

# *Um Robô Linguista que 'Ouve' e 'Fala': Geolinguística, PLN e Tabelas Hash em concurso*

A LINGUIST ROBOT THAT LISTEN AND TALKS:  
GEOLINGUISTIC, PLN AND HASH TABLES ON CONCURRENCY

Edio Roberto **Manfio**\*

Cinthyana Renata Sachs Camerlengo de **Barbosa**\*\*

Fabio Carlos **Moreno**\*\*\*

**Resumo:** Este estudo tem por objetivo apresentar o robô de conversação Professor Tical em uma dimensão mais ampla e com os recursos de *síntese e comandos por voz*. Tical, que se tornou operacional como protótipo durante do III Congresso Internacional de Dialectologia e Sociolinguística (CIDS) na UEL em 2014, continua sendo um campo de provas para aplicações de algumas teorias que mantêm entre si um caráter interdisciplinar como Geolinguística (CARDOSO et al., 2014a, 2014b), Processamento de Linguagem Natural (SCHILDT, 1989; RICH, 1993; MANFIO; MORENO; BARBOSA, 2014a, 2014b) e Processamento de Dados (ZIVIANI, 1999). O 'robô linguista', mesmo dispondo de um banco de dados bastante limitado e apresentando várias falhas típicas de sistemas dessa natureza e dotados desses recursos, serviu grandemente para realizar testes relativos a todas as áreas de conhecimento envolvidas e mostrou-se funcional o suficiente para suscitar a necessidade de dar continuidade ao projeto e às pesquisas: 'ouve', 'fala', realiza buscas rápidas e 'conhece linguística'.

---

\* Doutor em Estudos da Linguagem pela Universidade Estadual de Londrina – UEL (2016). Professor Doutor III na Faculdade de Tecnologia de Garça – Fatec. Contato: prof.ediorobertomanfio@gmail.com / edio.manfio@fatec.sp.gov.br.

\*\* Doutora em Ciência da Computação pelo Instituto Tecnológico de Aeronáutica. Professora Doutora no Centro de Ciências Exatas da Universidade Estadual de Londrina – UEL. Contato: sachs.cinthyana@gmail.com.

\*\*\* Mestrando em Ciência da Computação pela Universidade Estadual de Londrina – UEL. Contato: fbio\_moreno@yahoo.com.br / fabiocmoreno@gmail.com.

**Palavras-chave:** ALiB. Processamento de Linguagem Natural. Reconhecimento de voz.

**Abstract:** This study aims to present the conversational robot Teacher Tical in a wider dimension and with the resources of synthesis and voice commands. TICAL, which became operational as a prototype during the III International Congress of Dialectology and Sociolinguistics at the NFB in 2014, remains a proving ground for some theories that keeps each other an interdisciplinary character as Geolinguistics (CARDOSO et al., 2014a, 2014b), Natural Language Processing (SCHILDT, 1989; RICH, 1993; MANFIO; MORENO; BARBOSA, 2014a, 2014b) and Data Processing (ZIVIANI, 1999). The linguist ‘robot’, even featuring a fairly limited database and showing various typical failures of such systems and equipped with these capabilities, served largely to carry out tests on all the knowledge areas involved and proved to be functional enough to give rise to the need to give continuity to the project and research: ‘listen’, ‘talks’, performs quick searches and knows Linguistics.

**Keywords:** ALiB. Natural Language Processing. Voice recognition.

## Introdução

O Processamento de Linguagem Natural (doravante apenas PLN) não só é uma das áreas de conhecimento que compõem as bases da Inteligência Artificial como também uma das que mais evoluíram durante todo o século passado até os dias de hoje. O PLN está em quase todas as atividades vinculadas ao processamento das linguagens escrita e falada, principalmente quando estão em jogo conhecimentos léxicos, sintáticos e semânticos relativos à compreensão e geração de frases e/ou tradução idiomas, cuja acurácia depende em muito do processamento de linguagens naturais (SCHILDT, 1989; RICH, 1993). Entre variados exemplos de sistemas que se valem do conceito de PLN estão os robôs de conversação ou *chatbots*. Esses sistemas vêm sendo criados por muitos desenvolvedores desde a década de 1960 e, gradualmente, têm sido incrementados com as novas tendências tecnológicas: maior volume de processamento, funcionamento em rede, capacidade de processar sinais de diversas ordens, entre outros.

Joseph Weizenbaum, criador do Eliza (WEIZENBAUM, 1966), um dos robôs de conversação mais citados até hoje quando se fala em PLN, certamente ficaria fascinado com a possibilidade de disponibilizar sua criação na internet para que mais e mais pessoas pudessem testá-la. Mais que isso, dotar sua criação com recurso de *comandos* e *síntese de voz* seria simplesmente genial. Porém, na década de 1960, a tecnologia disponível não tornava isso possível. O Eliza, capaz de portar-se como um psicanalista ao simular diálogos com um interlocutor humano, foi desenvolvido para testar e aplicar conceitos teóricos de diferentes áreas. Toda a conversa, porém, tinha que ser realizada por meio da linguagem escrita, ou seja, utilizando o teclado alfabético.

Da mesma forma, como quase todos os projetos de cunho científico, incluindo o Eliza, o Professor Tical<sup>1</sup> (MANFIO; MORENO; BARBOSA, 2014a, 2014b) tinha como objetivo básico, em sua gênese, dar maior aplicabilidade a alguns elementos de ordem mais teórica. Para o caso de Tical, esses elementos dizem respeito às áreas de Processamento de Dados e de Linguística – especificamente, o Projeto Atlas Linguístico do Brasil – ALiB (COMITÊ NACIONAL DO PROJETO ALIB, 2001; CARDOSO et al., 2014a, 2014b; PROJETO ATLAS LINGUÍSTICO DO BRASIL, 2014). Assim sendo, sua concepção envolveu desde sempre um procedimento interdisciplinar, interdepartamental e interinstitucional, o grau de aplicação das diversas teorias envolvidas acabou, de variadas formas, apresentando resultados, ora no formato de pequenas soluções, ora como dados tabulados em diagramas e tabelas.

Respectivamente, em outras fases de desenvolvimento e pesquisa, Tical passou de um robô de conversação convencional para um *chatbot* capaz de operar com o aplicativo para troca de mensagens WhatsApp e, em uma versão mais atual, foram associadas ao ‘professor linguista’ as capacidades de operar com comandos e síntese de voz. A partir desse ponto da trajetória evolutiva, uniram-se no projeto *Processamento de Linguagem Escrita, Processamento de Linguagem Falada, Linguística Geral, Inteligência Artificial, Desenvolvimento de Sistemas, Tabelas Hash, Educação, redes sociais, entretenimento, difusão cultural* entre outros.

