUCRS, 2009.
- [PRI13] Prikladnicki, R; Carmel, E; "Is time-zone proximity an advantage for software development? the case of the brazilian IT industry.", *ICSE 2013*, 2013.
- [RAU06] Raux A., Bohus D., Langner B., Black A.W., Eskenazi M. "Doing research on a deployed spoken dialogue system: One year of let's go! Experience" , *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, Language Technologies Institute, Carnegie Mellon University, Pittsburgh, United States*, 2006.
- [ROE92] Roe, D.B.; Pereira, F.C; Sproat, R.W.; Riley, M.D. "Efficient grammar processing for a spoken language translation system," in *Proc. ICASSP, San Francisco, CA*, vol. 1, pp. 213–216, Mar. 1992.
- [SCO13] "What does Scopus cover?" Disponível em: <<http://www.info.sciverse.com/scopus/scopus-in-detail/facts>>. Acesso em: 1 mar. 2013.
- [SET94] Seto S., Kanazawa H., Shinchi H., Takebayashi Y. "Spontaneous speech dialogue system TOSBURG II and its evaluation" , *Speech*

Communication, Toshiba Corporation, Kansai Research Laboratory, 6-26, Motoyama-Minami-cho 8 chome, Higashinada-ku, Kobe-shi, 658, Japan; Toshiba Software Engineering Co. Ltd., 1, Komukai Toshiba-cho, Saiwai-ku, Kawasaki-shi, 210, Japan; Toshiba Corporation, Research and Development Center, 1, Komukai Toshiba-cho, Saiwai-ku, Kawasaki-shi, 210, Japan, 1994.

- [SHU01] Shull, F.; Carver, J.; Travassos, G. H. An empirical methodology for introducing software processes. *SIGSOFT Softw. Eng. Notes*, 26(5):288–296, September 2001.
- [SIL10] Silva P., Batista P., Neto N., Klautau A. "An open-source speech recognizer for Brazilian Portuguese with a windows programming interface" , *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Signal Processing Laboratory, Universidade Federal do Pará, Rua Augusto Correa. 1, 660750110 Belém, PA, Brazil, 2010.
- [SLO84] Slocum, J. *Machine translation: its history, current status, and future prospects*. 1984.
- [SMI13] Smitea, D.; Wohlina, C.; Aurumc, A.; Jabangwea, R.; Numminena, E. "Offshore insourcing in software development: Structuring the decision-making process", *Elsevier The Journal of Systems and Software*, 2013
- [STA10] Stallard, D.; Prasad, R.; Ananthakrishnan, S.; Choi, F.; Saleem, S.; Natarajan, P. Evaluating different confirmation strategies for speech-to-speech translation systems. In *Acoustics Speech and Signal Processing (ICASSP)*, 2010 IEEE International Conference on, pages 5218–5221. IEEE, March 2010.
- [TRA06] Trancoso I., Nunes R., Neves L., Viana C., Moniz H., Caseiro D., Mata A.I. "Recognition of classroom lectures in European Portuguese" , *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, L2F INESC-ID/IST, Lisbon, Portugal; L2F INESC-ID, Lisbon, Portugal; CLUL, Lisbon, Portugal*, 2006.

- [TSO11] Tsontzos G., Orglmeister R. "CMU Sphinx4 speech recognizer in a Service-oriented Computing style" , Proceedings - 2011 IEEE International Conference on Service-Oriented Computing and Applications, SOCA 2011, Electronics and Medical Signal Processing, Technical University of Berlin, Berlin, Germany, 2011.
- [WAI91] Waibel, A.; Jain, A.N.; McNair, A.E.; Saito, H.; Hauptmann, A.G.; Tebelskis, J. "JANUS: A Speech-to-speech translation using connectionist and symbolic processing strategies," in Proc. ICASSP, Toronto, Canada, pp. 793–796, May 1991.
- [WAI08] Waibel, A.; Fugen, C. Spoken language translation. Signal Processing Magazine, IEEE, 25(3):70–79, May 2008.
- [WOH00] Wohlin, C.; Runeson P.; Höst, M.; Ohlsson, M.C.; Regnell, B.; Wesslén A.; "Experimentation in Software Engineering – An Introduction". In Kluwer Academic Publishers, 2000
- [WU12] Wu, S. Overview of communication in global software development process. In Service Operations and Logistics, and Informatics (SOLI), 2012 IEEE International Conference on, pages 474–478. IEEE, July 2012.
- [VAN12] Van Delden S., Umrysh M., Rosario C., Hess G. "Pick-and-place application development using voice and visual commands" , Industrial Robot, Division of Mathematics and Computer Science, University of South Carolina Upstate, Spartanburg, SC, United States; Livingston and Haven Manufacturing, Charlotte, NC, United States, 2012.
- [VYS95] Vysotsky G.J. "VoiceDialingSM - The first speech recognition based service delivered to customer's home from the telephone network" , Speech Communication, NYNEX Science, Technology, Inc., 500 Westchester Avenue, White Plains, NY 10604, United States, 1995.
- [YAN11] Yang, Che-Yu. Cross-Language instant messaging with automatic translation. In Ubi-Media Computing (U-Media), 2011 4th International Conference on, pages 222–226. IEEE, July 2011.

- [YAS11] Yasuoka, M.; Bjorn, P. Machine translation effect on communication: What makes it difficult to communicate through machine translation? In Culture and Computing (Culture Computing), 2011 Second International Conference on, pages 110–115. IEEE, October 2011.
- [YU12] Yu Y. "Research on speech recognition technology and its application" , Proceedings - 2012 International Conference on Computer Science and Electronics Engineering, ICCSEE 2012, Department of Electronics and Information Engineering, Putian University, Putian, Fujian, 351100, China, 2012.
- [YUA11] Yuan X.J., Fan J. "Design and implementation of voice controlled tetris game based on Microsoft SDK" , 2011 International Conference on Multimedia Technology, ICMT 2011, School of Photoelectric Information and Communication Engineering, Beijing Information Science and Technology University, Beijing, 100101, China, 2011.

ANEXO 1: Artigos resultantes da revisão sistemática

ID	Artigo
1	Prabuwono A.S., Sheikh Abdullah S.N.H., Ganason R., Shamsudin H.C. "Restricted voice recognition for robot movement control" , International Journal of Advancements in Computing Technology, Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor Darul Ehsan, Malaysia; Department of Electrical Engineering, Politeknik Sultan Haji Ahmad Shah, Semambu, 25350 Kuantan, Pahang Darul Makmur, Malaysia, 2013.
2	Federico M., Furini M. "An automatic caption alignment mechanism for off-the-shelf speech recognition technologies" , Multimedia Tools and Applications, Servizio Accoglienza Studenti Disabili, Università di Modena e Reggio Emilia, Modena, Italy; Dipartimento di Comunicazione ed Economia, Università di Modena e Reggio Emilia, Reggio Emilia, Italy, 2012.
3	Lamel L. "Multilingual speech processing activities in quæro: Application to multimedia search in unstructured data", Frontiers in Artificial Intelligence and Applications, CNRS-LIMSI, France, 2012.
4	Stefanovic M., Cetic N., Kovacevic M., Kovacevic J., Jankovic M. "Voice control system with advanced recognition", 2012 20th Telecommunications Forum, TELFOR 2012 - Proceedings, Faculty of Technical Sciences, Trg Dositeja Obradovica 6, 21000 Novi Sad, Serbia, 2012.
5	Pereira M.H.R., Padua F.L.C., Silva G.D., Assis G.T., Zenha T.M. "A multimedia information system to support the discourse analysis of video recordings of television programs" , Iberian Conference on Information Systems and Technologies, CISTI, Lab. de Pesquisas Interdisciplinares

	Em Informação Multimídia, Centro Federal de Educação Tecnológica de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil; Instituto de Ciências Exatas e Biológicas - ICEB, Universidade Federal de Ouro Preto - UFOP, Ouro Preto, Minas Gerais, Brazil, 2012.
6	Meng J., Zhang J., Zhao H. "Overview of the speech recognition technology" , Proceedings - 4th International Conference on Computational and Information Sciences, ICCIS 2012, School of Control and Computer Engineering, North China Electric Power University, Baoding, China, 2012.
7	Canovas A., Tomas J., Lloret J., Garcia M. "Statistical speech translation system based on voice recognition optimization using multimodal sources of knowledge and characteristics vectors" , Computer Standards and Interfaces, Instituto de Investigacion para la Gestion Integrada de Zonas Costeras, Universidad Politecnica de Valencia, Spain, 2012.
8	Van Delden S., Umrysh M., Rosario C., Hess G. "Pick-and-place application development using voice and visual commands" , Industrial Robot, Division of Mathematics and Computer Science, University of South Carolina Upstate, Spartanburg, SC, United States; Livingston and Haven Manufacturing, Charlotte, NC, United States, 2012.
9	Da Silva D.D.C., Vasconcelos C.R., Neto B.G.A., Fachine J.M. "Evaluation of the impact in reducing the number of parameters for continuous speech recognition for Brazilian Portuguese" , 2012 ISSNIP Biosignals and Biorobotics Conference: Biosignals and Robotics for Better and Safer Living, BRC 2012, Telematics Coordination, Federal Institute of Technological Education - IFPB, Campina Grande, Brazil; Computer and Electrical Engineering Center, Federal University of Campina Grande - UFCG, Campina Grande, Brazil, 2012.
10	Oliveira A.L.C., Silva E.S., Macedo H.T., Matos L.N. "Brazilian Portuguese speech-driven answering system [Sistema de atendimento com interação de fala para o Português do Brasil]" , Proceedings of the

