

Interpretação automática ou tradução automática de fala: conceitos, definições e arquitetura de *software*

Automatic interpretation or automatic translation of speech: concepts, definitions and software architecture

Flávio de Sousa Freitas*
Marileide Dias Esqueda**

Resumo: A interpretação automática (IA), ou tradução automática de fala, é uma tecnologia que traduz discurso oral de uma língua para outra através de três funcionalidades acopladas em um único sistema computacional: reconhecimento automático de fala, tradução automática e síntese de fala. Apresentado pela primeira vez em 1983, durante a convenção ITU Telecom, em Genebra, o conceito de IA veicula a ideia de sistemas capazes de promover a comunicação entre pessoas que falam línguas diferentes de forma espontânea e eficaz (LEE 2015). A IA e seus

* Graduando pela Universidade Federal de Uberlândia/MG.

** Docente da Universidade Federal de Uberlândia/MG.

sistemas, todavia, ainda são escassamente estudados, especialmente na área de Estudos de Interpretação, sendo difundidos, em sua maioria, na área de Ciências da Computação (PÖCHHACKER 2004). Desse modo, o objetivo geral deste trabalho é apresentar os resultados de uma pesquisa bibliográfico-documental, que teve como proposta investigar como a IA é concebida pelos teóricos que a estudam.

Palavras-chave: Interpretação Automática; tradução automática de fala; tecnologias de interpretação automática.

Abstract: Machine Interpreting (MI) or Speech-to-Speech Translation (S2S) is a new technology that converts spoken utterances from one language into another through three functionalities grouped in only one software: Automatic Speech Recognition (ASR), Machine Translation (MT) and Speech Synthesis (or TTS - Text-To-Speech). Firstly presented in 1983, during ITU Telecom Convention, in Geneva, MI consists in a software that enables natural and efficient communication between people from different languages (LEE, 2015). Nevertheless, MI and its components are scarcely studied, especially in Interpreting Studies, generally investigated by the areas of Computer Sciences (PÖCHHACKER 2004). Following this trend, this paper aims to present the result of a bibliographical research whose proposal was to investigate how MI is conceived by scholars.

Keywords: Machine Interpreting; Speech-to-Speech Machine Translation; Speech Translation Technology.

1. Introdução

É provável que a maioria das pessoas já tenha ouvido falar em tradução automática (doravante TA), não sendo um exagero afirmar que praticamente todos conhecem ou já usaram, pelo menos uma vez, o Google Tradutor ou outra ferramenta de TA online para traduzir automaticamente algum tipo de texto. O aumento de recursos linguísticos tais como dicionários eletrônicos e *corpora* paralelos, aliados às tecnologias da informação e comunicação, têm contribuído para a popularização e evolução da tecnologia de TA (BARREIRO et al. 2014).

A interpretação automática (doravante IA), por outro lado, ainda não atingiu uma popularidade considerável, apesar de já ter sido tema de várias obras de ficção científica. Na famosa série norte-americana *Star Trek*, o *Universal Translator* é usado para decifrar e interpretar línguas alienígenas através do escaneamento de ondas cerebrais. Outro exemplo é o *Babel Fish*, da série *O guia do mochileiro das galáxias*, um pequeno dispositivo em formato de peixe, que ao ser introduzido no ouvido possibilita que qualquer pessoa compreenda qualquer língua do universo (DECAMP E ZETSCHE 2015).

No mundo real, todavia, a IA, ou tradução automática de fala, e a criação de sistemas computacionais que a operem, cumprindo com todas as exigências de um cenário de interpretação de conferências, ainda é uma incógnita, e tem sido investigada há mais de duas décadas pelas áreas de Ciências da Computação e Inteligência Artificial (PÖCHHACKER 2004).

Para o autor chinês da Universidade de Hong Kong, Lee Tan (2015), a interpretação automática é uma tecnologia que converte um discurso oral de uma língua para outra, possibilitando a comunicação entre pessoas de diferentes línguas, por meio de um *software* que congrega reconhecimento automático de fala, tradução automática e síntese de voz. Lee (2015) explica que, atualmente, espera-se que essa tecnologia avance no sentido de traduzir

oralmente sentenças inteiras, mantendo o ritmo de um discurso natural e contínuo, com alto grau de inteligibilidade e qualidade.

Nos últimos anos, as tecnologias de IA evoluíram para além do meio acadêmico, deixando de ser tema exclusivo de discussões teóricas para serem empregadas em situações em que antes somente a inteligência humana poderia intervir (AMINZADEH E SHEN 2009).

Waibel e Fügen (2008), Nakamura (2009) e vários outros pesquisadores informam que o aumento da importância dada à IA, na atualidade, deve-se ao processo de globalização que a humanidade está vivendo: “A economia global sem fronteiras torna a comunicação entre falantes de diferentes línguas potencialmente importante. A tecnologia de interpretação automática [...] tem sido desde há muito tempo um sonho para a humanidade” (NAKAMURA 2009: 35, tradução nossa)¹.

Liu e Zhang (2015) destacam que há um interesse cada vez maior em tecnologias de comunicação multimodal e multilíngue através dos meios de comunicação social e comunidades online. Essa conjuntura tem permitido o desenvolvimento de sistemas de tradução e interpretação automáticas (TA e IA), inseridos no contexto de um circuito global de tecnologias de processamento de conteúdo inteligente, em cujo centro encontra-se a TA (LIU E ZHANG 2015).

A disponibilidade de tecnologias capazes de realizar traduções de qualidade gera expectativas não só para o presente. Em 2004, por exemplo, um artigo da revista MIT Technology Review elegeu a IA entre as dez tecnologias que mudariam o mundo. Yuqing Gao (2004) estimava que no prazo de dez anos as tecnologias de tradução universal estariam presentes em dispositivos móveis e celulares ao alcance de qualquer pessoa.

¹ The global, borderless economy has made it critically important for speakers of different languages to be able to communicate. Speech translation technology [...] has long been a dream of humankind (NAKAMURA 2009: 35).

As previsões de Gao não estavam equivocadas, pois já nos anos 2000 as empresas Google e Microsoft, líderes no mercado global de tecnologias, anunciaram o lançamento de aplicativos gratuitos de IA (CASTELLUCCIO 2014; RAWLINSON 2015).

Assim como tem acontecido com a TA, as discussões sobre a IA versam, em sua grande maioria, a respeito da confiabilidade das traduções produzidas por tais tecnologias (BARREIRO et al. 2014). Grande parte dos sistemas disponíveis no mercado trabalha apenas com um determinado número de pares linguísticos e são programados para atuar em domínios especializados (turismo, negócios, saúde etc.).

Por mais que esse tipo de tecnologia possa provocar algum temor aos intérpretes humanos, Jekat e Klein (1996) já mencionavam a importância de se enfatizar que os sistemas de IA não podem, nem pretendem, substituir o intérprete humano. Faz parte desse processo, entretanto, a exploração dos mecanismos pelos quais a interpretação humana se realiza, visto que a comparação entre a tradução humana e a automática é uma das principais exigências para a descrição dos possíveis resultados práticos alcançados por um sistema de IA (JEKAT E KLEIN 1996).

Nesse sentido, não apenas as pesquisas sobre os aspectos teórico-práticos quanto à IA encontram-se em construção, mas a própria terminologia que a constitui. Embora tenham sido utilizados no título, e ao longo deste trabalho, os termos “Interpretação Automática”, percebe-se que ainda não parece haver uma terminologia consistente que a conceitue e defina. Pesquisas junto aos mais diversos catálogos de artigos, teses, dissertações e livros, tanto em meio impresso quanto eletrônico, revelam a baixa incidência de trabalhos científicos sobre a IA em contexto brasileiro. Como grande parte das publicações se encontram atreladas a países da União Europeia, EUA e Japão, propõe-se aqui implementar um estudo bibliográfico sobre a IA que poderá estabelecer diálogos entre teorias ocidentais e orientais (LEE 2015), vislumbrando seu possível estudo também em cenário brasileiro. Quando

trazidas ao âmbito nacional, as discussões podem contribuir para o fenômeno de expansão teórica da IA para além das instituições de pesquisa do exterior (AMINZADEH E SHEN 2009), lançando as bases para estudos futuros que abordem os impactos da tecnologia no trabalho do intérprete profissional.

A partir do exposto, o objetivo geral deste trabalho é realizar um levantamento bibliográfico sobre a IA, com vistas a compreender como os teóricos a concebem e descrevem no cenário atual. Com base no exame bibliográfico, construir-se-á um inventário de candidatos a termos (FINATTO 2002; FROMM 2005) que a denominam e conceituam, para seu possível uso em português do Brasil.

Espera-se que este estudo possa converter-se nos primeiros passos para a futura construção do estado da arte da IA, ainda inexistente no Brasil.

2. Procedimentos metodológicos

Tendo por base o objetivo de realizar um levantamento bibliográfico sobre a IA, a metodologia utilizada neste trabalho é a pesquisa bibliográfico-documental, de caráter inventariante e descritivo. Esse tipo de pesquisa, no campo das ciências humanas, incluindo dissertações de mestrado, teses de doutorado, publicações em periódicos e comunicações em anais de congressos e de seminários, vem sendo bastante difundida e utilizada nas últimas duas décadas (FERREIRA 2002).

A pesquisa bibliográfica tenta responder quais aspectos e dimensões vêm sendo contemplados pelas investigações científicas de determinada área do conhecimento em diferentes épocas e lugares, as diferentes formas com que o objeto de estudo é abordado e em que condições essas pesquisas se desenvolvem (FERREIRA 2002).

Na obra *Introducing Interpreting Studies*, Pöchhacker (2004) cita a IA como uma linha de pesquisa voltada para tecnologias futuristas e faz diversas referências a três termos que foram utilizados nesta pesquisa como palavras-chave aplicadas aos mecanismos de busca selecionados: *automatic interpreting/interpretation*, *machine interpreting/interpretation* e *speech-to-speech machine translation*.