Este estudo, portanto, tem o objetivo apresentar o ‘robô linguista’ em uma dimensão mais ampla e com todas as funcionalidades que a implementação

---

<sup>1</sup> Acrônimo de “Tecnologia Interativa Conversacional sobre Assuntos Linguísticos”.

tornou possíveis até o fechamento deste artigo. Nessa versão, Tical mostra-se como um projeto bastante relevante em termos científicos justamente por concentrar vários diferenciais: é um dos únicos robôs que ‘versam’ sobre linguística disponíveis, está todo em Português Brasileiro, opera em dispositivos móveis a partir de aplicativo de mensagem instantânea, responde com síntese de voz e atende a comandos por voz. Por esses motivos, seu grau de aplicabilidade é bastante alto e esse quesito colabora grandemente para que continue a ser incrementado com o intuito de evoluir e manter-se compatível com tecnologias futuras.

### **Implementação Básica e Funcionamento**

Originalmente, como dito há pouco, Tical tornou-se operacional a partir conceitos básicos de PLN e com um banco de dados sobre a Geolinguística no Brasil. Importante lembrar que o PLN se divide em *Processamento de Linguagem Escrita* e *Processamento de Linguagem Falada* (SCHILDT, 1989; RICH, 1993; BARBOSA, 2004) – doravante PLE e PLF, respectivamente. Para que Tical se tornasse operacional também em *Processamento de Linguagem Falada*, uma vez que, tal como o Eliza, de Weizenbaum, operava apenas com texto escrito por meio do teclado alfabético, foi necessário lançar mão de outros programas e aplicativos, além daqueles já utilizados para a implementação básica de Tical (MANFIO; MORENO; BARBOSA, 2014a, 2014b).

Válido salientar que, nem de longe, havia a pretensão de fazer com que essa aplicação suscitasse nos usuários a sensação de estar falando de fato com um linguista, tal como ocorreu com o Eliza, confundido com um psicanalista em 1966 – seria muita pretensão em nossa conjuntura e, mesmo que possível, definitivamente fora dos objetivos.

Embora isso possa ocorrer de alguma forma, dependendo do grau de evolução que o projeto adquira, ainda é improvável que pessoas confundam máquinas com humanos: a tecnologia avança, mas as pessoas, proporcionalmente, ficam impregnadas com a Tecnologia da Informação vigente. A despeito disso, Alan Mathison Turing, criador de um procedimento capaz distinguir, em sua época, a inteligência artificial da inteligência humana, jamais imaginaria que, passados quase 70 anos, seu método, hoje conhecido como Teste de Turing, estaria paradoxalmente popular e em risco, ao mesmo

tempo. Quase duas décadas após o Teste de Turing, o Eliza de Weizenbaum (1966), ‘portando-se’ como um psicanalista, já colocava à prova o famoso teste. Interessantemente, Descartes (2001), 313 anos antes de Turing, alertava: (i) não teriam (as máquinas) nossa habilidade linguística e (ii) não possuiriam nosso conhecimento de mundo.

Note-se que a discussão sobre autômatos e a possibilidade de eles substituírem os humanos em tarefas cotidianas é relativamente antiga. Não se pode afirmar, no entanto, que Descartes estava prevendo o surgimento da era da eletrônica com a invenção da válvula termiônica e sua substituição após a criação do transistor. Descartes, assim como outros pensadores, estava sensível à evolução tecnológica que, amiúde, acompanha alguns dos sonhos dos homens.

O *Watson* da IBM, no entanto, certamente faria Descartes e Turing, ao menos, reverem alguns conceitos. Operando a partir de linguagem natural, esse moderno sistema tem os recursos do aprendizado dinâmico, da geração de hipóteses, e é considerado pelos próprios desenvolvedores como “uma extensão natural daquilo que os humanos podem fazer em seu melhor” (IBM WATSON, 2015). O Watson talvez seja o mais próximo daquilo que pode ser considerado hoje como *compreensão por voz* – vai além do *reconhecimento de voz*.

Desenvolvido para dar respostas a perguntas no *Jeopardy!* – um dos mais conhecidos programas norte-americanos de televisão – o Watson “representa uma revelação no processamento da linguagem e a analítica” (IBM, 2015): conseguiu êxito competindo contra dois dos campeões mais famosos no programa, em um concurso transmitido ao longo de três noites consecutivas, em 2011. Importante lembrar que ele, longe de ser portátil como um computador convencional ou telefone celular, concentra várias poderosas CPUs em sala climatizada e requer uma quantidade bastante grande de energia para funcionar.

Entretanto, não é demais dizer que esse exemplo, embora conveniente para ilustrar a diversidade tecnológica em PLN, representa o extremo daquilo que o engenho humano pode prover nos dias atuais. Os investimentos feitos em tecnologia por grandes corporações ou pelos diferentes governos dos países mais desenvolvidos do mundo estão muito além da realidade que pesquisadores autônomos ou grupos de pesquisa em universidades em nosso país podem experimentar.

Assim sendo, livre do compromisso ou da necessidade de passar no teste de Turing, superar o *Watson*, impressionar Weinzenbaun ou surpreender Descartes, Tical tem um objetivo muito menos pretensioso – mas não menos importante. Enquanto o Eliza permitia testar alguns conhecimentos teóricos, Tical vai além, podendo disponibilizar informações – linguísticas – ao interlocutor. Tem também a vantagem de não requerer – ainda – grandes investimentos, poderosas CPUs ou salas climatizadas para funcionar. Mais que isso, nessa última versão, a interação não precisa ser exclusivamente realizada por meio da linguagem escrita, ou seja, utilizando o teclado alfabético. Tical, então, um *chatbot* similar a muitos outros, ganhou outro diferencial: PLF – *Processamento de Linguagem Falada*.

Como há vários modos de fazer com que comandos e síntese de voz se tornem efetivamente funcionais em termos de Interação Homem Máquina (IHC), foi necessário escolher uma das opções disponíveis para realizar tal tarefa. Entre todas as possibilidades de escolha, o Coruja (LAPS, 2014) foi eleito.