	6th Euro American Conference on Telematics and Information Systems, EATIS 2012, Universidade Federal de Sergipe, Av. Marechal Rondon, s/n, CEP 49100-000 São Cristovão - SE, Brazil, 2012.
11	Khalil R.T., Khalifeh A.F., Darabkh K.A. "Mobile-free driving with Android phones: System design and performance evaluation" , International Multi-Conference on Systems, Signals and Devices, SSD 2012 - Summary Proceedings, German University in Cairo, Egypt; German Jordan University, Jordan; University of Jordan, Jordan, 2012.
12	Derbali M., Jarrah M., Wahid M.T. "A review of speech recognition with Sphinx engine in language detection" , Journal of Theoretical and Applied Information Technology, Information Systems Department, Faculty of Computing and Information Technology King Abdulaziz University, Jeddah, Saudi Arabia; Information Technology Department, Faculty of Computing and Information Technology King Abdulaziz University, Jeddah, Saudi Arabia; Faculty of Computer Science and Information Systems Universiti Teknologi Malaysia, Malaysia, 2012.
13	Yu Y. "Research on speech recognition technology and its application" , Proceedings - 2012 International Conference on Computer Science and Electronics Engineering, ICCSEE 2012, Department of Electronics and Information Engineering, Putian University, Putian, Fujian, 351100, China, 2012.
14	Kalnikaite V., Ehlen P., Whittaker S. "Markup as you talk: Establishing effective memory cues while still contributing to a meeting" , Proceedings of the ACM Conference on Computer Supported Cooperative Work, CSCW, Interactables, Cambridge, United Kingdom; AT and T, San Francisco, CA, United States; University of California, Santa Cruz, CA, United States, 2012.
15	Wrigley S.N., Hain T. "Making an automatic speech recognition service freely available on the web" , Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH,

	Department of Computer Science, University of Sheffield Regent Court, 211 Portobello, Sheffield S1 4DP, United Kingdom, 2011.
16	Tsontzos G., Orglmeister R. "CMU Sphinx4 speech recognizer in a Service-oriented Computing style" , Proceedings - 2011 IEEE International Conference on Service-Oriented Computing and Applications, SOCA 2011, Electronics and Medical Signal Processing, Technical University of Berlin, Berlin, Germany, 2011.
17	Wrigley S.N., Hain T. "Web-based automatic speech recognition service - WebASR" , Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, Department of Computer Science, University of Sheffield, Regent Court, 211 Portobello, Sheffield S1 4DP, United Kingdom, 2011.
18	Chang Y.-S., Hung S.-H., Wangy N.J.C., Linz B.-S. "CSR: A cloud-assisted speech recognition service for personal mobile device" , Proceedings of the International Conference on Parallel Processing, Department of Computer Science and Information Engineering, National Taiwan University, Taipei, 100, Taiwan; Inda Mobile Technologies, Inc., India; Department of Information Management, National Taiwan University of Science and Technology, Taipei, 100, Taiwan, 2011.
19	Yuan X.J., Fan J. "Design and implementation of voice controlled tetris game based on Microsoft SDK" , 2011 International Conference on Multimedia Technology, ICMT 2011, School of Photoelectronic Information and Communication Engineering, Beijing Information Science and Technology University, Beijing, 100101, China, 2011.
20	Hassan Z., Mohamad A.R., Kalil M.R., Murah M.Z. "Evaluation of Microsoft speech recognition in controlling robot soccer" , Proceedings of the 2011 International Conference on Pattern Analysis and Intelligent Robotics, ICPAIR 2011, Electrical Engineering Department, Ungku Omar Polytechnic, Jalan Raja Musa Mahadi, 31400 Ipoh, Perak, Malaysia; Faculty of Information Science and Technology, Universiti Kebangsaan

	Malaysia, Bangi, Malaysia, 2011.
21	Neto J., Meinedo H., Viveiros M. "A media monitoring solution" , ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, Spoken Language Systems Lab., INESC-ID, Portugal; Instituto Superior Técnico, Technical University of Lisbon, Portugal; VoiceInteraction, Speech Processing Technologies, SA, Portugal, 2011.
22	Sundermeyer M., Nussbaum-Thom M., Wiesler S., Plahl C., El-Desoky Mousa A., Hahn S., Nolden D., Schluter R., Ney H. "The RWTH 2010 Quaero ASR evaluation system for English, French, and German" , ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, Human Language Technology and Pattern Recognition, Computer Science Department, RWTH Aachen University, Germany, 2011.
23	Chang C.A., Strahan R., Jolley D. "Non-clinical errors using voice recognition dictation software for radiology reports: A retrospective audit" , Journal of Digital Imaging, Southern Health Department of Diagnostic Imaging, 246 Clayton Road, Clayton, VIC 3168, Australia, 2011.
24	Neto N., Patrick C., Klautau A., Trancoso I. "Free tools and resources for Brazilian Portuguese speech recognition" , Journal of the Brazilian Computer Society, Federal University of Pará, Augusto Correa, 1, Belém, Brazil; IST/INESC-ID, Alves Redol, 9, Lisbon, Portugal, 2011.
25	Silva P., Batista P., Neto N., Klautau A. "An open-source speech recognizer for Brazilian Portuguese with a windows programming interface" , Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), Signal Processing Laboratory, Universidade Federal do Pará, Rua Augusto Correa. 1, 660750110 Belém, PA, Brazil, 2010.
26	Schiller J. "Modern meeting management and information retrieval: Automatic protocol generation and meeting workflow support" ,

	Proceedings of the Conference on Object-Oriented Programming Systems, Languages, and Applications, OOPSLA, Technische Universitaet Muenchen, Siemens AG, Germany, 2009.
27	Mardiana B., Hazura H., Fauziyah S., Zahariah M., Hanim A.R., Noor S.M.K. "Homes appliances controlled using speech recognition in wireless network environment" , ICCTD 2009 - 2009 International Conference on Computer Technology and Development, Faculty of Electronics and Computer Engineering, Universiti Teknikal Malaysia Melaka, Ayer Keroh, Malacca, Malaysia, 2009.
28	Garner P.N., Dines J., Hain T., El Hannani A., Karafiat M., Korchagin D., Lincoln M., Wan V., Zhang L. "Real-time ASR from meetings" , Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, Idiap Research Institute, Martigny, Switzerland; Speech and Hearing Research Group, University of Sheffield, United Kingdom; Speech Processing Group, Brno University of Technology, Czech Republic; Centre for Speech Technology Research, University of Edinburgh, United Kingdom, 2009.
29	Ortega A., Garcia J.E., Miguel A., Lleida E. "Real-time live broadcast news subtitling system for Spanish" , Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, Communications Technology Group (GTC), Aragon Institute for Engineering Research (I3A), University of Zaragoza, Spain, 2009.
30	Rybach D., Gollan C., Heigold G., Hoffmeister B., Loof J., Schluter R., Ney H. "The RWTH Aachen University open source speech recognition system" , Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, Human Language Technology and Pattern Recognition, Computer Science Department, RWTH Aachen University, Germany, 2009.
31	Atayero A.A., Ayo C.K., Nicholas I.-O., Ambrose A. "Implementation of

	'ASR4CRM': An automated speech-enabled customer care service system" , IEEE EUROCON 2009, EUROCON 2009, 2009.
32	Baker J.M., Li D., Glass J., Khudanpur S., Lee C.-H., Morgan N., O'Shaughnessy D. "Research developments and directions in speech recognition and understanding, Part 1" , IEEE Signal Processing Magazine, Saras Institute, West Newton, MA, United States; Microsoft Research, Redmond, WA, United States; MIT, Computer Science and Artificial Intelligence Laboratory, United States; GWC Whiting School of Engineering, Johns Hopkins University, Baltimore, MD, United States; Georgia Institute of Technology, Atlanta, GA, United States; University of California, Berkeley, CA, United States; INRS-EMT, University of Quebec, Montreal, QC, Canada, 2009.
33	Hain T., El Hannani A., Wrigley S.N., Wan V. "Automatic speech recognition for scientific purposes - WebASR" , Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, Dept. of Computer Science, University of Sheffield, Regent Court, 211 Portobello, Sheffield S1 4DP, United Kingdom, 2008.
34	Tur G., Stolcke A., Voss L., Dowding J., Favre B., Fernandez R., Frampton M., Frandsen M., Frederickson C., Graciarena M., Hakkani-Tur D., Kintzing D., Leveque K., Mason S., Niekrasz J., Peters S., Purver M., Riedhammer K., Shriberg E., Tien J., Vergyri D., Yang F. "The calo meeting speech recognition and understanding system" , 2008 IEEE Workshop on Spoken Language Technology, SLT 2008 - Proceedings, SRI International, Menlo Park,CA, United States; Stanford University, Center for the Study of Language and Information (CSLI), Stanford, CA, United States; International Computer Science Institute (ICSI), Berkeley, CA, United States, 2008.
35	Pezzullo J.A., Tung G.A., Rogg J.M., Davis L.M., Brody J.M., Mayo-Smith W.W. "Voice recognition dictation: Radiologist as transcriptionist" , Journal of Digital Imaging, Department of Diagnostic Imaging, Warren