Esses termos citados por Pöchhacker (2004) foram traduzidos para o português pelos autores deste trabalho e aplicados aos mecanismos de busca das bases dos periódicos pertencentes aos bancos de pesquisa online da CAPES, SciELO e Google Acadêmico, a fim de encontrar possíveis fontes em português. Essa etapa, todavia, resultou em apenas dois trabalhos em português brasileiro, a dissertação de Duarte (2014) e o artigo de Pinto e Prikladnicki (2015), dois trabalhos em português europeu, a dissertação de Gusmão (2008), e o artigo de Barreiro et al. (2014), do instituto INESC-ID, de Lisboa. A partir desses primeiros resultados, verificando-se a baixa incidência de pesquisas sobre IA em português, o foco passou a ser as produções científicas relativas ao tema em outros países, como uma motivação a mais para realizar um estudo do tipo bibliográfico, voltado para uma temática ainda pouco explorada no Brasil.

Pöchhacker (2004) cita o artigo *Machine interpretation: Open Problems and Some Solutions*, de Jekat e Klein (1996), publicado no periódico *Interpreting (International Journal of Research and Practice in Interpreting*, da editora John Benjamins) como sendo um dos trabalhos inaugurais sobre a temática. A partir da leitura inicial de Pöchhacker (2004) e tomando por base a técnica de pesquisa *snowball*, citada por Hale e Napier (2013), contatou-se via e-mail a pesquisadora austríaca Alexandra Klein, coautora do artigo, que concedeu uma cópia impressa do seu trabalho, enviada por correio terrestre, uma vez que seu artigo não se encontra disponível na internet.

A partir do artigo de Jekat e Klein (1996), o acesso a uma série de novas referências bibliográficas concernentes à IA fez expandir a lista de

palavras-chave, que passou a conter *speech translation* e *spoken language translation*, resultando num total de cinco palavras-chave aplicadas aos bancos de pesquisa.

Estabelecidas as palavras-chave, deu-se início à elaboração do *corpus* de dissertações de mestrado, teses de doutorado, publicações em periódicos e comunicações em anais de congressos e de seminários. Inicialmente, o primeiro banco de pesquisa selecionado foi o Google Acadêmico, ferramenta que oferece acesso a bancos de dados e bibliotecas digitais sem que seja necessário que o usuário acesse esses recursos diretamente. Obteve-se então acesso aos primeiros 70 trabalhos científicos que compõem o *corpus* desta pesquisa.

Além do Google Acadêmico, aplicou-se as palavras-chave aos mecanismos de busca da biblioteca digital IEEE Xplore², citada por Lee (2015), estabelecendo-se assim o segundo banco de pesquisa de relevância. A partir deste momento, descartou-se a palavra-chave *automatic interpreting/interpretation*, já que os resultados obtidos nas pesquisas junto à biblioteca digital IEEE Xplore referiam-se à tecnologia de interpretação automática de imagens e não à interpretação automática de fala.

Segundo Pöchhacker (2004), as pesquisas sobre o uso de tecnologia na interpretação se deu a partir dos anos 80, quando foram desenvolvidos sistemas de interpretação por telefone e tele interpretação. A partir desse dado, definiu-se realizar a busca para a pesquisa junto à biblioteca digital IEEE Xplore referente ao período de 1980 a 2015. Dessa forma obteve-se acesso a mais 141 trabalhos científicos.

² A biblioteca digital IEEE Xplore oferece acesso a conteúdos técnico-científicos publicados pelo instituto norte-americano IEEE (Institute of Electrical and Electronics Engineers) e seus parceiros. Segundo informações do portal, a biblioteca digital abriga mais de três milhões de documentos na íntegra, provenientes das publicações mais citadas da Engenharia Elétrica, Ciência da Computação e Eletrônica. A Universidade Federal de Uberlândia tem livre acesso aos documentos da IEEE Xplore. Qualquer aluno ou docente que acesse a biblioteca digital a partir da rede de internet da UFU pode realizar o *download* dos artigos sem ter que pagar a taxa para usuários com assinatura privada.

FREITAS, F. de S.; ESQUEDA, M. D. — Interpretação automática ou tradução automática de fala: conceitos, definições e arquitetura de *software*

A partir da leitura da maior parte desses trabalhos, a relevância das pesquisas realizadas pela Universidade de Carnegie Mellon, nos Estados Unidos, tornou-se evidente. Retomando a técnica de pesquisa *snowball* (HALE E NAPIER 2013), contatou-se via e-mail o professor Alex Waibel, um dos principais pesquisadores de IA, responsável pelo desenvolvimento de diversos sistemas e abordagens de IA. Além de conceder uma cópia digital do artigo *Spoken Language Translation - Enabling Cross-Lingual Human-Human Communication* (WAIBEL E FÜGEN 2008), a equipe do professor Waibel enviou o endereço eletrônico do portal interACT, do laboratório *The Interactive Systems Lab*, da Universidade de Carnegie Mellon e do Instituto de Tecnologia de Karlsruhe, da Alemanha.

O contato com o professor Waibel foi de extrema importância para esta pesquisa, pois possibilitou o acesso ao interACT³, o terceiro banco de pesquisa estabelecido para a elaboração do *corpus*, através do qual obteve-se acesso a 53 trabalhos científicos.

Segundo Pöchhacker (2004), embora a IA não seja do interesse de intérpretes profissionais, que apostam na impossibilidade de desempenho dos sistemas, pesquisadores da Ciência da Computação, Linguística, Processamento de Fala e Inteligência Artificial têm se esforçado para a implementação de sistemas dessa natureza. Com base nessa informação, o quarto banco de pesquisa foi selecionado, a biblioteca digital da Association for Computing Machinery (ACM), Associação de Pesquisadores das Ciências da Computação, também dos Estados Unidos, denominada ACM-DL. Nesse banco de dados, foram encontrados 17 trabalhos científicos concernentes à IA.

Ao passo que o *corpus* era elaborado, realizou-se uma busca no referencial bibliográfico de todo o material levantado, com vistas a encontrar a produção científica mais antiga sobre a IA. Em *Automatic Speech Translation at ATR*, Morimoto e Kurematsu (1993) citam o artigo *A Speech*

³ <http://isl.anthropomatik.kit.edu/cmu-kit/english/index.php>.

Driven Language Translation System, de Stentiford e Steer (1987). Contatou-se via e-mail o professor Fred Stentiford, da University College London, da Inglaterra, que concedeu uma cópia digital do artigo *A speech driven language translation system*, de 1987, e outro de 1988, com o título *Machine translation of speech*.

A composição final do *corpus* de pesquisa, portanto, consta da somatória de todo o material coletado nos bancos de pesquisa, os artigos de Waibel e Fügen (2008), Stentiford e Steer (1987, 1988), enviados por e-mail e o artigo de Jekat e Klein (1996), enviado por correio terrestre. O Quadro 1 resume, portanto, a composição do *corpus* de pesquisa.

Período	Bases de dados	N°
1987-2015	Google Acadêmico	70
	IEEE Xplore	141
	interACT	53
	ACM-DL	17
	E-mail	3
	Correio terrestre	1
	Total	285

Quadro 1 - Composição do *corpus* de pesquisa

Como é possível observar, embora o período de 1980 a 2015 tenha sido aplicado às ferramentas de busca das bases de pesquisa consultadas, o resultado obtido compreende o período de 1987 a 2015, o que está relacionado com o fato de a IA ter começado a ganhar força apenas no final da década de 80 (HUTCHINS 2015).

A delimitação do *corpus*, todavia, não significa que os pesquisadores sejam capazes de controlar o seu objeto de investigação. Segundo Ferreira (2002), é ilusório pensar que, se tomar apenas o material encontrado em uma única base de pesquisa, o pesquisador esteja escrevendo a história de

determinada área do conhecimento. No caso da IA, isto seria ainda menos possível, devido à grande quantidade de produções científicas produzidas no Japão, na Alemanha e na China, que por não estarem em inglês, não puderam compor o *corpus* desta pesquisa.

3. Por um inventário de candidatos a termos da IA: conceitos e definições

Finatto (2002: 74) explica que:

O enunciado que define uma noção, processo ou objeto é um elemento-chave na constituição da veiculação do conhecimento especializado, tecnológico ou científico. Afinal, expressa um segmento de relações de significação de uma determinada área do saber. Para compreender a importância do seu papel, basta dizer que muito do que já se escreveu e do que ainda se tem escrito, por exemplo, em Química, depende de definições como as de *átomo* ou de *ácido*. Além disso, definições, ao constituírem textos particularizados, mostram e identificam facetas de compreensão de fenômenos no seio de uma determinada ciência.

Para Finatto (2002), à medida que a Terminologia se aproxima do texto, vemos um movimento de independência e de relativização da tradição lógico-gramatical. A autora argumenta que a terminologia atual considera os elementos culturais, comunicativos, textuais e discursivos presentes nos enunciados, sendo este um passo importante rumo a uma visão mais integradora da comunicação.

Por se tratar de uma pesquisa bibliográfico-documental, de caráter inventariante e descritivo, o vocabulário terminológico a ser apresentado a seguir foi extraído dos textos investigados, nos termos de Finatto (2002), e configura-se em um inventário de candidatos a termos mais frequentes

encontrados na literatura consultada, almejando revelar como os autores de fato denominam a IA e as tecnologias utilizadas para operá-la.

Utiliza-se a denominação inventário de candidatos a termos como um passo inicial e que antecede uma pesquisa terminológica mais apurada, podendo-se futuramente lançar mão de um dos procedimentos metodológicos de construção de um vocabulário técnico explicitados por Fromm (2005), quiçá esclarecendo a profusão terminológica presente na teorização sobre a IA.

Segundo Fügen (2008), a confusão entre tradução e interpretação é recorrente, e em muitas ocasiões a interpretação simultânea é chamada erroneamente de tradução simultânea e o intérprete é chamado de tradutor. Essa flutuação terminológica também está presente nas teorizações sobre IA no interior das discussões da área de Ciências da Computação, já que os termos *machine translation* (MT) ou *automatic translation*, traduzidos para português como tradução automática (TA), são usados tanto para sistemas que traduzem texto quanto para sistemas que traduzem fala. De acordo com o autor, isso se deve ao fato de que, no passado, o principal objeto de estudo da TA era a tradução do texto. A tradução de língua falada (*Spoken Language Translation* - SLT), por sua vez, é mais recente. A fim de diminuir essa confusão, ao referir-se à IA, Fügen (2008) utiliza os termos tradução de fala simultânea (*simultaneous speech translation*) ou simplesmente tradução simultânea (*simultaneous translation*).