O Coruja, de acordo com os próprios desenvolvedores, “é um sistema de reconhecimento de voz com suporte a ditado” (LAPS, 2014). Comparado com sistemas menos complexos, fundamentados em gramáticas, tem a vantagem de permitir o desenvolvimento de aplicativos como transcrição forense, criação automática de legenda, entre outros. O motivo da utilização do Coruja é que ele se mostra eficaz para a aplicação em questão, opera em Português do Brasil e está disponível para uso em pesquisa científica.

O Loquendo (NUANCE, 2014) foi empregado para teste em função da qualidade e acurácia da síntese de voz, tal como foi feito com o *Balabolka* (BALABOLKA, 2012) em outro estudo (MANFIO, 2012). Outros sintetizadores estão disponíveis, mas sua qualidade é baixa e isso implica diretamente o grau de ambiência amigável do aplicativo. Válido lembrar que ele não é gratuito e para versões futuras do Tical, em que estejam implicadas outras finalidades que não apenas a pesquisa científica, provavelmente será necessário o uso de um recurso próprio desenvolvido especificamente para esse fim.

O sistema de busca das respostas, por sua vez, presente na gênese de Tical, conta com o *hash*, uma estrutura de dados específica que associa palavras-chave a elementos correlacionados. Essas ‘tabelas de dispersão’ ou ‘tabelas

de espalhamento’ – como também são conhecidas – têm o diferencial realizar buscas rápidas e obter o elemento correlacionado a partir de uma entrada simples, que pode ser uma única palavra. *Hash* significa ‘picadinho’ ou, em uma acepção mais ampla, ‘bagunça’, em função das partes estarem distribuídas em ordem não convencional (ZIVIANI, 1999), o que torna a busca veloz, uma das razões pela qual foi escolhida para essa aplicação desde a criação de Tical (MANFIO; MORENO; BARBOSA, 2014a, 2014b). Uma vez encontrada a resposta no banco, resta ao sintetizador de voz transformá-la em um som audível.

O modo como todos esses sistemas se integram é relativamente simples. O Coruja funciona a partir de Tical, ou seja, ao acionar o executável do robô clicando no ícone, automaticamente são acionados o banco de dados básico do Professor – contendo todas as respostas por escrito cadastradas – e o banco de dados de reconhecimento de voz – contendo as palavras-chave que compõem possíveis perguntas a Tical. A Figura 1 apresenta um diagrama em blocos que ilustra bem isso.



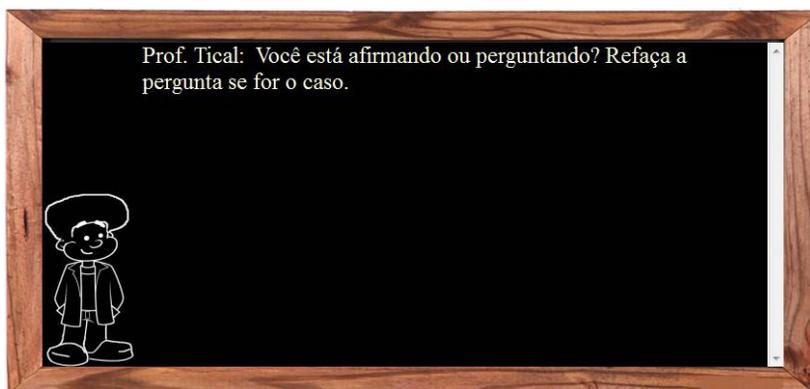
**Figura 1** – Representação em blocos da integração dos sistemas

Em uma das primeiras versões com o recurso de comandos por voz, todas as perguntas feitas oralmente a Tical deviam ser seguidas de um comando também oral para ativação da busca. Após formular oralmente uma pergunta como “o que é o alib?”, era necessário também oralmente dar o comando “buscar” para que o sistema iniciasse a consulta ao banco, tal como exposto no Quadro 1.

**Quadro 1** – Sequência de dois comandos para uma busca

Comando por voz I	Comando por voz II	Ação do sistema
‘o que é o ALiB?’	‘buscar’	efetua a busca

Um detalhe que merece ressalva nessa etapa é que a parte do sistema responsável por cuidar dos comandos por voz foi programada para considerar como perguntas todos os comandos orais, ou seja, o *chatbot* adiciona automaticamente o ponto de interrogação em todas as frases reconhecidas como comandos. Isso deve acontecer porque o sistema de busca por *hashing* foi programado para considerar como perguntas apenas as frases que tenham o ponto de interrogação ao final; caso contrário0 ele retorna como resposta uma outra pergunta: “você está afirmando ou perguntando? Refaça a pergunta se for o caso”, tal como representado na Figura 2.



**Figura 2** – Tela de Tical questionando o interlocutor

Esse cuidado foi necessário desde a primeira versão de Tical – que operava apenas com o teclado alfabético – porque, de um modo geral, nem tudo o que o interlocutor tem a fazer são perguntas. Se a proposta da criação de um *chatbot* é tornar a interface amigável, entre os recursos para conseguir este fim está a capacidade mínima de o robô distinguir perguntas de afirmações, por exemplo.

No cotidiano, há uma diferença enorme entre “O alib é um atlas?” – pergunta – e “O alib é um atlas” – afirmação. Não é exagero lembrar que a pontuação grafemática não dá conta de metade das nuances de prosódia possíveis e distinguíveis em um idioma. A ironia, por exemplo, perceptível prosodicamente à maioria dos ouvidos atentos, precisa ser indicada por extenso num texto escrito: “... disse, ironicamente...”. O escritor Luis Fernando Veríssimo, a propósito, comentou algumas vezes em entrevistas que uma das dificuldades de se trabalhar com a ironia em textos humorísticos é que, se o leitor não a entende ou não a percebe, não há o que fazer, afinal, é um despropósito colocar um aviso entre parênteses ou aspas como ‘ironia’ para indicar que ela está ali.

Como o propósito do sistema – por enquanto – não é considerar o valor linguístico das diferentes entonações possíveis de um mesmo grupo de fonemas ou, mais simplificada e falando, distinguir entre as diferentes prosódias de uma mesma frase, a ênfase manteve-se centrada nos elementos grafemáticos de que depende o sistema de busca pelas tabelas de dispersão ou *hashing*. Os dois exemplos do Quadro 2 são suficientes para ilustrar a maior parte das ocorrências que podem resultar em respostas diferentes quando o ponto de interrogação se faz ausente.

**Quadro 2** – Comparativo entre perguntas e afirmações

<b>Afirmação</b>	<b>Pergunta</b>
o Alib é um atlas	o Alib é um atlas?
Há outras denominações para tangerina no Brasil	Há outras denominações para tangerina no Brasil?