	Alpert School of Medicine, Brown University, Providence, RI, United States; Department of Diagnostic Imaging, Rhode Island Hospital, 593 Eddy Street, Providence, RI 02903, United States, 2008.
36	O'Shaughnessy D. "Invited paper: Automatic speech recognition: History, methods and challenges" , Pattern Recognition, INRS-EMT, University of Quebec, 800 de la Gauchetiere West, Montreal, Que. H5A 1K6, Canada, 2008.
37	Bacchiani M., Beaufays F., Schalkwyk J., Schuster M., Strobe B. "Deploying GOOG-411: Early lessons in data, measurement, and testing" , ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, Google, Inc., 2008.
38	Neto J., Meinedo H., Viveiros M., Cassaca R., Martins C., Caseiro D. "Broadcast news subtitling system in Portuguese" , ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, L2F, Spoken Language Systems Lab., INESC-ID; Institute Superior Tecnico, Technical University of Lisbon, 2008.
39	Cohen J. "Embedded speech recognition applications in mobile phones: Status, trends, and challenges" , ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, SRI International, 333 Ravenswood Road, Menlo Park, CA 94025, United States, 2008.
40	Renais S., Hain T., Boudard H. "Recognition and understanding of meetings the AMI and AMIDA projects" , 2007 IEEE Workshop on Automatic Speech Recognition and Understanding, ASRU 2007, Proceedings, Centre for Speech Technology Research, University of Edinburgh, Edinburgh EH3 5EU, United Kingdom; Dept. of Computer Science, University of Sheffield, Sheffield S1 4DP, United Kingdom; IDIAP Research Institute, 1920 Martigny, Switzerland, 2007.

- | | |
|----|---|
| 41 | Bain K., Hines J., Lingras P., Qin Y. "Using speech recognition and intelligent search tools to enhance information accessibility" , Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), Liberated Learning Consortium, Saint Mary's University, Halifax, NS B3H 3C3, Canada; Mathematics and Computer Science Department, Saint Mary's University, Halifax, NS B3H 3C3, Canada; 3IBM China Human Ability and Accessibility Center, 25/F, IBM Tower, 2A Gong Ti Bei Road, Chaoyang District, Beijing 100027, China, 2007. |
| 42 | Benzeghiba M., De Mori R., Deroo O., Dupont S., Erbes T., Jouvét D., Fissore L., Laface P., Mertins A., Ris C., Rose R., Tyagi V., Wellekens C. "Automatic speech recognition and speech variability: A review" , Speech Communication, Multitel, Parc Initialis, Avenue Copernic, B-7000 Mons, Belgium, 2007. |
| 43 | Odell J., Mukerjee K. "Architecture, user interface, and enabling technology in Windows Vista's speech systems" , IEEE Transactions on Computers, Speech Components Group, Microsoft Corporation, One Microsoft Way, Redmond, WA 98052, United States, 2007. |
| 44 | Lamel L., Gauvain J.-L., Adda G., Barras C., Bilinski E., Galibert O., Pujol A., Schwenk H., Zhu X. "The LIMSI 2006 TC-STAR EPPS transcription systems" , ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, Spoken Language Processing Group, LIMSI-CNRS, BP 133, 91403 Orsay cedex, France, 2007. |
| 45 | Trancoso I., Nunes R., Neves L., Viana C., Moniz H., Caseiro D., Mata A.I. "Recognition of classroom lectures in European Portuguese" , Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, L2F INESC-ID/IST, Lisbon, Portugal; L2F INESC-ID, Lisbon, Portugal; CLUL, Lisbon, Portugal, 2006. |

46	Campbell J.D., Naclerio C. "Instant Messaging comprehension with non-keyboard composition" , Conference on Human Factors in Computing Systems - Proceedings, UMBC, 1000 Hilltop Circle, Baltimore, MD 21250, United States; Colgate University, Hamilton, NY 13346, United States, 2006.
47	Loof J., Bisani M., Gollan Ch., Heigold G., Hoffmeister B., Plahl Ch., Schluter R., Ney H. "The 2006 RWTH parliamentary speeches transcription system" , Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, Informatik 6, Computer Science Dept., RWTH Aachen University, Aachen, Germany, 2006.
48	Raux A., Bohus D., Langner B., Black A.W., Eskenazi M. "Doing research on a deployed spoken dialogue system: One year of let's go! Experience" , Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, Language Technologies Institute, Carnegie Mellon University, Pittsburgh, United States, 2006.
49	Giannakopoulos T., Tatlas N.-A., Ganchev T., Potamitis I. "A practical, real-time speech-driven home automation front-end" , IEEE Transactions on Consumer Electronics, Informatics and Telecommunications Dept., University of Athens, 15784 Athens, Greece; Electrical and Computer Engineering Dept., University of Patras, 26500 Patras, Greece; Technological Educational Institute of Crete, Department of Music Technology and Acoustics, 74100 Rethymno Crete, Greece, 2005.
50	Lahtinen S., Peltonen J. "Adding speech recognition support to UML tools" , Journal of Visual Languages and Computing, Institute of Software Systems, Tampere University of Technology, P.O. Box 553, FIN-33101 Tampere, Finland, 2005.
51	Minker W., Haiber U., Heisterkamp P., Scheible S. "The SENECA spoken language dialogue system" , Speech Communication, University of Ulm, Department of Information Technology, Albert-Einstein-Allee 43, 89081

	Ulm, Germany; DaimlerChrysler, Res. and Technology, 89013 Ulm, Germany; Temic - SDS GmbH, 89077 Ulm, Germany, 2004.
52	Minker W., Haiber U., Heisterkamp P., Scheible S. "Intelligent dialog overcomes speech technology limitations: The SENECa example" , International Conference on Intelligent User Interfaces, Proceedings IUI, DaimlerChrysler - Res./Technology, 89013 Ulm, Germany; Temic - Sprachverarbeitung GmbH, 89077 Ulm, Germany, 2003.
53	Nguyen L., Matsoukas S., Davenport J., Kubala F., Schwartz R., Makhoul J. "Progress in transcription of broadcast news using Byblos" , Speech Communication, BBN Technologies, Verizon Communications Inc., Cambridge, MA 02138, United States, 2002.
54	Frederking R., Rudnicky A., Hogan C., Lenzo K. "Interactive speech translation in the DIPLOMAT project" , Machine Translation, Language Technologies Institute, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, United States, 2000.
55	Levin L., Lavie A., Woszczyna M., Gates D., Gavalda M., Koll D., Waibel A. "The JANUS-III translation system: Speech-to-speech translation in multiple domains" , Machine Translation, Language Technologies Institute, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213-3891, United States, 2000.
56	Amengual J.C., Castano A., Castellanos A., Jimenez V.M., Llorens D., Marzal A., Prat F., Vilar J.M., Benedi J.M., Casacuberta F., Pastor M., Vidal E. "The EUTRANS-I spoken language translation system" , Machine Translation, Depto. de Leng. y Sist. Informaticos, Campus Riu Sec, Universitat Jaume I, 12071 CastellÃ³n de la Plana, Spain, 2000.
57	Lau R. "WebGALAXY: Beyond point and click - A conversational interface to a browser" , Computer Networks, Spoken Language Systems Group, MIT Laboratory for Computer Science, 545 Technology Square, Cambridge, MA 02139, United States, 1997.