Dada a instabilidade terminológica, implementou-se um trabalho de identificação dos termos a partir de suas frequências, que foi realizado através da função “pesquisa avançada”, da ferramenta Adobe Acrobat Pro. Além de possibilitar a leitura dos arquivos de texto salvos em PDF, essa ferramenta oferece várias funções de pesquisas avançadas que permitem a extração de termos e trechos de texto de vários documentos simultaneamente, informando a frequência desses termos em cada documento e na totalidade de documentos pesquisados.

FREITAS, F. de S.; ESQUEDA, M. D. — Interpretação automática ou tradução automática de fala: conceitos, definições e arquitetura de *software*

De acordo com os achados desta pesquisa, os termos mais recorrentes nos 285 trabalhos são *speech translation*, *speech-to-speech translation*, *spoken language translation* e *simultaneous translation*. Termos como *speech driven language translation* e *machine interpreting* acabaram caindo em desuso e o termo *face-to-face dialog translation* é usado unicamente por Wahlster (1993), para descrever o sistema *Verbmobil*. Além dos termos utilizados por Fügen (2008), durante o desenvolvimento desta pesquisa foram encontrados vários outros termos e expressões sinônimas referentes à IA, totalizando 14 termos, conforme mostra o quadro abaixo:

N.	Termos encontrados na literatura	Referência(s)	Recorrência (Número de vezes em que aparece o termo na literatura)
1.	<i>Speech Driven Language Translation</i>	STENTIFORD E STEER (1987)	1
2.	<i>Machine Interpreting</i>	WAHLSTER (2000a)	2
3.	<i>Face-To-Face Dialog Translation</i>	WAHLSTER (1993)	3
4.	<i>Simultaneous Speech Translation</i>	FÜGEN (2008)	7
5.	<i>Automatic Interpreting</i>	JEKAT E KLEIN (1996)	12
6.	<i>Speech-To-Speech Dialog Translation</i>	KITANO (1994)	14
7.	<i>Spoken Language Machine Translation</i>	JIMENEZ, CASTELLANOS E VIDAL (1995)	18
8.	<i>Automatic Interpretation</i>	WATANABE et al. (2000)	34
9.	<i>Spontaneous Speech Translation</i>	MORIMOTO E KUREMATSU (1993)	44
10.	<i>Machine Interpretation</i>	MATSUBARA, TOYAMA, INAGAKI (1999)	66

FREITAS, F. de S.; ESQUEDA, M. D. – Interpretação automática ou tradução automática de fala: conceitos, definições e arquitetura de *software*

11.	<i>Simultaneous Translation</i>	FÜGEN (2008)	144
12.	<i>Spoken Language Translation</i>	KITANO (1991)	656
13.	<i>Speech-To-Speech Translation</i>	TOMITA, TOMABECHI E SAITO (1989)	1211
14.	<i>Speech Translation</i>	TOMITA, TOMABECHI E SAITO (1989); LEE (2015)	2482

Quadro 2 - Inventário de candidatos a termos da área da IA

Além dos 14 termos sinônimos identificados na literatura consultada e explicitados no quadro 2, da menor para a maior recorrência, identificaram-se também outros termos relacionados aos sistemas computacionais que realizam operações de interpretação automática. Tais termos expandem o campo de estudo sobre IA e tratam das tecnologias e procedimentos utilizados para a construção de *softwares*, como mostrado no quadro 3.

N.	Termos relacionados aos sistemas ou aplicações utilizadas durante uma IA	Referência(s)
1.	<i>Bi-directional speech translation system</i>	ISOTANI et al. (2003)
2.	<i>Computer-based speech translation technology</i>	LEE (2015)
3.	<i>Corpus-based speech translation</i>	TAKEZAWA et al. (2002)
4.	<i>Corpus-based paraphrasing system</i>	KIKUI et al. (2006)
5.	<i>Corpus-based statistical machine translation</i>	MRINALINI E VIJAYALAKSHMI (2015)
6.	<i>Domain-limited speech translation systems</i>	WAIBEL E FÜGEN (2008)
7.	<i>Domain-specific system</i>	KITANO (1994); CARTER et al. (1997), HUTCHINS (1997)
8.	<i>Domain-unlimited</i>	WAIBEL E FÜGEN

FREITAS, F. de S.; ESQUEDA, M. D. – Interpretação automática ou tradução automática de fala: conceitos, definições e arquitetura de *software*

	<i>speech translation systems</i>	(2008)
9.	<i>Hybrid Approach to Spoken Language Translation</i>	PAUL et al. (2005)
10.	<i>Interactive spoken dialogue</i>	LEE (2015)
11.	<i>Natural language generator</i>	TOMITA, TOMABECHI E SAITO (1989)
12.	<i>Natural language parsing</i>	GLICKMAN (1995)
13.	<i>Natural language processing</i>	KITANO (1991)
14.	<i>Real-time speech processing system</i>	TOMITA, TOMABECHI E SAITO (1989)
15.	<i>Speech-to-speech machine translation</i>	LEVIN et al. (2000b)
16.	<i>Speech-to-speech translation</i>	PAULIK (2010)
17.	<i>Spontaneous speech translation systems</i>	WAIBEL E FÜGEN (2008)
18.	<i>Statistical model of speech-to-text translation</i>	MATHIAS E BYRNE (2006)
19.	<i>Statistical-based machine translation</i>	FÜGEN (2008)
20.	<i>Text-to-text machine translation</i>	ZHOU et al. (2002)

Quadro 3 - Termos relacionados aos *softwares* de IA

O termo empregado para essa tecnologia em português do Brasil é tradução de voz e de Portugal, tradução de fala, que correspondem ao termo *speech translation* (BARREIRO et al. 2014; DUARTE 2014; PINTO E PRIKLADNICKI 2015). Durante a realização desta pesquisa bibliográfica, todavia, não foram encontradas referências ao termo “interpretação automática” em português brasileiro. O título deste trabalho, portanto, foi adotado com base em Pöchhacker (2004), que, como teórico versado nos Estudos de Interpretação, refere-se à tecnologia de tradução de fala por meio do termo *machine interpreting* (MI), traduzido aqui como interpretação automática (IA).

Embora a IA possa ser identificada por meio de vários termos e expressões diferentes, todos convergem para uma única definição: a IA é uma tecnologia que realiza a conversão automática de trechos de fala de uma dada língua em trechos de fala para outra língua, possibilitando que pessoas que

falam línguas diferentes se comuniquem de forma natural e espontânea (LEE 2015: 619).

Os sistemas de IA podem apresentar-se de várias formas, instalados em servidores de alta performance ou em dispositivos móveis tais como *smartphones*, *laptops*, computadores de bolso e computadores vestíveis (*wearable computers*). (FREDERKING, RUDNICKY E HOGAN 2000; ZHANG 2003). Grande parte desses sistemas, todavia, encontram-se instalados em computadores pessoais de médio desempenho, equipados com placa de áudio e microfone, que realizam a captura do conteúdo produzido pelo falante e reproduzem a fala traduzida para o ouvinte (LEE 2015).

De acordo com Fügen (2008: 1), é possível observar que a evolução dos sistemas de IA passou por diversas fases:

Primeiramente, apenas frases únicas e isoladas podiam ser reconhecidas e traduzidas. As frases tinham que ser pronunciadas de maneira clara e controlada, conforme uma determinada gramática e apenas frases previamente selecionadas podiam ser traduzidas. Nas fases seguintes, aumentaram-se as restrições relativas ao estilo de fala, levando ao surgimento de sistemas de reconhecimento de discurso conversacional e espontâneo. Concomitantemente, aumentaram-se as possibilidades discursivas no que diz respeito a vocabulário e frases permitidas.⁴

Em consequência desses avanços, os sistemas de IA de amplo vocabulário de fala contínua tornaram-se realidade. O que também pode ser observado no âmbito da TA (FÜGEN 2008: 1).

Com o advento da internet e popularização das redes sem fio, os sistemas de IA tornaram-se portáteis. É possível, por exemplo, que os usuários tenham acesso a uma interface móvel que se comunique com um servidor, de

⁴ At first, only single, isolated phrases could be recognized and translated. The phrases had to be spoken in a clean and controlled manner adhering to a predetermined grammar, and only previously seen phrases could be translated. In the next phases, the restrictions on the speaking style were lifted, leading to the emergence of recognition systems for conversational and spontaneous speech. At the same time, the allowed discourse in terms of vocabulary and sentences increased (FÜGEN 2008: 1).

FREITAS, F. de S.; ESQUEDA, M. D. — Interpretação automática ou tradução automática de fala: conceitos, definições e arquitetura de *software*

forma simples e intuitiva, sem que seja preciso que o usuário final tenha conhecimentos de programação para calibrar o sistema ou resolver problemas específicos (ISOTANI et al. 2002; WAIBEL et al. 2003).

O processo de interpretação nesses sistemas envolve três componentes básicos e integrados entre si: reconhecimento automático de fala (ASR - *Automatic Speech Recognition*), tradução automática (TA) e síntese de fala (TTS - *Text-To-Speech*). E ainda que esses componentes possuam suas próprias arquiteturas, no sistema de IA essas arquiteturas encontram-se conectadas. O produto final dessa somatória de componentes, portanto, é um único sistema, a ser comercializado em uma única interface (WAIBEL E FÜGEN 2008).

Há sistemas de tradução de fala, no entanto, que não se encaixam na descrição acima, pois não possuem ora o componente de ASR, ora o componente de TTS. Esses sistemas, portanto, não podem ser considerados sistemas de IA autênticos (ZHANG 2003). O *Phraselator*, desenvolvido pela empresa norte-americana *VoxTec*, por exemplo, não possui o componente de ASR e apesar de reproduzir faixas de áudio pré-gravadas, limita a conversação ao conteúdo disponível nessas faixas pré-gravadas (RICCI 2002; SRI 2009).