Note-se que tanto as afirmações quanto as perguntas são pertinentes, ou seja, são passíveis de serem feitas e de serem respondidas. Em um futuro próximo, o Professor Tical terá de dar conta de todas as quatro de modo minimamente eficiente e isso não gera maiores problemas ao banco de dados: basta que as informações estejam à disposição do sistema. Válido lembrar que, pensando exclusivamente no interlocutor, há três possibilidades imediatas: (i) faz uma pergunta a Tical *com* ponto de interrogação; (ii) faz uma pergunta a Tical *sem* ponto de interrogação; (iii) faz uma afirmação.

Para (i), Tical simplesmente responde. Esse recurso é o básico e, desde a primeira versão, em que o *chatbot* operava apenas com o teclado alfabético, ele faz isso – afinal, era justamente para essa função que foi programado.

Para (ii), Tical retorna “você está afirmando ou perguntando? Refaça a pergunta se for o caso”. Esse recurso passou a ser necessário, pois já se pensava desde sua gênese que, evoluindo, a interação teria de ser mais completa e seria necessário oferecer ao interlocutor a possibilidade de entrar também com comentários – afirmações. O Eliza de Weizembaun, interessantemente, dá a impressão de lidar bem com isso – embora o autor tenha comentado que se tratava de um ‘truque’.

No entanto, para (iii), considerando-se apenas a condição de uso do teclado alfabético, Tical tem muito pouco preparo. Levando em conta também o recurso de comandos por voz, a situação fica muito pior. Por enquanto, no máximo, pode-se incrementar as linhas de programação para que ele dê como resposta a um comando de voz inusitado algo como “Não entendi. Poderia reformular a pergunta com outros termos?”.

Assim, voltando à questão do funcionamento básico, como a conversão do texto oral – comando de voz – para texto escrito nessa versão de Tical é imediata e a busca pelas respostas obedece a condição em que o Professor operava apenas com o teclado alfabético – como a maioria dos *chatbots* que se conhece –, nada mais adequado que o ponto de interrogação seja adicionado automaticamente quando o comando é reconhecido.

Simplificando, por ora, o usuário só precisa tomar o cuidado de inserir o ponto de interrogação quando solicita uma informação a Tical por escrito. Em se tratando de uma solicitação por voz, basta pronunciar a pergunta próximo ao microfone. Em sistemas como o clássico – e descontinuado – *IBM ViaVoice*, o ponto de interrogação precisava ser inserido pronunciando a

expressão “ponto de interrogação”, ou pressionando o tipo correspondente no teclado. Ao menos nesse aspecto, Tical é mais prático.

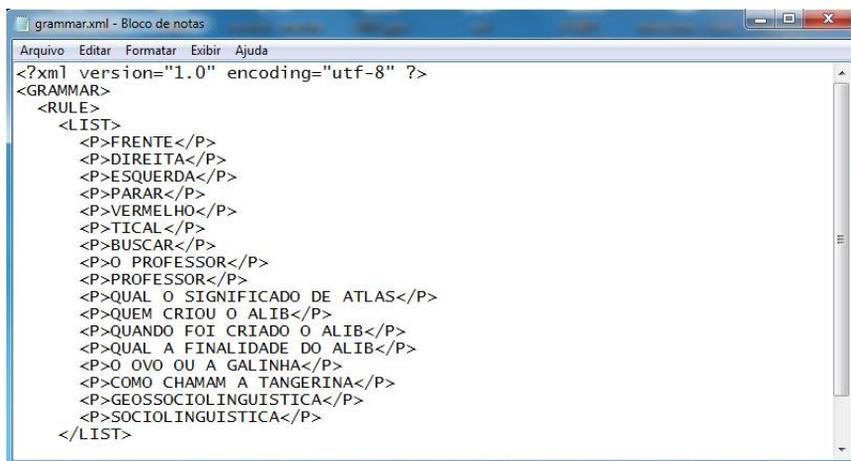
Quanto à capacidade do sistema em distinguir entre um tom de pergunta e um tom de afirmação na voz humana – tarefa simples para qualquer criança, mas um desafio enorme para a máquina – há um estudo em andamento. Ao redor do planeta existem sistemas bastante sofisticados que fazem isso com significativa acurácia. Porém, nenhum deles está disponível em Português do Brasil.

De qualquer forma, na versão aqui comentada, como dito anteriormente, basta formular a pergunta tal como é feita a um interlocutor humano e o sistema inicia a busca. Embora isso pareça óbvio ao leitor-pesquisador, não se pode perder de vista que alguns itens de funcionalidade sempre são levados mais em conta que outros quando se trata de um processo de implementação.

## Aplicação

Antes de colocar o sistema à prova, alguns ajustes como a *taxa de aprovação de reconhecimento*, que implica a proximidade em termos de similaridade daquilo que é pronunciado e aquilo que deve ser reconhecido, são necessários para melhor funcionamento. Os usuários que farão os testes com Tical com comandos por voz, evidentemente, não precisarão se preocupar com esses ajustes, pois tal atividade colocaria em xeque o grau de usabilidade, que se tornaria baixíssimo e inviável para a aplicação em questão.

Para incrementar o banco de dados de reconhecimento de voz de Tical, basta incluir a palavra, expressão ou frase junto à gramática do Coruja que, de imediato, passa a identificar os fonemas falados a partir dos elementos grafemáticos digitados. A Figura 3, a seguir, apresenta alguns dos comandos inseridos e prontos para funcionar. Note-se que há, além das perguntas típicas feitas a Tical, como “qual o significado de atlas” ou “quem criou o alib”, elementos tematicamente isolados, como as indicações vetoriais “frente”, “direita”, “esquerda”, “parar”, e mesmo cores, como “vermelho” entre outros. Essa variedade semântico-lexical deve-se à liberdade dos desenvolvedores em poder testar palavras isoladas e tematicamente mais neutras quando da realização de testes gerais.



**Figura 3** – Tela exibindo uma lista disponível no banco de dados

Após incluídos os itens na gramática, ele passa a reconhecer algumas pronúncias equivalentes. Para o caso de palavras específicas e diretamente relacionadas à área de atuação de Tical, como ‘sociolinguística’ ou ‘geossociolinguística’, o que se espera é que sejam reconhecidas imediatamente e é o que acontece. Quando elas ocorrem na expressão que compõe a pergunta como “o que é sociolinguística?” ou “o que é geossociolinguística?”, cuidados como pronúncia mais empostada e pausada são necessários para melhor funcionamento.