58	Perdue R.J. "The way we were: Speech technology, platforms and applications in the 'Old' AT&T" , Speech Communication, Lucent Technologies, Bell Laboratories, 6200 E. Broad St., Columbus, OH 15303, United States, 1997.
59	Gorin A.L., Riccardi G., Wright J.H. "How may I help you?" , Speech Communication, AT and T Labs.-Research, Florham Park, NJ, United States, 1997.
60	Vysotsky G.J. "VoiceDialingSM - The first speech recognition based service delivered to customer's home from the telephone network" , Speech Communication, NYNEX Science, Technology, Inc., 500 Westchester Avenue, White Plains, NY 10604, United States, 1995.
61	Seto S., Kanazawa H., Shinchu H., Takebayashi Y. "Spontaneous speech dialogue system TOSBURG II and its evaluation" , Speech Communication, Toshiba Corporation, Kansai Research Laboratory, 6-26, Motoyama-Minami-cho 8 chome, Higashinada-ku, Kobe-shi, 658, Japan; Toshiba Software Engineering Co. Ltd., 1, Komukai Toshiba-cho, Saiwai-ku, Kawasaki-shi, 210, Japan; Toshiba Corporation, Research and Development Center, 1, Komukai Toshiba-cho, Saiwai-ku, Kawasaki-shi, 210, Japan, 1994.
62	Anastasiou D. "Speech-to-speech translation in an assisted living lab" , ACM International Conference Proceeding Series, Department of Computer Science/Languages and Literary Studies, SFB/TR 8, University of Bremen, Bremen, Germany, 2011.
63	Tomas J., Canovas A., Lloret J., Garcia M. "Speech translation statistical system using multimodal sources of knowledge" , Proceedings - 5th International Multi-Conference on Computing in the Global Information Technology, Universidad Politécnica de Valencia,, Gandia, Spain, 2010.

ANEXO 2 - IBM ViaVoice: Protocolo para avaliação da qualidade do reconhecimento de voz

Informações gerais sobre o teste	
Número de participantes:	<ul style="list-style-type: none"> • A tecnologia foi avaliada pelo próprio pesquisador.
Treinamento do sistema de reconhecimento de voz:	<p>O IBM ViaVoice realiza reconhecimento de voz conhecimento “<i>speaker-dependent</i>”. Neste caso é necessário o treinamento do sistema para reconhecimento da voz de um usuário específico.</p> <p>O sistema oferece 4 opções de treinamento:</p> <ol style="list-style-type: none"> 1) Bentinho e Capitu (1): 70 frases, em torno de 20 minutos de leitura. 2) Bentinho e Capitu (2): 185 frases, em torno de 30 minutos de leitura. 3) A ilha do tesouro (1): 70 frases, em torno de 20 minutos de leitura. 4) A ilha do tesouro (2): 167 frases, em torno de 30 minutos de leitura. <p>Para o treinamento foi executada a opção 2 listada acima.</p>
Versão do sistema:	<p>IBM ViaVoice for Windows Pro USB Edition – Release 9.</p> <p>Atualmente o IBM ViaVoice é um sistema descontinuado.</p>
Funcionalidade avaliada:	<p>Está sendo avaliada a funcionalidade de reconhecimento de ditado. Essa funcionalidade pode ser acessada através de menu ViaVoice, opção Ditar, seguido de opção “Diretamente para aplicativo”.</p> <p>O aplicativo chamado de “Bloco de notas” está sendo utilizado em conjunto para receber a saída do texto sendo reconhecido.</p>
Objetivo do teste:	<p>Avaliar se o IBM ViaVoice pode ser utilizado para permitir reconhecimento de fala em tempo real para o idioma português</p>

	brasileiro em um contexto de engenharia de software. Se demonstrado que o IBM ViaVoice apresenta bons resultados, ele poderá ser utilizado como componente de reconhecimento de voz em uma máquina de tradução de voz.										
Critérios de avaliação											
Tempo de reconhecimento:	<p>Deve ser medido o tempo que o sistema leva para reconhecer uma frase completa. Será contabilizado o tempo apenas ao final da pronuncia completa da frase.</p> <p>Limitação: o tempo será cronometrado manualmente visto que o ViaVoice não possui mecanismo que permita a avaliação do tempo de reconhecimento.</p>										
Avaliação do reconhecimento:	<p>Cada frase transcrita pelo IBM Via Voice terá sua precisão classificada em um dos critérios abaixo:</p> <table border="1"> <thead> <tr> <th>Valor</th> <th>Descrição</th> </tr> </thead> <tbody> <tr> <td>4</td> <td> <p><i>Completamente preciso</i></p> <p>A transcrição claramente reflete a informação contida na sentença original. Ela é perfeitamente clara, inteligível, gramaticalmente correta e pode ser lida como um texto qualquer.</p> </td> </tr> <tr> <td>3</td> <td> <p><i>Bastante precisa</i></p> <p>A transcrição, em geral, reflete a informação contida na sentença original, apesar de algumas imprecisões. Ela é geralmente clara e inteligível e qualquer um pode quase que imediatamente entender o que ela significa.</p> </td> </tr> <tr> <td>2</td> <td> <p><i>Um pouco imprecisa</i></p> <p>A transcrição reflete de forma pobre a informação contida na sentença original. Ela contém erros gramaticais e palavras incorretas. A ideia geral é inteligível somente depois de considerável análise.</p> </td> </tr> <tr> <td>1</td> <td> <p><i>Completamente imprecisa</i></p> <p>A transcrição não é inteligível e não é possível obter</p> </td> </tr> </tbody> </table>	Valor	Descrição	4	<p><i>Completamente preciso</i></p> <p>A transcrição claramente reflete a informação contida na sentença original. Ela é perfeitamente clara, inteligível, gramaticalmente correta e pode ser lida como um texto qualquer.</p>	3	<p><i>Bastante precisa</i></p> <p>A transcrição, em geral, reflete a informação contida na sentença original, apesar de algumas imprecisões. Ela é geralmente clara e inteligível e qualquer um pode quase que imediatamente entender o que ela significa.</p>	2	<p><i>Um pouco imprecisa</i></p> <p>A transcrição reflete de forma pobre a informação contida na sentença original. Ela contém erros gramaticais e palavras incorretas. A ideia geral é inteligível somente depois de considerável análise.</p>	1	<p><i>Completamente imprecisa</i></p> <p>A transcrição não é inteligível e não é possível obter</p>
Valor	Descrição										
4	<p><i>Completamente preciso</i></p> <p>A transcrição claramente reflete a informação contida na sentença original. Ela é perfeitamente clara, inteligível, gramaticalmente correta e pode ser lida como um texto qualquer.</p>										
3	<p><i>Bastante precisa</i></p> <p>A transcrição, em geral, reflete a informação contida na sentença original, apesar de algumas imprecisões. Ela é geralmente clara e inteligível e qualquer um pode quase que imediatamente entender o que ela significa.</p>										
2	<p><i>Um pouco imprecisa</i></p> <p>A transcrição reflete de forma pobre a informação contida na sentença original. Ela contém erros gramaticais e palavras incorretas. A ideia geral é inteligível somente depois de considerável análise.</p>										
1	<p><i>Completamente imprecisa</i></p> <p>A transcrição não é inteligível e não é possível obter</p>										

	<p>a informação contida na sentença original. Analisar o conteúdo da transcrição não produz resultados, e mesmo com algum contexto do que pode ter sido dito entende-se que qualquer tentativa de adivinhar o significado do que está sendo dito teria um resultado pouco confiável.</p>
Estratégia	
<p>Descrição:</p>	<p>Foi definido um conjunto de frases. Cada uma dessas frases deve ser lida em tom de voz normal e velocidade normal. A leitura da frase deve ser contínua e não pausada (palavra por palavra). Cada palavra deve ser dita com todas as letras evitando omissão de partes das palavras.</p>
<p>Número de testes</p>	<p>Será executado uma tentativa para cada frase. Em caso de problema de pronuncia, a frase poderá ser repetida.</p>
Conteúdo a ser reconhecido	
<p>O vocabulário classificado como simples é composto de frases curtas do dia a dia. Já o vocabulário classificado como complexo contém frases maiores e relacionadas a engenharia de software.</p>	
Frases de testes	Vocabulário
Bom dia	Simples
Boa tarde	Simples
Sim	Simples
Não	Simples
Concordo com você	Simples
Discordo de você	Simples
Eu acho que sim.	Simples

Eu acho que não.	Simple
Não tenho certeza.	Simple
O que você acha?	Simple
Como você está?	Simple
Me fale sobre a tarefa que está trabalhando.	Complexo
Você termina essa tarefa até amanhã?	Complexo
Qual a situação do projeto?	Complexo
Você tem algo que te impeça de continuar suas atividades?	Complexo
Terminou o defeito que estava trabalhando?	Complexo
Vamos entregar o projeto dentro do prazo planejado.	Complexo
O projeto está dentro do orçamento e prazo.	Complexo
Não temos problemas de qualidade.	Complexo
Esse projeto possui 10 iterações.	Complexo
Após a fase de desenvolvimento, realizamos a qualificação do produto.	Complexo
Devo terminar essa atividade até amanhã.	Complexo
Estou com bastante dificuldade para resolver esse defeito.	Complexo
Não existe quase informação sobre essa nova tecnologia.	Complexo
Estou atualmente trabalhando em uma investigação.	Complexo
Atualmente estou terminando os testes da aplicação.	Complexo
Não foi encontrado nenhum defeito na funcionalidade desenvolvida.	Complexo
O cliente quer essa nova funcionalidade no produto.	Complexo