Segundo Lee (2015: 619), um sistema de IA hipotético bastante simples poderia funcionar como um “dicionário falante”⁵, que realizaria traduções palavra por palavra, desconsiderando as relações gramaticais e demais propriedades linguísticas das palavras. Tal sistema, todavia, apresentaria um nível tecnológico muito baixo e as possibilidades de aplicá-lo em um cenário de interpretação realístico seriam quase nulas.

Na prática, as expectativas com relação aos sistemas de IA são bem mais exigentes. Para que uma sessão de interpretação se desenvolva de forma confortável tanto para o ouvinte quanto para o conferencista, em uma conferência, por exemplo, é necessário que a interpretação seja feita em

⁵ *Speaking dictionary* (LEE 2015: 619).

tempo real e com alta qualidade sonora, com o mínimo possível de erros, para que toda a sessão não seja comprometida (FÜGEN 2008).

Para que isso se dê, é necessário que o sistema seja capaz de traduzir no nível da sentença e lide com fala contínua e espontânea, de modo que o produto final soe natural, inteligível e com a qualidade semelhante à da voz humana (LEE 2015).

Após os primeiros testes para a viabilização da IA, os pesquisadores começaram a explorar meios para tornar os sistemas e protótipos em tecnologia utilizável, ao alcance dos usuários comuns. Um dos meios encontrados foi a especialização dos sistemas em determinados domínios, pois como afirma Grazina (2010: 5, tradução nossa):

Embora os seres humanos falem de forma espontânea e, geralmente, não consigam lidar com a imposição de restrições linguísticas, há várias tarefas do dia-a-dia em que os domínios discursivos são inerentemente limitados. Ao explorar esses limites, os investigadores esperavam criar uma forma prática de usar a tradução automática de fala.⁶

A especialização dos sistemas de IA em domínios faz com que a tecnologia alcance maiores margens de equivalência, além de tornar o processamento de dados mais rápido e seguro. A IA tem que estar o mais próximo possível da interpretação humana e a tecnologia tem que ser capaz de traduzir de forma simultânea, mantendo o mesmo fluxo da fala humana (TAKEZAWA E KIKUI 2004; PAULIK E WAIBEL 2008).

Da teorização à aplicação prática da IA, nota-se que o interesse pelo desenvolvimento desses sistemas é quase exclusivo de órgãos governamentais e instituições acadêmicas, havendo poucas iniciativas por parte da indústria

⁶ Although humans speak in a spontaneous way and, generally, cannot deal effectively with imposed language constraints, there are many regular and daily tasks where the domain discourse is inherently limited. By exploiting these limitations, investigators hoped to create a practical way of using SLT (GRAZINA 2010: 5).

da tecnologia. E, até o momento, o emprego desses sistemas encontra-se restrito a alguns poucos ambientes (ZHANG 2003).

O governo dos EUA, por exemplo, tem financiado projetos de IA com a finalidade de desenvolver sistemas para serem usados em guerras e conflitos internacionais (AIKEN, SIMMONS E BALAN 2010), como é o caso dos sistemas *TransTac* e *IraqComm*, usados por soldados norte-americanos no Iraque, no Afeganistão e no Paquistão (BACH et al. 2007; SRI 2009).

A agência norte-americana de pesquisas DARPA (*Defense Advanced Research Projects Agency* - Agência de Projetos de Pesquisa Avançada de Defesa) possui projetos de larga escala para o desenvolvimento de sistemas de IA cuja finalidade é a tradução de diálogos típicos de consultas médicas, a serem empregados em situações em que o médico e o paciente não falam a mesma língua e estão entrando em contato pela primeira vez. Alguns destes projetos voltam-se para línguas ainda poucos exploradas, como por exemplo o pastó, uma das línguas oficiais do Afeganistão (NARAYANAN et al. 2003; KATHOL et al. 2005). Há ainda pesquisas sobre o emprego de IA em ambientes hospitalares e em organizações de saúde por parte da indústria de tecnologia, que veem boas oportunidades nesse tipo de negócio (SELIGMAN E DILLINGER 2012; Shin et al. 2015).

O governo da Alemanha, por sua vez, financiou um sistema de IA para planejamento de viagens, agendamento de consultas e manutenção remota de computadores via telefone, com a tradução consecutiva de diálogos e negociações entre falantes de japonês, alemão e inglês (KLEIN 1997).

O Japão, principalmente devido ao isolamento geográfico do país, adota medidas para o estudo permanente de IA, já que esse sistema pode ser uma boa opção para diminuir as barreiras linguísticas existentes entre as línguas ocidentais e o japonês (NAKAMURA 2009).

Os sistemas de IA também podem facilitar a comunicação entre turistas e cidadãos locais durante eventos internacionais de grande porte, quer seja

para fazer reservas em hotéis e restaurantes, quer seja durante o evento, tornando a experiência da viagem mais atraente e agradável, permitindo importantes trocas culturais (STÜKER et al. 2006).

A especialização em domínios, como sugerido na literatura, facilita o desenvolvimento de sistemas capazes de atender as demandas em cenários realísticos com maior segurança. Há, todavia, projetos como o *MASTOR*, da IBM e o *TC-STAR*, da Comissão Europeia⁷, voltados para a interpretação de qualquer tipo de diálogo, sem restrições discursivas de qualquer natureza (ZHANG 2003; STÜKER et al. 2006).

Sistemas como esses, embora atinjam pontuações menores quando submetidos a testes avaliativos, podem servir para os mais variados fins, já que não há, teoricamente, limites de vocabulário e os algoritmos utilizados para o processamento de fala são capazes de lidar com as mais variadas estruturas gramaticais (FÜGEN et al. 2006; AIKEN, SIMMONS E BALAN 2010). Fügen (2008), por exemplo, apresenta um sistema de IA voltado para palestras e discursos, capaz de interpretar em tempo real e projetado para lidar com os ruídos de fundo típicos de palestras.

Evidentemente, as possibilidades de emprego da IA são praticamente infinitas, assim como são as possibilidades comunicativas inerentes ao ser humano. Além da sua relevância para a engenharia de *softwares* e para a ciência como um todo, os impactos econômicos e culturais proporcionados pela IA são notórios. Até mesmo um sistema simples e limitado pode ser útil para turistas, seja para comprar refeições em restaurantes, adquirir passagens aéreas, encontrar determinados lugares em uma cidade e uma série de outras pequenas tarefas. Segundo Kitano (1994), sistemas deste tipo facilitariam muito a vida dos turistas, além de tornar a experiência de viagem mais interessante, conforme explicitam Kitano (op.cit.) e Stüker et al. (2006), já citados.

⁷ A Comissão Europeia é uma instituição politicamente independente, que representa e defende os interesses da União Europeia (UE).

A realidade, todavia, é que os seus benefícios ainda são pouco explorados devido às limitações tecnológicas. E embora o presente trabalho englobe um primeiro referencial bibliográfico sobre a IA, de 1987 até o presente, e muito embora se tenha buscado construir um modesto inventário de termos a ela relacionados, as colocações acima não são as únicas possibilidades vislumbradas. Além do mais, não se pode deixar de salientar que, conforme as pesquisas avançam, novos recursos são adicionados à tecnologia já desenvolvida e novos públicos são alcançados, de forma que seu emprego se encontra em constante processo de transformação, em conformidade com os demais progressos tecnológicos da contemporaneidade.

4. O que já se sabe sobre a tecnologia de IA até o momento

A partir do inventário de candidatos a termos, foi possível investigar nos 285 trabalhos coletados informações específicas concernentes à arquitetura de um *software* de IA. Com todos os textos digitalizados e convertidos em PDF, foi possível, mais uma vez utilizando a ferramenta Adobe Acrobat Pro, uma busca específica sobre tal descrição.

Constatou-se, como já mencionado, que a IA só é possível graças a três componentes tecnológicos que são integrados em uma só interface: Reconhecimento Automático de Fala (ASR - *Automatic Speech Recognition*); Tradução automática (TA) e Síntese de voz (TTS - *Text-To-Speech*).

A Figura 1 ilustra a arquitetura básica de um sistema de IA bidirecional para o par de idiomas A e B. O trecho de fala do idioma A é inserido no mecanismo e é reconhecido pelo componente de ASR, a fim de produzir-se um texto contendo a transcrição do trecho, ainda no idioma A. A transcrição então é traduzida pelo componente de TA que produz uma tradução textual

FREITAS, F. de S.; ESQUEDA, M. D. — Interpretação automática ou tradução automática de fala: conceitos, definições e arquitetura de *software*

no idioma B. Por fim, o texto traduzido passa pelo componente de TTS e então o trecho de fala contendo o enunciado traduzido é gerado e verbalizado (LEE 2015).

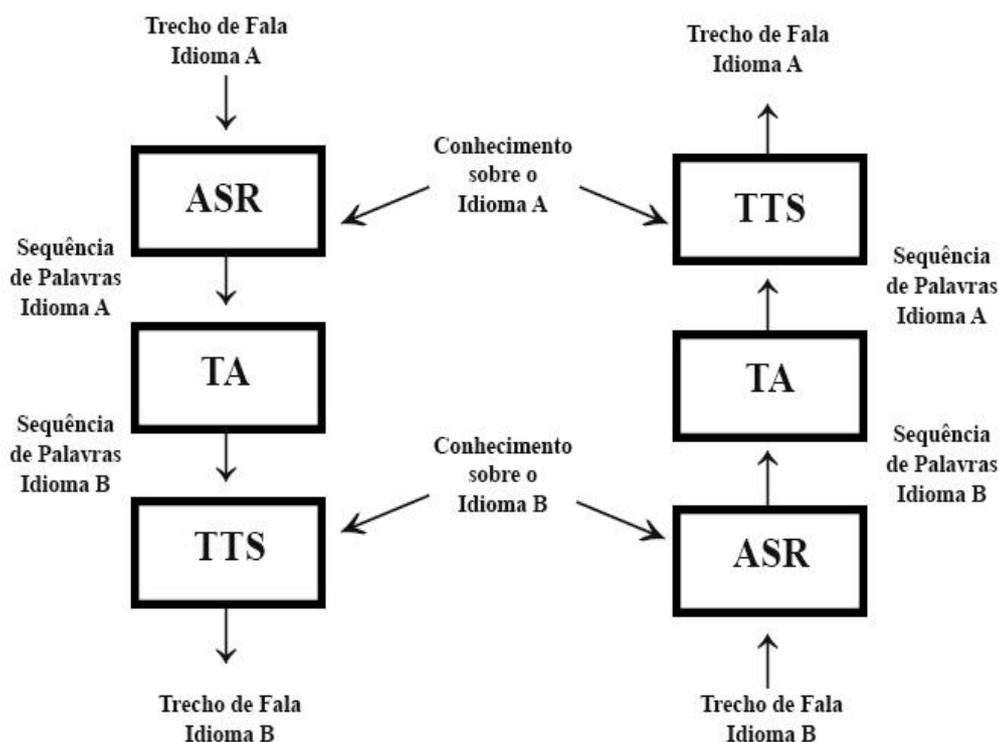


Figura 1 - Arquitetura de um sistema de IA bidirecional
Fonte: LEE (2015); nossa tradução.

