INTERFACE TECNOLÓGICA

CHATBOT COMO INTERFACE PARA A APRENDIZAGEM DA LÍNGUA INGLESA

CHATBOT AS INTERFACE FOR ENGLISH LANGUAGE LEARNING

Edinei Edson Antunes – edinei.antunes01@fatec.sp.gov.br
Mirela de Lima Piteli Picchi – mirela.piteli@fatetq.edu.br
Giuliano Scombatti Pinto – giuliano.pinto@fatectq.edu.br
Adriana Valim Ferreira Ragazani – adriana.ragazani@fatectq.edu.br
Faculdade de Tecnologia de Taquaritinga (FATEC) – SP – Brasil

DOI: 10.31510/infa.v15i1.357

RESUMO

Neste artigo foi proposto o uso de um *chatbot* para a aprendizagem e aprimoramento da língua inglesa, fazendo uso da grande presença dos dispositivos móveis atuais. Realizou-se uma revisão e análise bibliográfica sobre a aprendizagem de um idioma, a forma com que a tecnologia tem influenciado a aquisição da língua inglesa e sobre a possibilidade de transcender as salas de aula no que concerne ao aprendizado de uma nova língua. Além disso, buscou-se apresentar, também, o rápido crescimento da mobilidade de dispositivos eletrônicos, o cenário das plataformas móveis e as tecnologias que se destacam para o desenvolvimento de aplicativos nativos, hibridização como estratégia para o desenvolvimento de aplicativo móvel, e a viabilidade do uso de *chatbot* para o aprimoramento da língua inglesa. Para tanto, utilizou-se um renomado corpus linguístico estatístico e optou-se pelo *IBM Watson* como ferramenta de análise a partir do qual constatou-se o grande desafio de viabilizar a interação com o usuário de maneira mais eficaz. Os resultados preliminares de análise apontam que existe a real possibilidade de treinamento do *chatbot* com o corpus da língua inglesa para tal interação, através da frequência de uso das palavras em frases nos diálogos em inglês, para fins de aquisição e aprimoramento do referido idioma.

Palavras-chave: Dispositivos Móveis. Chatbot. Aprendizagem de Língua Inglesa.

ABSTRACT

In this article we have proposed the use of a chatbot for learning and improving the English language, making use of the great presence of the current mobile devices. A literature review and analysis was carried out on language learning, on the way technology has influenced the acquisition of the English language and the possibility of transcending classrooms in terms of learning a new language. In addition, we sought to present the rapid growth of electronic device mobility, the mobile platforms scenario and the technologies that stand out for the development of native applications, hybridization as a strategy for mobile application development, and chatbot use feasibility for the improvement of the English language. To do so, we used a renowned statistical linguistic corpus and chose IBM Watson as an analysis tool from which it was verified the great challenge of interacting with the user more effectively. The preliminary results have shown that there is an actual possibility of chatbot training with the corpus of the English language for such interaction, through the frequent use of selected words in sentences

in English dialogues, for the purpose of acquiring and improving that language.

Keywords: Mobile Devices. Chatbot. English Language Learning.

1 INTRODUÇÃO

Devido à importância que a língua inglesa possui no contexto da Tecnologia da Informação, da Pesquisa Científica, e da facilidade de mobilidade dos dispositivos eletrônicos, como *smartphones* e *notebooks* com acesso à internet, faz-se necessário buscar as ferramentas adequadas para a aprendizagem do referido idioma. De acordo com Crystal (2003), uma língua alcança um *status* verdadeiramente global quando desenvolve um papel especial reconhecido em todos os países e o inglês, claramente, alcançou tal papel. Além disso, é em inglês que se fecham praticamente todos os negócios internacionais, bem como se registram os mais importantes avanços científicos, em revistas e periódicos referenciados mundialmente (SCHWARTSMAN, 2011). Assim, a justificativa para a realização desta pesquisa jaz exatamente nessa necessidade de se aprender uma língua internacional, bem como explorar as facilidades advindas da mobilidade e conectividade, buscando as ferramentas mais apropriadas para tal aprendizado. Nesse contexto, os aplicativos para dispositivos móveis têm se mostrado muito úteis e apropriados, em especial os que utilizam o *chatbot* como uma das ferramentas para a prática e aquisição de línguas estrangeiras, de modo geral. Desta feita, o objetivo principal deste artigo é realizar um levantamento bibliográfico das principais teorias que embasam a elaboração de desenvolvimento de aplicativos híbridos voltados para a aprendizagem de língua estrangeira (inglês). Além disso, traz uma breve apresentação do *IBM Watson™ Assistant*, seu funcionamento básico e sua aplicação para a aprendizagem do inglês, propondo o desenvolvimento de um aplicativo para aprendizagem e aprimoramento da língua inglesa, fazendo uso da ferramenta *chatbot.* A estrutura do trabalho consiste em uma seção destinada à revisão bibliográfica dos principais conceitos teóricos que embasam o estudo (fundamentação teórica), seguida pelos procedimentos metodológicos que caracterizam a pesquisa (seção 3). Na sequência, são apresentados os resultados preliminares acerca do desenvolvimento do aplicativo com uso de *chatbot* e, por fim, algumas considerações finais.