Palavras não existentes oficialmente no léxico do Português do Brasil também podem ser reconhecidas. Para tanto, basta proceder da mesma forma como é feito com palavras existentes, adicionando-as na gramática do sistema, tal como ilustra a Figura 3. Esquisitices como ‘paculafife’ e ‘mocolotate’, por exemplo, não encontráveis – ainda – em sistemas de busca na internet, apresentaram bons resultados. Da mesma forma, unidades lexicais como ‘larango’, ‘sorvetone’, ‘ovonese’, entre outras, criadas por outros processos de formação de palavras denominados *splinters* (ALVES; GONÇALVES, 2014) funcionam bem.

Isso é um claro indício de que o tratamento reservado ao reconhecimento dos fones<sup>2</sup> do Português do Brasil no Coruja foi bastante criterioso. Para constatar e dar mais lastro a essa afirmação, que pode parecer genérica e inconveniente a alguns linguistas, basta realizar um experimento relativamente simples e rápido, bastante acessível a pesquisadores pertencentes ou não das áreas aqui envolvidas. A sugestão é fazer um teste que, embora esdrúxulo e potencialmente risível, é efficientíssimo quando se torna necessário parametrizar critérios de análise nesse âmbito: usar um reconhecedor e/ou sintetizador de voz de outro idioma.

O *Balabolka* pode ser uma das escolhas. Não se trata de um reconhecedor de voz, mas é um bom sintetizador com o recurso de conversão TTS – *text to speech*, ou texto-fala. Depois de baixado – é gratuito – e instalado, é possível utilizá-lo para reproduzir vozes masculina ou feminina em inglês, por exemplo. Com isso em mãos – e ouvidos –, torna-se até divertido ouvir como o sintetizador se ‘embanana’ todo quando o expomos à nossa organização grafemática. Dito de outro modo, ele naturalmente – porque foi programado para isso – vai reproduzir as palavras digitadas em português levando em conta a fonologia do inglês. Nessa condição, a palavra ‘geolinguística’ torna-se um despropósito sonoro. Guardadas as devidas proporções, ‘macaxeira’, que, definitivamente não é um anglicismo e, portanto, inexistente no léxico da língua inglesa, fica quase ininteligível.

Mesmo que pareça uma covardia fazer esse experimento, pelo fato de, notoriamente, o sintetizador em inglês não ter a capacidade de reproduzir os fonemas do Português do Brasil, ele serve para alertar o pesquisador de que o que diferencia um reconhecedor/sintetizador nas diferentes línguas não é simplesmente o banco de dados, mas, paralelamente, o tratamento reservado ao reconhecimento dos fonemas relativos ao idioma em questão. Caso a diferenciação seja apenas o banco de dados, haverá um problema enorme quando um usuário de um sistema como Tical formular uma pergunta que

---

<sup>2</sup> **Fones** estão para os **sons** assim como **fonemas** estão para as **realizações** da fala. Embora muitas vezes coincidam, ou seja, o som passível de ser reproduzido pela voz humana é o mesmo que pode ser transcrito, nem sempre são os mesmos. Em se tratando de reconhecimento e comandos por voz, tanto os fones quanto os fonemas interessam igualmente.

contenha estrangeirismos – a dicotomia saussureana *langue* e *parole* constitui apenas um entre as centenas de exemplos da área de Linguística.

Assim sendo, esse tipo de teste com palavras oficialmente inexistentes não é fortuito. Nas diferentes áreas de conhecimento são recorrentes os verbetes que ainda não estão incorporados no léxico e que, portanto, não figuram nos dicionários de Português Brasileiro. As palavras “geolinguística” e “geossociolinguística”, por exemplo, até o fechamento deste texto, não estavam presentes nos dicionários institucionalizados mais conhecidos<sup>3</sup>. Isso não significa que não possam ser cadastradas no banco de dados de Tical ou que não possam ser reconhecidas grafemática e foneticamente pelas duas frentes do sistema – PLE e PLF.

Os problemas dos ruídos de ambiência, entretanto, responsáveis por muitas falhas em dispositivos comandados por voz, tal como o clássico *IBM Via Voz* (MANFIO, 2014), precisam ser levados em consideração. Lugares com muita reverberação ou naturalmente ruidosos certamente acarretarão mau funcionamento do sistema, seja utilizando o microfone interno do computador ou um sistema externo – microfone de mesa ou incorporado a fones de ouvido. É necessário, também, considerar as diferenças fisiológicas existentes entre pessoas: timbre, potência sonora, entonação, prosódia, entre outros, podem ser empecilhos ao bom funcionamento.

Desconsiderados esses obstáculos – justamente porque, para o teste inicial, foram utilizadas apenas as vozes dos pesquisadores<sup>4</sup> em um ambiente relativamente silencioso – e ajustados os parâmetros comentados há pouco, o aplicativo funciona bastante bem.

---

<sup>3</sup> Mesmo considerando que praticamente não há dicionários cujos discursos não sejam institucionalizados, ou seja, reconhecidos como legítimos por nossas instituições (NUNES, 2002), salvo alguns glossários e listas vocabulares alternativos, todos têm limitações, entre elas a impossibilidade de incorporar rapidamente todas as palavras novas mais utilizadas.

<sup>4</sup> O aporte linguístico para todas as versões de Tical foi dado pelo Professor Edio Roberto Manfio, mas o desenvolvimento ficou sempre a cargo dos professores Fabio Carlos Moreno e Cinthyan R. S. C. de Barbosa, que também contribuíram com suas vozes em muitos experimentos.

## Desempenho com Variantes do Português do Brasil

Qualquer sistema de reconhecimento de voz, ao menos até a data atual, apresenta falhas, como já comentado. Fosse possível desconsiderar – na prática não é – as diferenças fisiológicas existentes entre pessoas, como timbre, entonação, prosódia, que, como dito, são grandes empecilhos ao bom funcionamento de sistemas como esses, ainda seria necessário dar conta dos infinitos ruídos de ambiência originados por veículos, animais, máquinas, vento, chuva e mesmo emissões humanas idiossincráticas de valor linguístico discutível.

Solucionados os problemas da diferença fisiológica e do ruído de ambiência que, notoriamente subdividem-se cada qual em uma ampla gama de possibilidades que se entrecruzam e que têm o potencial quase mágico de criar possibilidades imprevisíveis – como é o caso de uma palavra pronunciada em ambiente que passou a gerar reverberação por ter sido esvaziado de móveis – ainda há a questão das variantes linguísticas. Estas, diferentemente das outras, embora também numerosas, são cientificamente mapeáveis e relativamente previsíveis – considerando que a linguagem humana é dinâmica.