Há sistemas de IA que são construídos a partir de componentes desenvolvidos de forma independente, acoplados em uma só interface através de *softwares* específicos (AIKEN, SIMMONS E BALAN 2010). Na maioria das vezes, todavia, esse tipo de acoplagem não produz *feedback* sobre os procedimentos isolados. Conseqüentemente, torna-se impossível que um erro seja corrigido pelos componentes antes do fim de todo o processo. Assim, um erro gerado pelo ASR é replicado para os demais componentes de forma sequencial, comprometendo o desempenho de todo o sistema (LEE 2015).

FREITAS, F. de S.; ESQUEDA, M. D. — Interpretação automática ou tradução automática de fala: conceitos, definições e arquitetura de *software*

Por outro lado, há sistemas de IA desenvolvidos a partir de uma integração mais sofisticada entre os componentes, visando um processamento coerente e um desempenho otimizado de ponta a ponta. Em sistemas desse tipo, o componente de ASR produz várias opções (hipóteses) para que o componente de TA possa escolher a que melhor se encaixa ao contexto do discurso com que se trabalha (CASACUBERTA et al. 2008). Hashimoto et al. (2012) afirmam, no entanto, que poucos sistemas têm proposto este tipo de abordagem para o componente de TTS, o que pode prejudicar a compreensão do conteúdo traduzido e verbalizado pelo sistema.

Segundo Barreiro et al. (2014), o estado da arte atual da IA caracteriza-se por

uma integração relativamente fraca entre os três módulos, não explorando as sinergias existentes entre o reconhecimento e a tradução, entre a tradução e a síntese e ainda entre o reconhecimento e a síntese. Por exemplo, o módulo de reconhecimento escolhe normalmente uma única hipótese de transcrição que será a entrada do módulo de tradução. Se, em alternativa a esta hipótese, for oferecida ao módulo de tradução uma lista de possíveis transcrições, este pode decidir qual a mais adequada aos modelos de tradução. Por outro lado, o módulo de síntese assume que receberá como entrada um texto fluente, o que usualmente não acontece quando essa entrada resulta de um módulo de tradução automática (BARREIRO et al. 2014: 78).

Nessa perspectiva, o reconhecimento automático de fala (ASR) refere-se ao processo computacional de converter um sinal acústico de fala em uma sequência de palavras digitalizadas. Devido à complexidade desta tarefa, os primeiros sistemas ASR eram capazes de reconhecer apenas números e, ainda assim, o locutor tinha que enunciar pausadamente e de forma clara. O *Audrey*, desenvolvido pelos Laboratórios Bell dos Estados Unidos, na década de 50, é um exemplo desse tipo de sistema (SAMPAIO NETO 2011).

Em 1962, a IBM fez uma demonstração pública do *Shoe Box*, sistema capaz de reconhecer 16 palavras do inglês (YADAV E MUKHEDKAR 2013). Também

na década de 60, além dos EUA e da Inglaterra, países como o Japão e a URSS começaram a desenvolver sistemas ASR capazes de reconhecer um número limitado de consoantes e vogais.

Nesses primeiros anos, a tecnologia ASR foi trabalhada através da abordagem linguística, segundo a qual a comunicação oral humana se dá nos níveis acústico, fonético, lexical, sintático, semântico e pragmático. Um sistema capaz de compreender fala humana, portanto, deveria ser capaz de lidar com esses níveis. Em consequência disso, a maioria das abordagens aplicada à ASR na década de 60 tinham por base a segmentação dos trechos de fala em unidades fonéticas (PIERACCINI E LUBENSKY 2005).

Na década de 70, a despeito de todas as dificuldades e limitações impostas pela tecnologia da época, a agência de pesquisa norte-americana DARPA (*Defense Advanced Research Projects Agency*) interessou-se em financiar pesquisas sobre ASR, além de investir na melhoria dos sistemas já desenvolvidos. Em 1975, surgiu então a técnica *dynamic time-warping* (DTW), um algoritmo que compara e alinha duas sequências (neste caso de áudio) que variam em tempo e velocidade, determinando assim a semelhança que há entre uma e outra (SAMPAIO NETO 2011).

Nessa época, surgiram também o sistema *Harpy*, desenvolvido pela já mencionada Universidade de Carnegie Mellon, dos Estados Unidos, com o apoio da DARPA e da Threshold Technology, a primeira empresa norte-americana especializada em ASR. Utilizando-se algumas abordagens de sistemas anteriores ao *Harpy*, que era capaz de reconhecer 1.011 palavras, foi introduzido um método mais eficiente de busca denominado *beam search* (WAIBEL E LEE 1990).

No começo da década de 80, os sistemas ASR começaram a aplicar regras sintáticas. O mecanismo de reconhecimento de fala passou a ser feito a partir de palavras isoladas que eram captadas e analisadas. Caso as palavras

seguissem regras gramaticais e sintáticas já conhecidas pelo sistema, então era possível determinar quais eram essas palavras (SAMPAIO NETO 2011).

A fala humana, no entanto, é repleta de variáveis tais como sotaques, dialetos e exceções inerentes à língua, além de sofrer constante interferência de ruídos. Incapaz de lidar com essas variáveis, as abordagens linguísticas e aquelas baseadas em regras mostraram-se ineficientes. As transcrições lexicais possuíam muitos erros que se propagavam para as interpretações sintáticas e semânticas do conteúdo inserido na tecnologia (PIERACCINI E LUBENSKY 2005).

Na esteira desses acontecimentos, no fim da década de 80, popularizou-se a abordagem estatística, a partir da qual os sistemas ASR começaram a reconhecer um número cada vez maior de palavras, além de serem empregados em várias áreas. Em 1987, por exemplo, a empresa norte-americana *Worlds of Wonders* implantou um dispositivo ASR na boneca Julie, que podia ser treinada para reconhecer a voz da criança e responder determinadas perguntas (YADAV E MUKHEDKAR 2013).

Segundo Aminzadeh e Shen (2009), o estado da arte da tecnologia ASR da atualidade é representado basicamente pelo uso dos Modelos Ocultos de Markov (*Hidden Markov Model* - HMM⁸), método estatístico que analisa o sinal acústico inserido no sistema e faz uma série de observações, além de modelar uma lista de probabilidades para as palavras. Esses modelos são obtidos através do treinamento (*training*) de dados específicos que fornecem as estatísticas com as quais o sistema trabalhará (AMINZADEH E SHEN 2009).

Assim, o reconhecimento automático de fala torna-se um processo de busca pela sequência mais provável de palavras a partir de um amplo conjunto de sequências geradas pelo sistema. Para exemplificar como esse

⁸ Para maiores detalhes de como funcionam os modelos HMM em um sistema de ASR ver Ynoguti (1999).

FREITAS, F. de S.; ESQUEDA, M. D. — Interpretação automática ou tradução automática de fala: conceitos, definições e arquitetura de *software*

processo se dá em nível estatístico, Aminzadeh e Shen (2009) citam a equação matemática⁹ fundamental para este tipo de processo:

$$\mathbf{W}^* = \operatorname{argmax}_{\mathbf{W}} P(o | \mathbf{W})P(\mathbf{W})$$

Figura 2 - Equação matemática de sequência mais provável de palavras

Fonte: Aminzadeh e Shen (2009)

$P(o|\mathbf{W})$, comumente chamado de modelo acústico, é a função probabilística que molda \mathbf{W} para as observações acústicas \mathbf{O} . $P(\mathbf{W})$, também chamado de modelo de linguagem, é a probabilidade de se gerar a palavra w a partir da lista de palavras levantadas pelo modelo HMM. \mathbf{W}^* é a sequência final de palavras geradas pelo sistema (AMINZADEH E SHEN 2009; LEE 2015).

Sistemas ASR de larga escala podem recorrer a milhões de parâmetros (probabilidades relacionadas aos enunciados de fala) para treinar os modelos acústicos e de linguagem. Conseqüentemente, para se obter \mathbf{W}^* é preciso lançar mão de muitos recursos computacionais. Segundo Aminzadeh e Shen (2009), os sistemas da atualidade podem usar mais de dez vezes os recursos do computador durante o processo de reconhecimento. Sistemas ASR menores, todavia, como aqueles instalados em *smartphones* e *tablets* são computacionalmente mais econômicos (AMINZADEH E SHEN 2009).

O treinamento do modelo acústico é feito através de *corpus* de voz digitalizada, contendo as transcrições ortográficas das palavras e/ou as transcrições fonéticas, chamadas unidades de modelagem¹⁰. A forma com que as unidades são modeladas depende de cada sistema. Segundo Lee (2015), enquanto a modelagem no nível da palavra é mais adequada para sistemas de pequeno vocabulário, a modelagem no nível das sílabas e demais unidades

⁹ Devido a limitações de espaço, as equações matemáticas não serão detalhadas neste trabalho. Para maiores detalhes, o trabalho de Sampaio Neto (2011) poderá ser consultado.

¹⁰ Modeling units (LEE 2015).

constituintes das palavras é indicada para sistemas que trabalham com o reconhecimento de fala contínua com amplo vocabulário.