ANEXO 3 – IBM ViaVoice: Resultados da avaliação do reconhecimento de voz

Frases de testes	Texto reconhecido	Tempo	Avaliação Valor	Vocabulário
Bom dia	bom dia	1,3	4	Simple
Boa tarde	Botar	1,3	1	Simple
Sim	Sem	1	1	Simple
Não	Não	1	4	Simple
Concordo com você	concordo com você	1,6	4	Simple
Discordo de você	discordo de você	1,3	4	Simple
Eu acho que sim.	eu acho que sim	2	4	Simple
Eu acho que não.	eu acho que não	2	4	Simple
Não tenho certeza.	não tenho certeza	1,5	4	Simple
O que você acha?	o que voce acha	1,5	3	Simple
Como você está?	como você está	1,5	3	Simple
Me fale sobre a tarefa que está trabalhando.	me fale sobre a tarefa que está trabalhando	1,5	4	Complexo
Você termina essa tarefa até amanhã?	você teme nessa tarefa até amanhã	2	1	Complexo
Qual a situação do projeto?	o qual a situação do projeto	2,5	3	Complexo
Você tem algo que te impeça de continuar suas atividades?	você tem algo que te impeça de os nossos atividades	2	2	Complexo
Terminou o defeito que estava trabalhando?	Terminou o defeito que estava trabalhando?	1,8	4	Complexo
Vamos entregar o projeto dentro do prazo planejado.	vamos entregar o projeto dentro do prazo planejar	1,9	2	Complexo
O projeto está dentro do orçamento e prazo.	o projeto está dentro do orçamento e prazos	1,9	3	Complexo
Não temos problemas de qualidade.	Não temos problemas de qualidade.	1,8	4	Complexo
Esse projeto possui 10 iterações.	esse projeto possui dez interações	1,9	2	Complexo

Após a fase de desenvolvimento, realizamos a qualificação do produto.	após a fase de desenvolvimento realizamos a qualificação do produto	1,6	3	Complexo
Devo terminar essa atividade até amanhã.	deu terminara santidade amanhã	1,7	1	Complexo
Estou com bastante dificuldade para resolver esse defeito.	Estou com bastante dificuldade para resolver esse defeito.	1,9	4	Complexo
Não existe quase informação sobre essa nova tecnologia.	Não existe quase informação sobre essa nova tecnologia.	2	4	Complexo
Estou atualmente trabalhando em uma investigação.	Estou atualmente trabalhando em uma investigação.	2	4	Complexo
Atualmente estou terminando os testes da aplicação.	o aumento terminando os testes da aplicação	1,7	1	Complexo
Não foi encontrado nenhum defeito na funcionalidade desenvolvida.	não foi encontrado nenhum defeito na funcionalidade envolvido	1,8	2	Complexo
O cliente quer essa nova funcionalidade no produto.	o cliente quer essa nova função da idade no produto	2	1	Complexo

ANEXO 4 - Google: Protocolo para avaliação das tecnologias de reconhecimento e tradução.

Informações gerais sobre o teste

Tecnologias avaliadas	<ul style="list-style-type: none"> • Google Web Speech API – disponível no navegador Google Chrome. • Google Translator – disponível através de web service API. <p>Versão: testes realizados em dezembro de 2013.</p>
Participantes:	<ul style="list-style-type: none"> • O estudo foi realizado com a participação de 4 mestrandos do cursos de ciência da computação da universidade PUC-RS. Esses alunos fizeram prova de proficiência em inglês e por isso estão aptos a avaliarem se a tradução foi efetuada corretamente.
Funcionalidades avaliadas:	<ul style="list-style-type: none"> • Reconhecimento de voz para o português brasileiro. • Tradução de texto do português brasileiro para o inglês.
Objetivo do teste:	<ul style="list-style-type: none"> • Avaliar a qualidade do reconhecimento de voz da tecnologia Google Web Speech API para o idioma português brasileiro. • Avaliar a qualidade da tradução para o inglês produzida pelo Google Translator se recebido como entrada o texto, em português brasileiro, resultante do reconhecimento de voz realizado pelo Google Web Speech API.

Critérios de avaliação

Qualidade do reconhecimento de voz:	<p>Cada frase transcrita pela Google Web Speech API terá sua precisão classificada em um dos critérios abaixo:</p> <table border="1" style="width: 100%;"> <thead> <tr> <th style="text-align: center;">Valor</th> <th style="text-align: center;">Descrição</th> </tr> </thead> <tbody> <tr> <td style="text-align: center;">4</td> <td><i>Completamente preciso</i> A transcrição claramente reflete a informação contida na sentença original. Ela é perfeitamente clara, inteligível, gramaticalmente correta e pode ser lida como um texto qualquer.</td> </tr> <tr> <td style="text-align: center;">3</td> <td><i>Bastante preciso</i> A transcrição, em geral, reflete a informação contida na sentença original, apesar de algumas imprecisões. Ela é geralmente clara e inteligível e qualquer um pode quase que imediatamente</td> </tr> </tbody> </table>	Valor	Descrição	4	<i>Completamente preciso</i> A transcrição claramente reflete a informação contida na sentença original. Ela é perfeitamente clara, inteligível, gramaticalmente correta e pode ser lida como um texto qualquer.	3	<i>Bastante preciso</i> A transcrição, em geral, reflete a informação contida na sentença original, apesar de algumas imprecisões. Ela é geralmente clara e inteligível e qualquer um pode quase que imediatamente
Valor	Descrição						
4	<i>Completamente preciso</i> A transcrição claramente reflete a informação contida na sentença original. Ela é perfeitamente clara, inteligível, gramaticalmente correta e pode ser lida como um texto qualquer.						
3	<i>Bastante preciso</i> A transcrição, em geral, reflete a informação contida na sentença original, apesar de algumas imprecisões. Ela é geralmente clara e inteligível e qualquer um pode quase que imediatamente						

		entender o que ela significa.
	2	<i>Um pouco impreciso</i> A transcrição reflete de forma pobre a informação contida na sentença original. Ela contém erros gramaticais e palavras incorretas. A ideia geral é inteligível somente depois de considerável análise.
	1	<i>Completamente impreciso</i> A transcrição não é inteligível e não é possível obter a informação contida na sentença original. Analisar o conteúdo da transcrição não produz resultados, e mesmo com algum contexto do que pode ter sido dito entende-se que qualquer tentativa de adivinhar o significado do que está sendo dito teria um resultado pouco confiável.
Qualidade da tradução:	Cada frase traduzida para o inglês pela tecnologia Google Translator terá sua precisão classificada em um dos critérios abaixo:	
	Valor	Descrição
	4	<i>Completamente precisa</i> A tradução claramente reflete a informação contida na sentença original. Ela é perfeitamente clara, inteligível, gramaticalmente correta e pode ser lida como um texto qualquer.
	3	<i>Bastante precisa</i> A tradução, em geral, reflete a informação contida na sentença original, apesar de algumas imprecisões. Ela é geralmente clara e inteligível e qualquer um pode quase que imediatamente entender o que ela significa.
	2	<i>Um pouco imprecisa</i> A tradução reflete de forma pobre a informação contida na sentença original. Ela contém erros gramaticais e palavras incorretas. A ideia geral é inteligível somente depois de considerável análise.
	1	<i>Completamente imprecisa</i> A tradução não é inteligível e não é possível obter a informação contida na sentença original. Analisar o conteúdo da transcrição não produz resultados, e mesmo com algum contexto do que pode ter sido dito entende-se que qualquer tentativa de adivinhar o significado do que está sendo dito teria um resultado pouco confiável.
Limitações:	<ul style="list-style-type: none"> • A qualidade da tradução e reconhecimento será avaliada pelo próprio participante. • O número de participantes é baixo. Apenas quatro. • A simulação envolve a leitura de frases, não sendo avaliado o 	

	<p>desempenho da tecnologia em cenários simulando conversas entre um ou mais participantes. Optou-se pela leitura de frases ao invés de um diálogo real para simplificar a simulação dado o fato de a tecnologia ser bastante recente e não ter sido encontrado nenhum estudo anterior.</p> <ul style="list-style-type: none"> • O vocabulário utilizado não mistura palavras com idioma estrangeiro, o que não condiz com a realidade das equipes de desenvolvimento de software atuais.
Estratégia	
<p>Características do texto a ser reconhecido:</p>	<ul style="list-style-type: none"> • Para a simulação será utilizado um conjunto de 20 frases compostas da seguinte forma: <ul style="list-style-type: none"> ○ 5 frases de conteúdo simples extraídas do experimento de tradução de texto realizado em conjunto por esse grupo de pesquisa [CAL11, CAL12]. ○ 5 frases de estrutura mais complexa extraídas do experimento de tradução de texto realizado em conjunto por esse grupo de pesquisa [CAL11, CAL12]. ○ 5 frases de estrutura simples elaboradas pelo pesquisador envolvendo termos de engenharia de software. ○ 5 frases de estrutura mais complexa elaboradas pelo pesquisador envolvendo termos de engenharia de software.
<p>Regras relacionadas a leitura das frases:</p>	<ul style="list-style-type: none"> • Cada das frases deve ser lida em tom de voz normal e velocidade normal. • A leitura da frase deve ser contínua e não pausada (palavra por palavra). Cada palavra deve ser dita com todas as letras evitando omissão de partes das palavras. • Se durante a pronúncia das frases alguma palavra for pronunciada errada, será executada a leitura completa da frase novamente. • Todos os participantes utilizarão o mesmo microfone. • Não deve haver ruídos externos onde a simulação estiver sendo realizada.
<p>Ferramentas utilizadas na simulação:</p>	<ul style="list-style-type: none"> • Para o reconhecimento de voz foi utilizado o aplicativo de reconhecimento de voz disponibilizado pelo Google: <ul style="list-style-type: none"> ○ https://www.google.com/intl/en/chrome/demos/speech.html

- Para a tradução do texto foi criado uma aplicação java que recebe uma lista de frases no idioma português brasileiro e retorna a lista traduzida para o inglês. Essa aplicação acessa o Google Translator através de *web services* para realizar a tradução.