Muito do que já se pesquisou sobre as variantes do Português do Brasil são encontráveis no grande número de textos científicos do Projeto ALiB. Se forem considerados todos os trabalhos – e não apenas aqueles vinculados ao ALiB – sobre essas questões, há concretamente à disposição um significativo mapeamento de todo processo de variação linguística do país que, tranquilamente, tem o potencial de alimentar um sistema bastante robusto em termos de bibliotecas em aplicativos de reconhecimento de voz. Não é à toa que reconhecedores de voz no idioma inglês estão mais avançados em termos de desempenho: os conhecimentos linguísticos sobre essa língua estão sendo aplicados há mais tempo.

Nesse contexto, o Coruja, que dá a Tical o poder da ‘audição’, tem um bom desempenho quando exposto a algumas variantes do Português do Brasil. Depois de inseridas em sua gramática, tal como descrito há pouco, as palavras são reconhecidas com um grau de acurácia significativo, tanto na cadeia da fala – palavras pronunciadas em sequência constituindo uma sintaxe – quanto separadamente – como comandos simples.

Elementos vetoriais como *esquerda, direita, frente, trás e para*, funcionaram muito bem com diferentes variantes. Note-se que, embora ainda pouco usuais ao escopo e à biblioteca de Tical enquanto ‘robô linguista’, são muito úteis para operacionalizar qualquer veículo sobre rodas e, por extensão, menus de alguns softwares especialmente desenvolvidos para isso ou adaptados a uma acessibilidade de telas mais simplificada. O próprio Tical pode ser programado para receber esses comandos e tal procedimento pode auxiliar em muito pessoas com pouco controle motor sobre os membros.

Da mesma forma, nuances cromáticas como *vermelho, verde, azul, branco e amarelo*, já utilizadas em outros experimentos<sup>5</sup> com processadores de sinais da fala, obtiveram bons resultados na busca, tal como apresentado no Quadro 3.

**Tabela 3** – Relação dos comandos e cores utilizados no experimento

variante referência (gravada)	variantes simuladas como comandos (pronunciadas)				
	a	b	c	d	e
[veɾ'meliu]	[veχ'meliu]	[veh'meliu]	[veχ'meliu]	[veh'meliu]	[veɾ'meliu]
['veɾdʒi]	['veχdi]	['vehdi]	['veχdʒi]	['vehdʒi]	['verdʒi]

O Coruja também possibilitou a Tical sair-se melhor – ao menos em alguns aspectos – que o buscador por voz do Google, este avaliado em um experimento linguístico que estava em andamento durante a redação desse texto. Comparativamente, enquanto o referido buscador falhou ao tentar

<sup>5</sup> Os referidos experimentos foram feitos utilizando um dispositivo comercial distribuído pela Elechouse denominado *Voice Recognition Module – V2*. Nele, são gravados os comandos de voz que, depois, ao serem pronunciados, habilitam respectivas portas lógicas capazes de acionar outros dispositivos – como eletrodomésticos, por exemplo.

reconhecer palavras pertencentes a uma mesma esfera de conhecimento, a integração Tical-Coruja teve êxito com o mesmo conjunto lexical.

A exemplo disso, o Quadro 4 apresenta a relação entre um domínio de conhecimento – Linguística – e algumas das modalidades de variações mais conhecidas. Os itens marcados com asterisco foram aqueles que, no experimento com o buscador, não resultaram em sucesso de busca (esse critério também serve para o Quadro 5).

**Quadro 4 – Domínios da Linguística e variações**

<b>Domínio</b>	<b>Variações</b>
Linguística	Idioleto
Dialetologia	Socioleto
Sociolinguística	Diastrática
Linguística textual	Diatópica
Análise do discurso	Diafásica
Geolinguística*	Diamésica*
Geossociolinguística*	Diagenérica*
	Diageracional*
	Diassexual*

Tical não teve problemas com nenhum dos itens destacados com asterisco. Bastou apenas que, como já comentado anteriormente, as palavras fossem inseridas em seus dois bancos de dados – Tical e Coruja.

Outro domínio de conhecimento utilizado, bastante diverso da Linguística, mas equivalente em termos da disposição hierárquica das palavras envolvidas, foi a Eletroeletrônica. Novamente Tical superou o buscador e não teve dúvidas ao retornar a resposta adequada para todos os itens marcados em asterisco e que geraram confusão junto ao conhecido sistema que opera na rede.

**Quadro 5** – Grandezas elétricas e componentes eletrônicos

<b>Grandezas</b>	<b>Componentes</b>
Volt	resistor
Hertz	transistor
Ampére*	capacitor
Watt*	varistor
Ohm*	disjuntor
Farad*	capacitor variável
Coulomb*	

Válido lembrar que, em alguns dos testes feitos nos experimentos com o buscador do Google e, posteriormente, com Tical neste estudo, representados nos Quadros 4 e 5, predominam palavras isoladas, mas também ocorrem expressões como ‘Linguística Textual’, ‘Análise do Discurso’ e ‘capacitor variável’. Nesse âmbito, não é demais dizer que o sistema se sai bem com as vizinhanças fônicas ocasionadas por palavras pronunciadas em sequência na cadeia da fala, fenômeno conhecido pelos desenvolvedores como ‘fala contínua’.

Frases inteiras, no entanto, e não apenas expressões, apresentam algumas falhas cuja origem provavelmente não se encontra em elementos externos – diferenças fisiológicas e ruídos de ambiente – já citados, mas em particularidades do algoritmo do sistema. Interessantemente, Tical responde à pergunta ‘qual a finalidade do atlas’, mas não responde ‘qual a finalidade do alib’. Comparativamente, a proximidade grafemática e fonética das perguntas, tal como representadas no Quadro 6, é grande e, considerando que ambas foram devidamente inseridas na gramática do Coruja, as duas deveriam ser reconhecidas e não apenas uma.