Esse último tipo de modelagem tem a vantagem de oferecer maiores probabilidades de reconhecimento para palavras que não estão presentes no *corpus* de treinamento, haja vista que uma única unidade pode estar presente em várias palavras. Segundo Sampaio Neto (2011: 31) “estimar um bom modelo acústico é considerada a etapa mais difícil dentro do projeto de um sistema ASR”.

O treinamento do modelo de linguagem, por outro lado, é feito através de um vocabulário digital e um arquivo contendo frases das quais serão extraídas as estatísticas para o treinamento do modelo. Segundo Yadav e Mukhedkar (2013), o modelo de linguagem tenta capturar as propriedades do idioma e prever quais palavras serão ditas em dada sequência.

A relação entre o tamanho do vocabulário digital e o número de frases utilizadas para o treinamento do modelo não influenciam no desempenho do sistema. No sistema *TransTac 2007 - Eyes and Hands-Free*, desenvolvido pela Universidade Carnegie Mellon, por exemplo, enquanto o modelo de linguagem foi treinado a partir de vários *corpora*, totalizando 1,7 milhões de palavras, o vocabulário digital tinha apenas 11 mil palavras. Esse exemplo sugere que a extensão do arquivo de treinamento é mais importante do que o vocabulário (BACH et al. 2007).

Quando trazida para o contexto dos sistemas de IA, no entanto, é muito importante que a tecnologia ASR seja modelada para um acoplamento forte (*tight coupling*) com a tecnologia de TA, de forma a não produzir erros que afetem todo o processo de interpretação (GRAZINA 2010).

Assim como a tecnologia ASR, a TA tem sido implementada através de várias abordagens sucessivas e sua evolução é caracterizada pelo intenso aprendizado com os problemas e erros encontrados ao longo de aproximadamente 60 anos de pesquisas e testes (FERNANDES E BARTHOLAMEI JUNIOR

2009). Na IA, a tecnologia de TA encontra-se no centro de todo o processo, sendo a responsável pela tradução de um conteúdo originalmente oral que, após ser processado, retorna ao formato oral através do sistema TTS.

Inicialmente, ao lidar com o processamento da linguagem humana, observou-se que cada sistema linguístico possuía um conjunto de regras específico, de modo que cada idioma é regido por regras diferentes (ARNOLD et al. 1994). Partindo desta observação, é possível destacar algumas abordagens aplicadas à TA: a abordagem lexicográfica (*dictionary-based approach*), a abordagem exemplária (tradução de *example-based approach* utilizada por FERNANDES E BARTHOLAMEI JUNIOR, 2009), a abordagem interlíngua (*interlingua approach*) e a abordagem estatística (*statistical approach*) (HUTCHINS 2015).

Embora essas abordagens possuam cada qual a sua especificidade e tenham contribuído de forma relevante para a evolução da tecnologia de TA, a maioria dos pesquisadores da IA se referem à proeminência das abordagens interlíngua e estatística.

A partir dessa descrição geral da arquitetura de *softwares* de IA, pode-se dizer que esta ferramenta ainda carrega consigo vários desafios (BARREIROS et al. 2014: 77).

Embora representem um obstáculo para a realização de interpretações totalmente isentas de erros, os desafios que acompanham os sistemas de IA desde o seu surgimento, em parte herdados da tradução automática, não têm impedido que a indústria mundial de tecnologia continue apostando no conceito. Em um mundo cada vez mais globalizado, democrático e acessível a qualquer pessoa, a comunicação entre os povos desempenha um papel de vital importância. A fim de manterem-se competitivos e eficientes no cenário global, governos, empresas e instituições de pesquisa dependem de informações de alta qualidade, que devem ser disseminadas com naturalidade e eficácia. A maior parte das informações, todavia, não está disponível no

idioma de quem delas necessita e se não forem interpretadas, não poderão ser utilizadas, comprometendo todo o sistema em que as informações circulam (WAIBEL E FÜGEN 2008).

Palavras finais

Nos últimos 15 anos, a IA vem sendo apresentada em conferências internacionais com maior frequência, passando de um assunto estranho ao público em geral para um dos principais interesses de pesquisa para os estudiosos do processamento de fala. Segundo Waibel e Fügen (2008), a convergência de novas tecnologias e a crescente atenção à necessidade de se ter uma comunicação multilíngue em um mundo globalizado são os motivos que explicam o crescente interesse em tecnologias desse tipo.

Buscou-se, por meio deste trabalho, realizar um levantamento bibliográfico concernente aos estudos sobre IA, especificamente quanto a seus conceitos, definições e descrição da arquitetura desse tipo de tecnologia. A partir de um corpus de 285 artigos científicos que buscam definir, conceituar e descrever a IA e os sistemas que a constituem, foi possível elaborar um inventário de candidatos a termos. Verificou-se que *speech translation*, *speech-to-speech translation*, *spoken language translation* e *simultaneous translation* são termos recorrentes na literatura que definem e discutem essa tecnologia.

Em relação aos termos adotados em língua portuguesa, verificou-se que o termo “tradução de voz” é utilizado pelas pesquisas brasileiras, embora ainda evidentemente escassas, e que “tradução de fala” é utilizado por estudiosos de Portugal, em correspondência ao termo *speech translation* (BARREIRO et al. 2014; DUARTE 2014; PINTO E PRIKLADNICKI 2015).

FREITAS, F. de S.; ESQUEDA, M. D. — Interpretação automática ou tradução automática de fala: conceitos, definições e arquitetura de *software*

A partir deste trabalho, espera-se que os primeiros passos tenham sido dados rumo a um estado da arte sobre a IA, almejando veicular o que se tem estudado sobre essa emergente temática. Nas palavras de Grazina (2010), o panorama socioeconômico da atualidade requer, mais do que nunca, a construção de canais de comunicação e interação multilíngues em nível global, força motriz de grande importância para novas frentes de investigação tanto para os Estudos da Tradução e Interpretação quanto das Ciências da Computação.

Referências bibliográficas

- ACM DIGITAL LIBRARY [Internet]. New York: ACM. 2013 - [citado em 2016 jan. 12]. Disponível em: <<http://dl.acm.org/>>
- AIKEN, M.; Simmons, L. L.; Balan, S. “Automatic Interpretation of English Speech”. In: *Issues in Information Systems*, v. 11, n. 1, 2010.p.129-133. Disponível em: <http://iacis.org/iis/2010/129-133_LV2010_1335.pdf>. Acesso em 17 fev. 2016.
- AMINZADEH, A. R.; SHEN, W. “Advocate: A distributed Architecture for Speech-to-Speech Translation”. *Lincoln Laboratory Journal*. Lexington, v. 18, n. 1, 2009. Disponível em: <https://www.ll.mit.edu/publications/journal/pdf/vol18_no1/18_1_3_Aminzadeh.pdf>. Acesso em: 04 ago. 2015.
- ARNOLD, D. et al. *Machine Translation: An Introductory Guide*. Londres: NCC Blackwell, 1994. Disponível em: <<http://promethee.philo.ulg.ac.be/engdep1/download/bacIII/Arnold%20et%20al%20Machine%20Translation.pdf>>. Acesso em: 21 nov. 2015.

FREITAS, F. de S.; ESQUEDA, M. D. — Interpretação automática ou tradução automática de fala: conceitos, definições e arquitetura de *software*

- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. NBRISO/IEC9126-1 Engenharia de *software* - Qualidade de produto - Parte 1: Modelo de qualidade, 2003.
- BACH, N. et al. The CMU TransTac 2007 Eyes-free and Hands-free Two-way Speech-to-Speech Translation System. In: International Workshop On Spoken Language Translation (IWSLT), 2007, Trento. Proceedings... Trento: Fondazione Bruno Kessler, 2007. Disponível em: <https://www.cs.cmu.edu/~awb/papers/iwslt2007/iwslt2007_transta c.pdf>. Acesso em: 20 nov. 2015.
- BARREIRO, A. et al. Projetos sobre tradução automática do português no laboratório de sistemas de língua falada do INESC-ID. *Linguamática*. Lisboa, v. 6, n. 2, 2014. p.75-85. Disponível em: <<http://www.linguamatica.com/index.php/linguamatica/article/view/v6n2-6>>. Acesso em: 04 ago. 2015.
- CARTER, D. et al. Translation Methodology in the Spoken Language Translator: An Evaluation. In: Workshop On Spoken Language Translation, 1997, Madri. Proceedings... Madri: ACL, 1997. Disponível em: <<http://arxiv.org/pdf/cmp-lg/9705015v1.pdf>>. Acesso em: 10 jan. 2016.
- CASACUBIERTA, F. et al. “Recent Efforts in Spoken Language Translation”. In: *Signal Processing Magazine*, Maryland, v. 25, n.3, 2008. p. 80-88. Disponível em: <[http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=4490204 &newsearch=true&queryText=Recent%20Efforts%20in%20Spoken%20La nguage%20Translation](http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=4490204&newsearch=true&queryText=Recent%20Efforts%20in%20Spoken%20La nguage%20Translation)>. Acesso em: 21 jan. 2016.
- CASTELLUCCIO, Michael. Skype Universal Translator. *Strategic TechNotes*, Rancho Palos Verdes, v. 16, n. 10, 2014. Disponível em: <<http://www.imanet.org/technotes/TNVol16No10.pdf>>. Acesso em: 01 mar. 2016.