Conteúdo a ser reconhecido

Frases original	Origem	Categoria
olá tudo bem?	Experimento	Simples
para mim tambem.	Experimento	simples
não tinha visto.	Experimento	Simples
sim, concordo.	Experimento	Simples
qual o próximo?	Experimento	Simples
mas acho que uma das opções mais importantes da lista é a Agenda de contatos.	Experimento	complexo
vamos fechar a lista dos importantes?	Experimento	complexo
proponho baixar 10% de cada item para sobrar alguns pontos para o final.	Experimento	complexo
Quase ninguém usa controle de voz por isso disse que seria menos importante.	Experimento	complexo
Acho que o tempo já está acabando!!	Experimento	complexo
Esse requisito é prioridade.	Frase criada	Simples
Vou fechar esse defeito amanhã.	Frase criada	Simples
Qual a situação atual do projeto?	Frase criada	Simples
Isso é um defeito no código da aplicação.	Frase criada	Simples
Termine os testes da aplicação.	Frase criada	Simples
O plano de projeto define como o projeto é executado, controlado, monitorado e encerrado.	Frase criada	Complexo
Os testes da aplicação devem levar pelo menos uma semana e precisarão de no mínimo duas pessoas alocadas.	Frase criada	Complexo
Nesta iteração a equipe deve focar no fechamento de defeitos e não nas implementações de novas histórias de usuário.	Frase criada	Complexo
Cada funcionalidade deverá ser quebrada em várias histórias de usuário menores que serão priorizadas durante as iterações do projeto.	Frase criada	Complexo
O código está pronto, porém os testes da funcionalidade ainda não foram terminados.	Frase criada	Complexo

Execução

Fluxo de execução da simulação:	<p>Para cada participante, o pesquisador deve executar os seguintes passos:</p> <ul style="list-style-type: none"> • Abrir a aplicação web para realizar o reconhecimento de voz usando o Web Speech API: <ul style="list-style-type: none"> ◦ https://www.google.com/intl/en/chrome/demos/speech.html • Selecionar idioma português do brasil. • Para cada frase, o pesquisador deve fazer o seguinte:
---------------------------------	--

- Clicar no símbolo “microfone” para iniciar o reconhecimento de voz.
- Solicitar ao participante que leia a próxima frase
 - Se houve erro na pronuncia por parte dos participantes, o pesquisador deverá solicitar que a frase seja lida novamente.
- Clicar no símbolo “microfone” para encerrar o reconhecimento.
- Armazenar o texto resultante do reconhecimento.
- Neste momento, o participante pode deixar a simulação.
- Após todas as frases terem sido reconhecidas. O pesquisador deverá submeter a lista de frases reconhecidas a aplicação de tradução e armazenar os resultados da tradução.
- Por email, o pesquisador deve solicitar que o participante avalie a qualidade do reconhecimento em relação ao original e qualidade da tradução para o inglês também em relação ao original seguindo os critérios de avaliação de qualidade definidos anteriormente. O email incluirá:
 - Nome do participante.
 - Descrição sobre como deve ser avaliado a qualidade do reconhecimento de voz.
 - Descrição sobre como deve ser avaliado a qualidade da tradução do texto.
 - Tabela com os resultados do reconhecimento de voz contendo as seguintes colunas:
 - (1) Texto original em português brasileiro; (2) Texto reconhecido em português brasileiro; (3) Local para o participante entrar com a avaliação dele sobre a qualidade do reconhecimento.
 - Tabela com os resultados da tradução de texto contendo as seguintes colunas:
 - (1) Texto original em português brasileiro; (2) Texto traduzido para o inglês; (3) Local para o participante entrar com a avaliação dele sobre a qualidade da tradução.

ANEXO 5 – Google: Resultado da simulação

Resultados da avaliação de reconhecimento de voz realizada com o Google Web Speech API.

Origem	Categoria	Frases original	ASR - Reconhecido	Avaliação	Participante
Experimento	Simples	olá tudo bem?	Olá tudo bem	4	Estudante 1
Experimento	simples	para mim também.	Para mim também	4	Estudante 1
Experimento	Simples	não tinha visto.	Não tinha visto	4	Estudante 1
Experimento	Simples	sim, concordo.	Sim concordo	4	Estudante 1
Experimento	Simples	qual o próximo?	Qual o próximo	4	Estudante 1
Experimento	complexo	mas acho que uma das opções mais importantes da lista é a Agenda de contatos.	Mas acho que uma das opções mais importantes da vida é agenda de contatos	3	Estudante 1
Experimento	complexo	vamos fechar a lista dos importantes?	Vamos fechar a lista dos importantes	4	Estudante 1
Experimento	complexo	proponho baixar 10% de cada item para sobrar alguns pontos para o final.	Proponho baixar 10 por cento de cadeira para alguns	2	Estudante 1
Experimento	complexo	Quase ninguém usa controle de voz por isso disse que seria menos importante.	Quase ninguém no controle de voz por isso de ser importante	2	Estudante 1
Experimento	complexo	Acho que o tempo já está acabando!!	Acho que o tempo está acabando	4	Estudante 1
Frase criada	Simples	Esse requisito é prioridade.	Esse requisito é prioridade	4	Estudante 1
Frase criada	Simples	Vou fechar esse defeito amanhã.	Você já viu esse defeito amanhã	3	Estudante 1
Frase criada	Simples	Qual a situação atual do projeto?	Qual a situação atual do projeto	4	Estudante 1
Frase criada	Simples	Isso é um defeito no código da aplicação.	Isso é um defeito no colégio	2	Estudante 1
Frase criada	Simples	Termine os testes da aplicação.	Hermínio staff ligação	1	Estudante 1
Frase criada	Complexo	O plano de projeto define como o projeto é executado, controlado, monitorado e encerrado.	O plano de projeto define como deve estar do lado de fora do inter	2	Estudante 1
Frase criada	Complexo	Os testes da aplicação devem levar pelo menos uma semana e precisarão de no mínimo duas pessoas alocadas.	Os testes aplicação de levar pelo menos uma semana e precisaram de no mínimo duas pessoas	3	Estudante 1

Frase criada	Complexo	Nesta iteração a equipe deve focar no fechamento de defeitos e não nas implementações de novas histórias de usuário.	Net são as cidades se eu não nasci de novo história de usuário	1	Estudante 1
Frase criada	Complexo	Cada funcionalidade deverá ser quebrada em várias histórias de usuário menores que serão priorizadas durante as iterações do projeto.	Cidade de vera menores que serão realizadas durante as ações do projeto	1	Estudante 1
Frase criada	Complexo	O código está pronto, porém os testes da funcionalidade ainda não foram terminados.	O poder está pronto porém ainda não foram terminadas	1	Estudante 1
Experimento	Simple	olá tudo bem?	Olá tudo bem	4	Estudante 2
Experimento	simple	para mim também.	Para mim também	4	Estudante 2
Experimento	Simple	não tinha visto.	Não tinha visto	4	Estudante 2
Experimento	Simple	sim, concordo.	Sim concordo	4	Estudante 2
Experimento	Simple	qual o próximo?	Qual o próximo	4	Estudante 2
Experimento	complexo	mas acho que uma das opções mais importantes da lista é a Agenda de contatos.	Mas acho que uma das funções mais importantes da lista de agenda de contatos	3	Estudante 2
Experimento	complexo	vamos fechar a lista dos importantes?	Vamos fechar a lista dos mais importantes	4	Estudante 2
Experimento	complexo	proponho baixar 10% de cada item para sobrar alguns pontos para o final.	Proponho baixar 10 por cento de cada item para sobrar alguns pontos ponto final	4	Estudante 2
Experimento	complexo	Quase ninguém usa controle de voz por isso disse que seria menos importante.	Quase ninguém usar controle de voz por isso disse que seria menos importante	4	Estudante 2
Experimento	complexo	Acho que o tempo já está acabando!!	Acho que o tempo está acabando	4	Estudante 2
Frase criada	Simple	Esse requisito é prioridade.	Existe perda de	1	Estudante 2
Frase criada	Simple	Vou fechar esse defeito amanhã.	Vou fechar esse defeito amanhã	4	Estudante 2
Frase criada	Simple	Qual a situação atual do projeto?	Qual situação atual do projeto	4	Estudante 2
Frase criada	Simple	Isso é um defeito no código da aplicação.	Isso é um defeito código da atração	4	Estudante 2
Frase criada	Simple	Termine os testes da aplicação.	Ministério da educação	1	Estudante 2
Frase criada	Complexo	O plano de projeto define como o projeto é executado, controlado, monitorado e encerrado.	O plano de projeto define como projeto executado controlado monitorado encerrado	3	Estudante 2