## Quadro 6 – Comparativo fonético e grafemático entre duas frases semelhantes

Pergunta	Transcrição
(i) Qual a finalidade do atlas?	[kwawɛfinalidadʒidwatlɛs]
(ii) Qual a finalidade do alib?	[kwawɛfinalidadʒidwalibi]

Fácil perceber a partir do Quadro 6 que, até o décimo nono fonema, a transcrição pode ser praticamente a mesma – considerando, evidentemente a variante dos pesquisadores. A diferença está nos quatro últimos fonemas. Partindo da hipótese elementar de que o problema seria a última palavra, procedeu-se com a inserção em separado de cada uma delas – ‘alib’ e ‘atlas’, embora, aparentemente, apenas ‘alib’ apresentasse falha. Não houve sucesso. Então, pensando na cadeia da fala e na influência da vizinhança fônica, a suspeita recaiu sobre ‘finalidade do’ – [finalidadʒidw] – e o procedimento foi o mesmo: foram inseridas separadamente ‘finalidade’ e ‘do’. Novamente não houve sucesso.

Assim, acerca dessa curiosa falha, restaram apenas duas hipóteses plausíveis: (i) há um problema, falha ou deficiência no algoritmo; (ii) apenas as vozes – e todas as suas particularidades – dos pesquisadores não constituem uma amostragem válida para colocar o sistema à prova.

Como o objetivo desse estudo não era extinguir dúvidas sobre falhas específicas e pontuais como essa, é pertinente salientar que um estudo linguístico que trate tão somente dessas particularidades, não só é perfeitamente possível como produtivo, uma vez que o sistema encontra-se em operação e várias metodologias podem ser adotadas quando da busca por prováveis desvios durante o seu funcionamento.

### PLE, PLF e Outras Modalidades de Processamento

Observando os resultados gerais dos primeiros testes, foi possível pensar em algumas soluções para que os comandos por voz e a síntese por voz pudessem operar também de modo parcial, com um pouco mais de flexibilidade.

Nesse sentido, verificou-se que seria interessante o usuário poder escolher entre fazer a pergunta por escrito – utilizando o PLE – ou falada – utilizando o PLF. Para tanto, foram consideradas quatro opções de utilização, descritas a seguir.

A primeira opção seria um ambiente naturalmente ruidoso como uma avenida, uma cantina, grupo de pesquisas ou uma oficina em que o usuário, impossibilitado de fazer a pergunta oralmente, escreve-a utilizando-se do teclado e recebe a resposta por escrito na tela. A primeira versão de Tical foi projetada para isso e não estão envolvidos aqui os comandos por voz.

A segunda opção leva em conta as mesmas condições da primeira, com a diferença de o usuário fazer a escolha por receber a resposta ditada, ou seja, por síntese de voz. Há aí a conversão de linguagem escrita para falada, a já citada conversão TTS (MANFIO, 2012, 2013, 2014) e também não estão envolvidos aqui os comandos por voz.

Como terceira opção, o usuário está em casa pesquisando deitado em sua cama e, portanto, livre de quase todos os ruídos externos presentes em primeira e segunda opções. Então, escolhe perguntar oralmente ao Tical, mas receber a resposta por escrito para que não tenha o trabalho de digitar ou transcrever. Há aí a conversão de linguagem falada para escrita ou conversão fala-texto, conhecida como STT – *speech to text* (MANFIO, 2012, 2013, 2014).

Assim, a quarta opção desconsidera o texto escrito para pergunta ou resposta. O usuário faz a pergunta oralmente e recebe a resposta também oralmente. O que ocorre aqui é o reconhecimento de voz capaz de atender ao comando e trazer como resposta um texto em síntese de voz. Guardadas as devidas proporções e o profundo fosso existente entre a realidade e conceitos imagísticos, os recursos do aplicativo chegam bem próximos àqueles dos filmes de ficção científica com a vantagem de lidar tão somente com o Português Brasileiro e falar do nosso próprio atlas linguístico, o ALiB.

Essas quatro opções, no entanto, ainda não estão disponíveis para funcionamento em conjunto – operam parcialmente e em separado. Entretanto, estão sendo observadas com mais critério e em estudos futuros, quando o desenvolvimento das pesquisas em andamento atingirem um grau satisfatório, poderão estar todas concomitantemente operacionais no Professor.

A exemplo disso, Tical já se encontra disponível no WhatsApp, ou seja, é possível consultá-lo em qualquer lugar do Brasil. Para tanto, basta que

se tenha à disposição um smartphone e, evidentemente, uma linha funcional. O que se estuda nesse momento é uma maneira eficiente de incrementar essa capacidade de Tical em operar com esse aplicativo de mensagem multiplataforma para smartphones com PLF, ou seja, utilizar on-line os recursos de comandos e síntese por voz do Professor – terceira e quarta opções comentadas há pouco.

Outro exemplo seria dar a Tical o poder de reconhecer a face de seu interlocutor. Este estudo também está em andamento e tem o objetivo de, primeiramente, promover mais um aspecto de ambiência amigável. Fazendo isso, o sistema pode tornar-se muito mais atrativo, uma vez que a conversa pode ser iniciada com uma saudação por parte de Tical: “Bom dia, Saussure! Que bom que voltou. Não conversamos há dois dias. O que deseja saber?”. A partir daí, entre outras coisas, é possível fazer com que o sistema resgate uma espécie de lista de preferências do interlocutor de modo que Tical pode continuar a conversa de onde parou ou simplesmente agilizar uma nova busca, a critério do usuário.

### **Considerações Finais**

Weinzembaun, Turing e Descartes, cada qual por motivos diferentes, provavelmente observariam o funcionamento de Tical com bastante critério e curiosidade. Não porque seria total novidade ou porque estaria isento de críticas a despeito de seu funcionamento, mas porque constitui um modo de aplicação diversificado para várias tecnologias e áreas de conhecimento.

Nesse sentido, apresentar o robô de conversação Tical em uma dimensão mais ampla, ou seja, com todas as funcionalidades que a implementação tornou possíveis até o fechamento deste texto, foi muito produtivo. Constituiu, também, uma tarefa árdua, não porque tenha sido necessário começar do zero ou porque faltavam ferramentas ou conteúdo, mas em função de ter sido realizada por uma equipe pouco numerosa.

O esforço, no entanto, foi recompensado. Nessa versão, Tical se mostra como um projeto bastante relevante em termos científicos justamente por concentrar vários diferenciais já comentados aqui: é um dos únicos robôs que ‘versam’ sobre Linguística; está todo em Português Brasileiro; opera em dispositivos móveis a partir de aplicativo de mensagem instantânea; possui síntese de voz e atende a comandos por voz.

Por esses motivos, seu grau de aplicabilidade é bastante alto e o trabalho tornou-se, ao mesmo tempo, projeto e objeto de pesquisa. Em outras palavras, o trabalho serviu não apenas para aplicar teorias das áreas relacionadas, mas também para testar a funcionalidade de comandos por voz com Português Brasileiro e conteúdo afim: o ALiB.