FREITAS, F. de S.; ESQUEDA, M. D. — Interpretação automática ou tradução automática de fala: conceitos, definições e arquitetura de *software*

- DECAMP, J.; ZETZSCHE, J. “A History of Translation Technology in the United States”. In: CHAN, S. (Org.). *The Routledge Encyclopedia of Translation Technology*. Londres/Nova York: Routledge, 2015. p.120-136.
- DUARTE, T. S. Máquinas de tradução aplicada à comunicação em tempo real para desenvolvimento distribuído de *software*. 117f. Dissertação (Mestrado) - Pontifícia Universidade Católica Do Rio Grande Do Sul, Porto Alegre, 2014.
- FERNANDES, L. P.; BARTHOLAMEI JUNIOR, L. A. *Estudos da Tradução II*. Universidade Federal de Santa Catarina, Florianópolis, 2009. Disponível em: <http://www.libras.ufsc.br/colecaoLetrasLibras/eixoFormacaoEspecificica/estudosDaTraducaoII/assets/432/TEXTTO_BASE_-_Estudos_da_Traducao_II.pdf>. Acesso em: 2 maio 2016.
- FERREIRA, N. S. A. “As Pesquisas Denominadas ‘Estado Da Arte’”. In: *Educação & Sociedade*, Campinas, ano XXIII, n. 79, 2002. Disponível em: <<http://www.scielo.br/pdf/es/v23n79/10857.pdf>>. Acesso em: 2 maio 2016.
- FINATTO, M. J. B. O Papel Da Definição De Termos Técnico-Científicos. In: *Revista da ABRALIN*, v.1, n.1, 2002. p. 73-97.
- FREDERKING, R.; RUDNICKY, A.; Hogan, C. Interactive Speech Translation in the DIPLOMAT Project. In: ACL/EACL Workshop On Spoken Language Translation, 1997, Madri. Proceedings... Madri: Universidad Nacional de Educación a Distancia, 1997. Disponível em: <<http://www.aclweb.org/anthology/W97-0409> >. Acesso em: 27 fev. 2016.
- FROMM, G. “A Construção do Sentido em Vocabulários Técnicos: o Uso de Corpora e Outros procedimentos”. In: *Crop*, São Paulo, v. 10, 2005. p. 225-239.

FREITAS, F. de S.; ESQUEDA, M. D. – Interpretação automática ou tradução automática de fala: conceitos, definições e arquitetura de *software*

- FÜGEN, C. A system for simultaneous translation of lectures and speeches. 2008. 204f. Tese (PhD) - Universität Karlsruhe, Mannheim, 07 Nov. 2008. Disponível em: <http://dnb.info/1014223113/34?origin=publication_detailsam>. Acesso em: 27 fev. 2016.
- FÜGEN, C. et al. Open Domain Speech Translation: From Seminars and Speeches to Lectures. In: TC-Star Speech to Speech Translation Workshop, Barcelona, 2006. Disponível em: <http://www.cs.cmu.edu/~tanja/Papers/FuegenSchultz_JEP2006.pdf>. Acesso em: 20 nov. 2015.
- GAO, Y. “10 Emerging Technologies that will Change your World”. In: *Technology Review*. Massachusetts Institute of Technology, Boston, 2004. Disponível em: <http://www.rle.mit.edu/thz/documents/10_emerging_tech.pdf>. Acesso em: 04 ago. 2015.
- GLICKMAN, O. Using Domain Knowledge to Improve End to End Performance in a Speech Translation System. 1995. Dissertação (Mestrado) - Carnegie Mellon University, Pittsburgh, 1995. Disponível em: <http://isl.anthropomatik.kit.edu/cmu-kit/downloads/1995_Using_Context_in_Machine_Translation_of_Spoken_Language.pdf>. Acesso em: 21 jan. 2016.
- GRAZINA, N. M. M. Automatic Speech Translation. Tese (Mestrado) - Instituto Superior Técnico, Universidade de Lisboa, Lisboa, 2010. Disponível em: <<http://www.inesc-id.pt/pt/indicadores/Ficheiros/5512.pdf>>. Acesso em: 09 nov. 2015.
- GUSMÃO, A. J. A. Tradução Automática de Notícias Televisivas. 79f. Dissertação (Mestrado) - Universidade Técnica de Lisboa, Lisboa, 2008. Disponível em: <<https://www.l2f.inesc->

FREITAS, F. de S.; ESQUEDA, M. D. — Interpretação automática ou tradução automática de fala: conceitos, definições e arquitetura de *software*

id.pt/wiki/index.php/Alexandre_Gusm%C3%A3o>. Acesso em: 2 maio 2016.

HALE, S.; Napier, J. *Research Methods in Interpreting: a Practical Resource*. London and New York: Bloomsbury Academic, 2013.

HASHIMOTO, K. et al. “Impacts of machine translation and speech synthesis on speech-to-speech translation”. In: *Speech Communication*, v.54, n.7, p.857-866, 2012. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0167639312000283>>. Acesso em: 2 maio 2016.

HUTCHINS, J. “From first conception to first demonstration: the nascent years of machine translation, 1947-1954, A chronology”. In: *Machine Translation*, v. 12, n. 3, p. 195-252, 1997. Disponível em: <<http://www.hutchinsweb.me.uk/MTJ-1997.pdf>>. Acesso em 05 mar. 2016.

HUTCHINS, J. “Machine Translation, History of Research and Applications”. In: CHAN, Sin-Wai (Org.). *The Routledge Encyclopedia of Translation Technology*. Londres/Nova York: Routledge, 2015. p.120-136.

HUTCHINS, J. *Machine Translation: Past, Present, Future*. Chichester: Ellis Horwood, 1986. Disponível em: <<http://www.hutchinsweb.me.uk/PPF-TOC.htm>>. Acesso em: 04 ago. 2015.

IEEE Xplore Digital Library [Internet]. [s.l.]: IEEE. 1998 - [citado em 2016 jan. 12]. Disponível em: <<http://ieeexplore.ieee.org/Xplore/home.jsp>>.

INTERNATIONAL CENTER FOR ADVANCED COMMUNICATION TECHNOLOGIES - INTERACT [Internet]. Pittsburgh e Karlsruhe: CMU e KIT. 2004 - [citado em 2016 jan. 12]. Disponível em: <http://isl.anthropomatik.kit.edu/cmu-kit/english/2162_2673.php>

ISOTANI, R. et al. “Speech-to-Speech Translation Software on PDAs for Travel Conversation”. In: *NEC Research and Development*. Tóquio, v.44, n.2,

FREITAS, F. de S.; ESQUEDA, M. D. – Interpretação automática ou tradução automática de fala: conceitos, definições e arquitetura de *software*

Apr. 2002. P. 197-202. Disponível em: <
https://www.researchgate.net/publication/265217780_Speech-to-Speech_Translation_Software_on_PDAs_for_Travel_Conversation_Speech-to-Speech_Translation_Software_on_PDAs_for_Travel_Conversation>. Acesso em: 27 fev. 2016.

JEKAT, S.; KLEIN, A. “Machine Interpretation: Open Problems and Some Solutions”. In: *Interpreting*, Amsterdam, v. 1, n. 1, p. 7-20, 1996.

KATHOL, A. et al. Speech Translation for Low-Resource Languages: The Case of Pashto. In: INTERSPEECH, 2005, Lisboa. Proceedings... Lisboa: INESC ID's Spoken Language Systems Laboratory, 2005. Disponível em: <https://www.sri.com/sites/default/files/publications/speech_translation_for_low-resource_languages.pdf>. Acesso em: 27 fev. 2016.

KIKUI, G. et al. “Comparative study on corpora for speech translation”. In: *IEEE Transactions on Audio, Speech, and Language Processing*, v.14, n.5, 2006. Disponível em: <<http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=1677987&newsearch=true&queryText=Comparative%20study%20on%20corpora%20for%20speech%20translation&fname=&lname=&title=&volume=&issue=&spage=>>>. Acesso em: 2 maio 2016.

KITANO, H. ΦDMDialog: “An Experimental Speech-to-Speech Dialog Translation System”. In: *Computer*, v.24, n.6, p.36-50, 1991. Disponível em: <<http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=86837&newsearch=true&queryText=%CE%A6DMDialog:%20An%20Experimental%20Speech-to-Speech%20Dialog%20Translation%20System>>. Acesso em: 25 jan. 2016.

_____. *Speech-to-speech translation: a massively parallel memory-based approach*. 1. ed. Nova York: Springer Science & Business Media, 1994.

FREITAS, F. de S.; ESQUEDA, M. D. — Interpretação automática ou tradução automática de fala: conceitos, definições e arquitetura de *software*

ISBN - 978-1-4615-2732-9 (ebook). Disponível em: <<http://bookzz.org/ireader/2118327>>. Acesso em : 25 jan. 2016.

- KLEIN, A. “The Verbmobil Project: Automatic Interpreting of Business Negotiation Dialogues”. In: *Electronic Markets*, v. 7, n. 1, p. 29-30, 1997. Disponível em: <https://www.researchgate.net/publication/220505652_The_Verbmobil_Project_Automatic_Interpreting_of_Business_Negotiation_Dialogue_s>. Acesso em: 27 fev. 2016.
- LEE, T. “Speech Translation”. In: CHAN, S. (Org.). *The Routledge Encyclopedia of Translation Technology*. Londres/Nova York: Routledge, 2015. p. 619-631.
- LEVIN, L. et al. Lessons Learned from a Task-based Evaluation of Speech-to-Speech Machine Translation. In: Language Resources and Evaluation Conference (LREC), 2000, Atenas. Proceedings... Atenas: LREC, 2000. Disponível em: <<http://www.cs.cmu.edu/~aria/Papers/LREC-2000-Levin.pdf>>. Acesso em 2 maio 2016.
- LIU, Q.; ZHANG, X. “Machine Translation: General”. In: CHAN, S. (Org.). *The Routledge Encyclopedia of Translation Technology*. Londres/Nova York: Routledge, 2015. p. 105-119. Disponível em: <<http://bookzz.org/book/2470011/5925f6>>. Acesso em: 19 nov. 2015.
- MATHIAS, L.; BYRNE, W. Statistical Phrase-Based Speech Translation. In: ICASSP, 2006, Toulouse Proceedings... Toulouse: IEEE, 2006. Disponível em: <<http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=1660082&newsearch=true&queryText=Statistical%20Phrase-Based%20Speech%20Translation>>. Acesso em: 2 maio 2016.
- MORIMOTO, T.; KUREMATSU, A. Automatic Speech Translation at ATR. In: MT SUMMIT IV, 1993, Kobe. Proceedings... Kobe: AAMT, 1993. Disponível em: <<http://www.mt-archive.info/MTS-1993-Morimoto.pdf>>. Acesso em: 5 mar. 2016.