Frase criada	Complexo	Os testes da aplicação devem levar pelo menos uma semana e precisarão de no mínimo duas pessoas alocadas.	Sucesso da operação deve levar pelo menos uma semana precisarão de ônibus de sorocaba	2	Estudante 2
Frase criada	Complexo	Nesta iteração a equipe deve focar no fechamento de defeitos e não nas implementações de novas histórias de usuário.	Mesa de tração equipe deve focar no fechamento de defeitos e não implementação de novos gestores usuário	2	Estudante 2
Frase criada	Complexo	Cada funcionalidade deverá ser quebrada em várias histórias de usuário menores que serão priorizadas durante as iterações do projeto.	Funcionalidade deverá ser quebrado em várias histórias dos menores serão realizadas durante as gravações do projeto	3	Estudante 2
Frase criada	Complexo	O código está pronto, porém os testes da funcionalidade ainda não foram terminados.	Códigos prontos teste de funcionalidade ainda não foram treinados	2	Estudante 2
Experimento	Simples	olá tudo bem?	Olá tudo bem	4	Estudante 3
Experimento	simples	para mim também.	Para mim também	4	Estudante 3
Experimento	Simples	não tinha visto.	Não tinha visto	4	Estudante 3
Experimento	Simples	sim, concordo.	Sim concordo	4	Estudante 3
Experimento	Simples	qual o próximo?	Qual o próximo	4	Estudante 3
Experimento	complexo	mas acho que uma das opções mais importantes da lista é a Agenda de contatos.	Mas acho que uma das opções mais importantes da lista de agenda de contatos	3	Estudante 3
Experimento	complexo	vamos fechar a lista dos importantes?	Vamos fechar a lista dos importantes	4	Estudante 3
Experimento	complexo	proponho baixar 10% de cada item para sobrar alguns pontos para o final.	Proponho baixar 10 por cento de cada item para sobrar algum ponto final	3	Estudante 3
Experimento	complexo	Quase ninguém usa controle de voz por isso disse que seria menos importante.	Quase ninguém usar controle de voz por isso disse que seria importante	3	Estudante 3
Experimento	complexo	Acho que o tempo já está acabando!!	Acho que o tempo está acabando	3	Estudante 3
Frase criada	Simples	Esse requisito é prioridade.	Escrita por idade	1	Estudante 3
Frase criada	Simples	Vou fechar esse defeito amanhã.	Vou fechar esse defeito amanhã	4	Estudante 3
Frase criada	Simples	Qual a situação atual do projeto?	Qual a situação atual do projeto	4	Estudante 3
Frase criada	Simples	Isso é um defeito no código da aplicação.	Isso é um defeito código de aplicação	3	Estudante 3
Frase criada	Simples	Termine os testes da aplicação.	Ministério da educação	1	Estudante 3
Frase criada	Complexo	O plano de projeto define como o projeto é executado, controlado, monitorado e encerrado.	Plano de projeto define como projeto executado controlado monitorado encerrado	3	Estudante 3

Frase criada	Complexo	Os testes da aplicação devem levar pelo menos uma semana e precisarão de no mínimo duas pessoas alocadas.	Teste da aplicação deve levar pelo menos uma semana que precisarão de no mínimo duas pessoas bloqueadas	3	Estudante 3
Frase criada	Complexo	Nesta iteração a equipe deve focar no fechamento de defeitos e não nas implementações de novas histórias de usuário.	Nessa interação equipe deve focar no fechamento de defeito e não nas interpretações nova história de usuário	3	Estudante 3
Frase criada	Complexo	Cada funcionalidade deverá ser quebrada em várias histórias de usuário menores que serão priorizadas durante as iterações do projeto.	Cada funcionalidade de braço quebrado em várias histórias de usuário menores que serão realizadas durante as interações do projeto	2	Estudante 3
Frase criada	Complexo	O código está pronto, porém os testes da funcionalidade ainda não foram terminados.	Código está pronto porém essa funcionalidade ainda não foram internados	2	Estudante 3
Experimento	Simples	olá tudo bem?	Olá tudo bem	3	Estudante 4
Experimento	simples	para mim também.	Para mim também	4	Estudante 4
Experimento	Simples	não tinha visto.	Não tinha visto	4	Estudante 4
Experimento	Simples	sim, concordo.	Sim concordo	3	Estudante 4
Experimento	Simples	qual o próximo?	Qual próximo	2	Estudante 4
Experimento	complexo	mas acho que uma das opções mais importantes da lista é a Agenda de contatos.	Mas acho que uma das pessoas mais importantes da lista de contatos	2	Estudante 4
Experimento	complexo	vamos fechar a lista dos importantes?	Vamos fechar a lista dos importantes	3	Estudante 4
Experimento	complexo	proponho baixar 10% de cada item para sobrar alguns pontos para o final.	Problemas a deus por cento de cada item para sobrados ponto final	1	Estudante 4
Experimento	complexo	Quase ninguém usa controle de voz por isso disse que seria menos importante.	Quase ninguém controle de voz por isso disse que seria menos importante	3	Estudante 4
Experimento	complexo	Acho que o tempo já está acabando!!	Acho que o tempo está acabando	3	Estudante 4
Frase criada	Simples	Esse requisito é prioridade.	Existe de verdade	1	Estudante 4
Frase criada	Simples	Vou fechar esse defeito amanhã.	Vou fechar esse defeito amanhã	4	Estudante 4
Frase criada	Simples	Qual a situação atual do projeto?	Qual a situação atual do projeto	3	Estudante 4
Frase criada	Simples	Isso é um defeito no código da aplicação.	Isso é um defeito código aplicação	2	Estudante 4
Frase criada	Simples	Termine os testes da aplicação.	Terminal aplicação	1	Estudante 4
Frase criada	Complexo	O plano de projeto define como o projeto é executado, controlado, monitorado e encerrado.	O plano de projeto define como é descontado controlado monitorar encerrado	2	Estudante 4

Frase criada	Complexo	Os testes da aplicação devem levar pelo menos uma semana e precisarão de no mínimo duas pessoas alocadas.	Aplicação deve levar pelo menos uma semana episódio 1 de no mínimo duas pessoas atacadas	1	Estudante 4
Frase criada	Complexo	Nesta iteração a equipe deve focar no fechamento de defeitos e não nas implementações de novas histórias de usuário.	Neste diapasão equipe focada no fechamento de defeitos nos braços de novo dicionário	1	Estudante 4
Frase criada	Complexo	Cada funcionalidade deverá ser quebrada em várias histórias de usuário menores que serão priorizadas durante as iterações do projeto.	Funcionalidade deverá ser quebrado em várias histórias 18 anos que não precisa de interações do projeto	1	Estudante 4
Frase criada	Complexo	O código está pronto, porém os testes da funcionalidade ainda não foram terminados.	O código está pronto teste de funcionalidade não foram terminadas	2	Estudante 4

Resultados da avaliação da tradução realizada utilizando o Google Translator e tendo como entrada as frases reconhecidas pelo Google Web Speech API.

Origem	Categoria	Frases original	MT - Traduzido para Inglês	Avaliação	Participante
Experimento	Simple	olá tudo bem?	Hello all well	4	Estudante 1
Experimento	simple	para mim tambem.	Me too	4	Estudante 1
Experimento	Simple	não tinha visto.	Had not seen	3	Estudante 1
Experimento	Simple	sim, concordo.	Yes I agree	4	Estudante 1
Experimento	Simple	qual o próximo?	What next	4	Estudante 1
Experimento	complexo	mas acho que uma das opções mais importantes da lista é a Agenda de contatos.	But I think one of the most important choices in life is to Phonebook	3	Estudante 1
Experimento	complexo	vamos fechar a lista dos importantes?	Let's close the list of important	3	Estudante 1
Experimento	complexo	proponho baixar 10% de cada item para sobrar alguns pontos para o final.	I propose to download chair 10 percent for some	2	Estudante 1
Experimento	complexo	Quase ninguem usa controle de voz por isso disse que seria menos importante.	Almost nobody in the voice control so it is important	2	Estudante 1
Experimento	complexo	Acho que o tempo já está acabando!!	I think time is running out	4	Estudante 1