Reconhecidas plenamente as funcionalidades de Tical após essa fase prototipal, resta dar ênfase e continuidade ao projeto. Para tanto, existe como opção a montagem de uma equipe técnica que pode constituir-se de um grupo de pesquisa interdisciplinar – ou vários deles. A vantagem de grupos de pesquisa devidamente cadastrados junto a órgãos de fomento, por exemplo, é a de que vislumbram possibilidade de apoio financeiro às pesquisas, sem o qual, historicamente, muitas boas ideias tiveram como destino o grande baú do esquecimento.

Independentemente da existência de grupos ou equipes vindouros, o estudo já serviu para comprovar a dimensão funcional de Tical enquanto conceito, para permitir a utilização de conhecimentos teóricos adquiridos e para projetar, a partir dos resultados obtidos, alguns outros caminhos possíveis além daqueles listados há pouco: Tical pode ainda operar com recursos de Realidade Virtual e/ou Aumentada e uma versão holográfica do docente cibernético seria apenas mais alguns passos adiante.

## Referências

ALVES, R. S.; GONÇALVES, C. A. V. O processo de formação de palavras com os splinters –nese –nejo e –tone. *Entretextos*, Londrina, v. 14, n. 1, p. 27-42, jan./jun. 2014. Disponível em: <<http://bit.do/cWELP>>. Acesso em: 12 jan. 2015.

BALABOLKA. Disponível em: <<http://bit.do/cWELZ>>. Acesso em: 13 jan. 2012.

BARBOSA, C. R. S. C. de. *Técnicas de parsing para gramática livre de contexto lexicalizada da língua portuguesa*. 2004. Tese (Doutorado em Engenharia Eletrônica e Computação) – Instituto Tecnológico de Aeronáutica, São José dos Campos.

- CARDOSO, S. A. M. et al. *Atlas Linguístico do Brasil: Introdução*. v. 1. Londrina: Eduel, 2014a.
- CARDOSO, S. A. M. et al. *Atlas Linguístico do Brasil: Cartas Linguísticas I*. v. 2. Londrina: Eduel, 2014b.
- COMITÊ NACIONAL DO PROJETO ALIB. *Atlas Linguístico do Brasil: questionários*. Londrina: Editora UEL, 2001.
- DESCARTES, R. *Discurso do método*. Tradução de Maria Ermantina Galvão. São Paulo: Martins Fontes, 2001.
- IBM. Disponível em: <<http://bit.do/cWEMe>>. Acesso em: 03 jun. 2013.
- IBM. *Fazendo o Watson trabalhar*. Disponível em: <<http://bit.do/cWEMg>>. Acesso em: 22 mar. 2015.
- IBM Watson. *What is Watson*. Disponível em: <<http://bit.do/cWEMj>>. Acesso em: 13 jan. 2015.
- LAPS - Laboratório de Processamento de Sinais. Disponível em: <<http://bit.do/cWEMk>>. Acesso em: 11 out. 2014.
- MANFIO, E. R. Como funcionam alguns fonemas no aplicativo Balabolka. *Revista Via Litterae*, Anápolis, v. 4, n. 2, p. 191-204, jul./dez. 2012. Disponível em: <<http://bit.do/cWEMp>>. Acesso em: 21 jun. 2014.
- MANFIO, E. R. Aspectos fonéticos no DosVox enquanto aplicativo tipo texto-fala. In: SEMINÁRIO DO CELLIP, 21., 2013, Paranaguá. *Anais...* Paranaguá: FAFIPAR; UFPR Litoral, Instituto Federal, 2013. p. 358-372. Disponível em: <<http://bit.do/cWEMB>>. Acesso em: 21 jun. 2014.
- MANFIO, E. R. O ALiB no IBM Via Voice: pesquisa geossociolinguística associada a aplicativo de reconhecimento de fala. In: SEMINÁRIO DE DISSERTAÇÕES E TESES EM ANDAMENTO – SEDATA, 13., 2014, Londrina. *Caderno de Resumos*. Londrina: UEL, 2014. p. 35.

MANFIO, E. R.; MORENO, F. C.; BARBOSA, C. R. S. C. de. Tecnologia Interativa Conversacional sobre Assuntos Linguísticos – Tical: linguagem e significação. In: SELISIGNO, 9.; SIMPÓSIO DE LEITURA DA UEL, 10., 2014, Londrina. *Caderno de Resumos*. Londrina: UEL, 2014a. p. 54-55.

MANFIO, E. R.; MORENO, F. C.; BARBOSA, C. R. S. C. de. Professor Tical: robô de conversação sobre Dialetologia e Geossociolinguística. In: CONGRESSO INTERNACIONAL DE DIALETOLOGIA E SOCIOLINGUÍSTICA – CIDS, 3., 2014, Londrina. *Caderno de Resumos*. Londrina: UEL, 2014b. p. 48.

NUANCE – Loquendo. Disponível em: <<http://bit.do/cWEM8>>. Acesso em: 18 out. 2014.

NUNES, J. H. Dicionarização no Brasil: condições e processos. In: NUNES, J. H.; PETTER, M. (Org.). *História do saber lexical e a construção de um léxico brasileiro*. São Paulo: Humanitas; FFLCH; USP; Campinas: Pontes, 2002. p. 99-119.

PROJETO Atlas Linguístico do Brasil. Disponível em: <<http://bit.do/cWENn>>. Acesso em: 02 fev. 2013.

RICH, E. *Inteligência artificial*. Tradução de Maria Cláudia Santos Ribeiro Ratto. São Paulo: Makron Books, 1993.

SCHILDT, H. *Inteligência artificial utilizando linguagem C*. Tradução de Cláudio Gaiger Silveira e Mônica Soares Rufino. São Paulo: McGraw-Hill, 1989.

TURING, A. M. Computing machinery and intelligence. *Mind*, v. 59, n. 236, p. 433-460, Oct. 1950. Disponível em: <<http://bit.do/cWENp>>. Acesso em: 13 out. 2014.

ZIVIANI, N. *Projeto de algoritmos: com implementação em Pascal e C*. 4. ed. São Paulo: Pioneira, 1999.

WEIZENBAUM, J. ELIZA – A computer program for the study of natural language communication between man and machine. *Communications of the ACM*, v. 9, n. 1, p. 36-45, jan. 1966.

WHATSAPP. Disponível em: <<http://bit.do/cWENt>>. Acesso em: 17 out. 2014.

Recebido em: 01/03/2016

Aceito: 23/05/2016