FREITAS, F. de S.; ESQUEDA, M. D. — Interpretação automática ou tradução automática de fala: conceitos, definições e arquitetura de *software*

- MRINALINI, K.; VIJAYALAKSHMI, P. Hindi-English speech-to-speech translation system for travel expressions. In: International Conference On Computation Of Power, Energy Information And Communication, 2015, Chennai. Proceedings... Chennai: IEEE, 2015. Disponível em: <<http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=7259472&newsearch=true&queryText=Hindi-English%20speech-to-speech%20translation%20system%20for%20travel%20expressions>>. Acesso em: 15 fev. 2016.
- NAKAMURA, S. *Overcoming the language barrier with speech translation technology*. Science & Technology Trends. Tóqui, n.31, 2009. Disponível em: <<http://www.nistep.go.jp/achiev/ftx/eng/stfc/stt031e/qr31pdf/STTqr3103.pdf>>. Acesso em: 04 ago. 2015.
- NARAYANAN, S. et al. Transonics: A speech-to-speech system for English-Persian interactions. In: IEEE Workshop on Automatic Speech Recognition and Understanding, 2003, São Tomás. Proceedings... São Tomás: IEEE, 2003. Disponível em: <<http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=1318520&newsearch=true&queryText=Transonics:%20a%20speech%20to%20speech%20system%20for%20English-Persian%20interactions>>. Acesso em: 15 fev. 2016.
- PAUL, M. et al. Nobody is perfect: ATR's hybrid approach to spoken language translation. In: International Workshop on Spoken Language Translation (IWSLT), 2005, Pittsburgh. Proceedings... Pittsburgh: IWSLT, 2005. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=E0C592AAEECB2125F7F439B94CCAC1BF?doi=10.1.1.123.8541&rep=rep1&type=pdf>>. Acesso em: 2 maio 2016.

FREITAS, F. de S.; ESQUEDA, M. D. – Interpretação automática ou tradução automática de fala: conceitos, definições e arquitetura de *software*

- PAULIK, M. Learning speech translation from interpretation. 2010. PhD Thesis. Karlsruhe Institute of Technology, 2010. Disponível em: <http://isl.anthropomatik.kit.edu/cmu-kit/downloads/PhD_2010_01_Learning_Speech_Translation_from_Interpretation.pdf>. Acesso em: 2 maio 2016.
- _____; WAIBEL, A. Extracting clues from human interpreter speech for spoken language translation. In: ICASSP 2008 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2008 Las Vegas. Proceedings... Las Vegas: Caesars Palace, 2008. P. 5097-5100. Disponível em: <https://www.researchgate.net/publication/224313055_Extracting_clues_from_human_interpreter_speech_for_spoken_language_translation>. Acesso em: 27 fev. 2016.
- PIERACCINI, R.; LUBENSKY, D. Spoken Language Communication with Machines: The Long and Winding Road from Research to Business. In: ALI, M.; ESPOSITO, F. (Org.). IEA/AIE 2005, LNAI 3533, pp. 6 - 15, 2005. Disponível em: <<http://www.thepieraccinis.com/publications/2005/IEA-AIE-2005.pdf>>. Acesso em: 2 maio 2016.
- PINTO, J. H.; PRIKLADNICKI, R. Um estudo empírico sobre máquinas de tradução em tempo real para equipes distribuídas de desenvolvimento de *software*. In: Brazilian Conference on Software: Theory and Practice, 2015, Belo Horizonte. Proceedings... Belo Horizonte: CBSOFT.ORG, 2015. Disponível em: <<http://cbsoft.org/articles/0000/0528/WTDSOft.pdf>>. Acesso em: 2 maio 2016.
- PÖCHHACKER, F. *Introducing Interpreting Studies*. Londres: Routledge, 2004. p. 265.

FREITAS, F. de S.; ESQUEDA, M. D. — Interpretação automática ou tradução automática de fala: conceitos, definições e arquitetura de *software*

- RAWLINSON, K. *Google Translate ‘turns interpreter’ with voice function*. Londres: BBC News, 2015. Disponível em: <<http://www.bbc.com/news/technology-30812277>>. Acesso em: 20 nov. 2015.
- RICCI, L. “A Militarized PDA Voice-to-Voice Phrase Translator”. In: *Handheld & Wireless Solutions Journal*, v.2, Winter 2002. p.54-55. Disponível em: <http://projectile.sv.cmu.edu/research/public/talks/speechTranslation/otherPaper/ADS-Militarized_PDA-Story.pdf>. Acesso em: 27 fev. 2016.
- SAMPAIO NETO, N. C. Ferramentas e recursos livres para reconhecimento e síntese de voz em português brasileiro. 2011. 96 f. Tese (Doutorado) - Universidade Federal do Pará, Instituto de Tecnologia, Belém, 2011. Programa de Pós-Graduação em Engenharia Elétrica.
- SELIGMAN, M.; DILLINGER, M. Spoken Language Translation: Three Business Opportunities. In: The Tenth Biennial Conference of the Association for Machine Translation in the Americas (AMTA-12), 2012, San Diego. Proceedings... San Diego: Association for Machine Translation in the Americas, 2012. Disponível em: <<http://amta2012.amtaweb.org/AMTA2012Files/papers/SeligmanDillinger.pdf>>. Acesso em: 27 fev. 2016.
- SHIN, S. et al. Speech-to-Speech Translation Humanoid Robot in Doctor’s Office. In: 6th International Conference on Automation, Robotics and Applications (ICARA), 2015, Queenstown. Proceedings... Queenstown: IEEE, 2015. Disponível em: <<http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=7081196&newsearch=true&queryText=Speech-to-speech%20translation%20humanoid%20robot%20in%20doctor%27s%20office>>. Acesso em: 15 fev. 2016.

FREITAS, F. de S.; ESQUEDA, M. D. — Interpretação automática ou tradução automática de fala: conceitos, definições e arquitetura de *software*

- SRI. *Speech Translation Research at SRI International*. 2009. Disponível em: <<http://www.speech.sri.com/projects/translation/full.shtml>>. Acesso em: 27 fev. 2016.
- STENTIFORD, F. W. M.; STEER, M. G. A speech driven language translation system. In: European Conference on Speech Technology, 1987, Edinburgh. Proceedings... Edinburgh: [s.n.], 1987.
- _____, F. W. M.; Steer, M. G. “Machine translation of speech”. In: *British Telecom Technology Journal*, Londres, v.6, n.2, April 1988.
- STÜKER, S. et al. Speech-to-speech translation services for the olympic games 2008. In: RENALS, S.; BENGIO, S. e FISCUS, J.G. (Org.). *Machine Learning for Multimodal Interaction*. Bethesda: Springer, 2006. p. 297-308. Disponível em: <<http://nlpr-web.ia.ac.cn/cip/ZongPublications/2006%20DO-MLMI.pdf>>. Acesso em: 27 fev. 2016.
- TAKEZAWA, T. *et al.* Toward a Broad-coverage Bilingual Corpus for Speech Translation of Travel Conversations in the Real World. In: Third International Conference on Language Resources and Evaluation (LREC-2002), 2002, Las Palmas. Proceedings... Las Palmas: University of Las Palmas de Gran Canaria, 2002. Disponível em: <<http://hnk.ffzg.hr/bibl/lrec2002/pdf/305.pdf>>. Acesso em: 2 maio 2016.
- _____; KIKUI, G. A comparative study on human communication behaviors and linguistic characteristics for speech-to-speech translation. In: Language Resources and Evaluation Conference - LREC, 2004, Lisboa. Proceedings... Lisboa: Universidade Nova de Lisboa, 2004. p.1589-1592. Disponível em: <<http://www.mt-archive.info/LREC-2004-Takezawa.pdf>>: Acesso em: 27 fev. 2016.
- TOMITA, M.; TOMABECHI, H.; SAITO, H. SpeechTrans: an experimental real-time speech-to-speech translation. In: 1989 Spring Symposium of the

FREITAS, F. de S.; ESQUEDA, M. D. – Interpretação automática ou tradução automática de fala: conceitos, definições e arquitetura de *software*

American Association for Artificial Intelligence, 1989, Stanford. Proceedings... Disponível em: <<http://space.snu.ac.kr/bitstream/10371/85868/1/3.%202235656.pdf>>.

Acesso em: 09 nov. 2015.

WAIBEL, A. et al. Speechalator: two-way speech-to-speech translation on a consumer PDA. In: European Conference on Speech Communication and Technology, 2003, Genebra. Proceedings... Genebra: International Speech Communication Association, 2003. p. 369-372. Disponível em: <<https://www.cs.cmu.edu/~awb/papers/eurospeech2003/speechalator.pdf>>. Acesso em: 27 fev. 2016.

_____; FÜGEN, C. Spoken language translation - enabling cross-lingual human-human communication. In IEEE Signal Processing Magazine, n.3, 2008. Disponível em: <http://isl.anthropomatik.kit.edu/cmu-kit/english/2162_2673.php>. Acesso em: 09 nov. 2015.

_____; LEE, K. F. *Readings in Speech Recognition*. San Mateo: Morgan Kaufmann Publishers, 1990. p. 308-19.

YADAV, K. S.; Mukhedkar, M. M. “Review on Speech Recognition”. In: *International Journal of Science and Engineering*, v. 1, n. 2, 2013. p. 61-70. Disponível em: <http://www.ijse.org/admin/post_image/1409338463_Review_on_Speech_Recognition.pdf>. Acesso em: 5 mar. 2016.

ZHANG, Y. Survey of Current Speech Translation Research. In: Multilingual Speech-to-Speech Translation Seminar, 2003, Pittsburgh. Proceedings... Pittsburgh: Universidade de Carnegie Mellon, 2003. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.61.2089&rep=rep1&type=pdf>>. Acesso em: 27 fev. 2016.

ZHOU, B. et al. Statistical natural language generation for speech-to-speech machine translation. In: International Conference on Spoken Language

FREITAS, F. de S.; ESQUEDA, M. D. — Interpretação automática ou tradução automática de fala: conceitos, definições e arquitetura de *software*

Processing, 2002, Denver. Proceedings... Denver: ICSLP, 2002.
Disponível em:
<<https://pdfs.semanticscholar.org/611a/4df28acc714f4a8728d23676c65178c1f07f.pdf>>. Acesso em: 2 maio 2016.

Data de submissão: 03/05/2016

Data de aprovação: 03/02/2017