Frase criada	Simple	Esse requisito é prioridade.	This requirement is priority	4	Estudante 1
Frase criada	Simple	Vou fechar esse defeito amanhã.	Have you seen this defect tomorrow	2	Estudante 1
Frase criada	Simple	Qual a situação atual do projeto?	What is the status of the project	4	Estudante 1
Frase criada	Simple	Isso é um defeito no código da aplicação.	This is a defect in the college	2	Estudante 1
Frase criada	Simple	Termine os testes da aplicação.	Herminio staff liaison	1	Estudante 1
Frase criada	Complexo	O plano de projeto define como o projeto é executado, controlado, monitorado e encerrado.	The project plan defines how it should be outside of the inter	1	Estudante 1
Frase criada	Complexo	Os testes da aplicação devem levar pelo menos uma semana e precisarão de no mínimo duas pessoas alocadas.	The tests applied to take at least a week and required at least two people	3	Estudante 1
Frase criada	Complexo	Nesta iteração a equipe deve focar no fechamento de defeitos e não nas implementações de novas histórias de usuário.	Net are the cities if I was not born again user story	1	Estudante 1
Frase criada	Complexo	Cada funcionalidade deverá ser quebrada em várias histórias de usuário menores que serão priorizadas durante as iterações do projeto.	City of smaller vera that will be undertaken during project actions	1	Estudante 1
Frase criada	Complexo	O código está pronto, porém os testes da funcionalidade ainda não foram terminados.	The power is ready but not yet completed	1	Estudante 1
Experimento	Simple	olá tudo bem?	Hello all well	2	Estudante 2
Experimento	simple	para mim tambem.	Me too	4	Estudante 2
Experimento	Simple	não tinha visto.	Had not seen	2	Estudante 2
Experimento	Simple	sim, concordo.	Yes I agree	4	Estudante 2
Experimento	Simple	qual o próximo?	What next	4	Estudante 2
Experimento	complexo	mas acho que uma das opções mais importantes da lista é a Agenda de contatos.	But I think one of the most important functions of the list of phonebook	2	Estudante 2
Experimento	complexo	vamos fechar a lista dos importantes?	Let's close the list of the most important	3	Estudante 2
Experimento	complexo	proponho baixar 10% de cada item para sobrar alguns pontos para o final.	I suggest downloading 10 percent of each item to spare some points endpoint	3	Estudante 2
Experimento	complexo	Quase ninguem usa controle de voz por isso disse que seria menos importante.	Almost nobody uses voice control so he said it would be less important	4	Estudante 2

Experimento	complexo	Acho que o tempo já está acabando!!	I think time is running out	4	Estudante 2
Frase criada	Simple	Esse requisito é prioridade.	There is loss of	2	Estudante 2
Frase criada	Simple	Vou fechar esse defeito amanhã.	I'll close this defect tomorrow	4	Estudante 2
Frase criada	Simple	Qual a situação atual do projeto?	What is the current status of the project	4	Estudante 2
Frase criada	Simple	Isso é um defeito no código da aplicação.	This is a fault code of attraction	1	Estudante 2
Frase criada	Simple	Termine os testes da aplicação.	Ministry of Education	1	Estudante 2
Frase criada	Complexo	O plano de projeto define como o projeto é executado, controlado, monitorado e encerrado.	The project plan defines how the project performed controlled monitored closed	3	Estudante 2
Frase criada	Complexo	Os testes da aplicação devem levar pelo menos uma semana e precisarão de no mínimo duas pessoas alocadas.	Success of the operation will take at least a week need bus sorocaba	1	Estudante 2
Frase criada	Complexo	Nesta iteração a equipe deve focar no fechamento de defeitos e não nas implementações de novas histórias de usuário.	Table traction team should focus on closure of defects and failure to implement new user managers	2	Estudante 2
Frase criada	Complexo	Cada funcionalidade deverá ser quebrada em várias histórias de usuário menores que serão priorizadas durante as iterações do projeto.	Functionality should be broken into several stories of minors will be held during the recording project	3	Estudante 2
Frase criada	Complexo	O código está pronto, porém os testes da funcionalidade ainda não foram terminados.	Functionality testing ready codes have not been trained	3	Estudante 2
Experimento	Simple	olá tudo bem?	Hello all well	3	Estudante 3
Experimento	simple	para mim tambem.	Me too	4	Estudante 3
Experimento	Simple	não tinha visto.	Had not seen	3	Estudante 3
Experimento	Simple	sim, concordo.	Yes I agree	4	Estudante 3
Experimento	Simple	qual o próximo?	What next	3	Estudante 3
Experimento	complexo	mas acho que uma das opções mais importantes da lista é a Agenda de contatos.	But I think one of the most important options from the list of contacts book	3	Estudante 3
Experimento	complexo	vamos fechar a lista dos importantes?	Let's close the list of important	4	Estudante 3
Experimento	complexo	proponho baixar 10% de cada item para sobrar alguns pontos para o final.	I suggest downloading 10 percent of each item to spare some endpoint	2	Estudante 3

Experimento	complexo	Quase ninguem usa controle de voz por isso disse que seria menos importante.	Almost nobody uses voice control so said it was important	3	Estudante 3
Experimento	complexo	Acho que o tempo já está acabando!!	I think time is running out	3	Estudante 3
Frase criada	Simples	Esse requisito é prioridade.	Written by age	1	Estudante 3
Frase criada	Simples	Vou fechar esse defeito amanhã.	I'll close this defect tomorrow	4	Estudante 3
Frase criada	Simples	Qual a situação atual do projeto?	What is the status of the project	4	Estudante 3
Frase criada	Simples	Isso é um defeito no código da aplicação.	This is a faulty application code	4	Estudante 3
Frase criada	Simples	Termine os testes da aplicação.	Ministry of Education	1	Estudante 3
Frase criada	Complexo	O plano de projeto define como o projeto é executado, controlado, monitorado e encerrado.	Project plan defines how the project performed controlled monitored closed	3	Estudante 3
Frase criada	Complexo	Os testes da aplicação devem levar pelo menos uma semana e precisarão de no minimo duas pessoas alocadas.	Test the application will take at least a week that will require a minimum of two people locked	3	Estudante 3
Frase criada	Complexo	Nesta iteração a equipe deve focar no fechamento de defeitos e não nas implementações de novas histórias de usuário.	In this interaction team should focus on closing the defect and not the interpretations new user story	3	Estudante 3
Frase criada	Complexo	Cada funcionalidade deverá ser quebrada em várias histórias de usuário menores que serão priorizadas durante as iterações do projeto.	Each feature broken arm in several smaller user stories that will be performed during the design iterations	2	Estudante 3
Frase criada	Complexo	O código está pronto, porém os testes da funcionalidade ainda não foram terminados.	Code is ready but this functionality has not yet been admitted	3	Estudante 3
Experimento	Simples	olá tudo bem?	Hello all well	2	Estudante 4
Experimento	simples	para mim tambem.	Me too	4	Estudante 4
Experimento	Simples	não tinha visto.	Had not seen	2	Estudante 4
Experimento	Simples	sim, concordo.	Yes I agree	4	Estudante 4
Experimento	Simples	qual o próximo?	What next	3	Estudante 4
Experimento	complexo	mas acho que uma das opções mais importantes da lista é a Agenda de contatos.	But I think one of the most important people in the contact list	1	Estudante 4
Experimento	complexo	vamos fechar a lista dos importantes?	Let's close the list of important	4	Estudante 4

Experimento	complexo	proponho baixar 10% de cada item para sobrar alguns pontos para o final.	Problems god percent of each item to houses endpoint	1	Estudante 4
Experimento	complexo	Quase ninguem usa controle de voz por isso disse que seria menos importante.	Almost no voice control so he said it would be less important	2	Estudante 4
Experimento	complexo	Acho que o tempo já está acabando!!	I think time is running out	4	Estudante 4
Frase criada	Simple	Esse requisito é prioridade.	There is real	1	Estudante 4
Frase criada	Simple	Vou fechar esse defeito amanhã.	I'll close this defect tomorrow	4	Estudante 4
Frase criada	Simple	Qual a situação atual do projeto?	What is the status of the project	4	Estudante 4
Frase criada	Simple	Isso é um defeito no código da aplicação.	This application is a code defect	2	Estudante 4
Frase criada	Simple	Termine os testes da aplicação.	Terminal application	1	Estudante 4
Frase criada	Complexo	O plano de projeto define como o projeto é executado, controlado, monitorado e encerrado.	The project plan defines how controlled monitor discounted ended	2	Estudante 4
Frase criada	Complexo	Os testes da aplicação devem levar pelo menos uma semana e precisarão de no minimo duas pessoas alocadas.	Application must take at least one week episode 1 of at least two persons attacked	1	Estudante 4
Frase criada	Complexo	Nesta iteração a equipe deve focar no fechamento de defeitos e não nas implementações de novas histórias de usuário.	In this pitch team focused on closing defects in the arms of new dictionary	1	Estudante 4
Frase criada	Complexo	Cada funcionalidade deverá ser quebrada em várias histórias de usuário menores que serão priorizadas durante as iterações do projeto.	Functionality should be broken into several stories 18 who does not need design iterations	1	Estudante 4
Frase criada	Complexo	O código está pronto, porém os testes da funcionalidade ainda não foram terminados.	The code is ready functionality test were not completed	3	Estudante 4