2 CONCEITUAÇÃO DAS TECNOLOGIAS APLICADAS

O objetivo desta seção é apresentar os principais conceitos teóricos que embasam este trabalho, a saber: definição e uso de *chatbot*, aplicativos híbridos, multiplataformas, aprendizagem de línguas assistida pelo computador (do inglês *Computer Assisted Language Learning* - CALL) e aprendizagem de línguas assistida por dispositivos móveis (do inglês *Mobile Assisted Language Learning* – MALL).

2.1 Definição e uso de chatbot

Chatbots ou *chatterbots* são agentes virtuais de conversação, que possuem inteligência limitada, mas com capacidade de cognição (LAVEN, 2001). De acordo com Teixeira et al. (2005, p.2), trata-se de "um programa de computador que tenta simular um ser humano na conversação com as pessoas". *Chatbots* têm por objetivo, por exemplo, responder às perguntas de tal forma que se tenha a impressão de estar conversando com outra pessoa, e não com um programa de computador. Ainda segundo os autores, "após o envio de perguntas em linguagem natural, o programa consulta uma base de conhecimento e em seguida fornece uma resposta que tenta imitar o comportamento humano" (TEIXEIRA et al., 2005, p. 2). Eles podem ser empregados em diversos setores, sendo responsáveis por simples tarefas, tais como: atendimento em *sites* de *e-commerce*, de compra de passagem aérea, de dúvidas sobre um produto ou serviço e, também, no aprendizado de uma língua estrangeira.

Esse último uso dos *chatbots*, nos aproxima do que afirma Kenski (2007, p. 19) "para que o uso das tecnologias possa alterar o processo educativo, elas devem ser inseridas pedagogicamente, sendo eficaz sua aplicação para, assim, gerar impacto no aprendizado".

2.2 Aplicativos Híbridos

Esta seção tem como objetivo abordar as plataformas presentes hoje nos dispositivos móveis, bem como apresentar o que são aplicativos híbridos e suas vantagens em relação aos aplicativos nativos.

Segundo Guenveur (2017), os dispositivos móveis presentes no Brasil fazem uso das plataformas IOS (*Apple*) e *Android (Google)* em sua maioria, os quais somados, representam 97,3%. Percebe-se com esses números, que são diversas as tecnologias empregadas nestes

dispositivos. As que estão presentes no sistema da *Google*, de código aberto, rodam nativamente aplicações em *Java* e *Kotlin*. Já a plataforma proprietária da *Apple*, faz uso nativo do *Object-C* e do *Swift.*

Devido à variedade de produtos no cenário dos dispositivos móveis, surge o conceito de aplicativos híbridos, isto é, um único aplicativo desenvolvido para todas as plataformas, englobando, assim, os 97,3% de dispositivos que existem e suas devidas plataformas (GUENVEUR, 2017). Esta alternativa mais rápida e com menor custo, diminui o tempo de desenvolvimento de aplicativos.

2.3 Tecnologia e aprendizagem de línguas – CALL e MALL

A tecnologia tem feito parte do ensino e aprendizagem de línguas há décadas e tal relação se intensificou com os laboratórios de línguas nas décadas de 1960 e 1970. É nessa época, também, que começou a se desenvolver uma nova área no ensino de línguas, que estuda a sua relação com a tecnologia no aprendizado de línguas (MARTINS; MOREIRA, 2012). A partir dos anos de 1980, com o surgimento dos computadores pessoais, essa área de estudos passa a ter uma denominação própria: *Computer Assisted Language Learning* (CALL) – que em português significa Aprendizagem de Línguas Assistida pelo Computador.

No Brasil, de acordo com Reis (2008), as pesquisas começaram a surgir a partir de 1998, e, nas últimas décadas, com o aparecimento e o aprimoramento das novas tecnologias, aumentaram significativamente as funções desempenhadas por computadores, o que ampliou o campo de ensino de línguas mediado pelas ferramentas CALL. Sem dúvida, essa dimensão interativa, que é uma parte indispensável do processo de aprendizagem de línguas, trouxe uma série de benefícios práticos para essa atividade e, na última década, com a disseminação de *smartphones* e *tablets,* foi introduzida a expressão "aprendizagem de idiomas assistida por dispositivos móveis", do inglês *Mobile Assisted Language Learning* (MALL). Trata-se de uma subdivisão da aprendizagem de línguas assistida por computador CALL, que se dá a partir do uso de dispositivos móveis para apoiar tal aprendizagem (GUO, 2016). Em linhas gerais, o uso de dispositivos portáteis e pessoais diferencia o MALL do CALL, e tais dispositivos permitem a utilização contínua e espontânea, acesso e interação em diferentes contextos, promovendo novas formas de aprendizagem a partir de aplicativos dinâmicos e atrativos.

3 PROCEDIMENTOS METODOLÓGICOS

Este é um estudo de base bibliográfica, pois busca conhecer e aprofundar as principais contribuições teóricas sobre o tema. Segundo Köche (2002, p. 122), esse é o tipo de pesquisa "que se desenvolve tentando explicar um problema, utilizando o conhecimento disponível na área, identificando as teorias produzidas, analisando-as e avaliando sua contribuição para auxiliar a compreender ou explicar o problema objeto da investigação". Kochë também elucida que a relevância da pesquisa bibliográfica para um trabalho científico está no fato de que, ao realizar um levantamento bibliográfico, o pesquisador amplia seu grau de conhecimento em uma determinada área, a fim de utilizá-lo como "fundamentação na construção de um modelo teórico explicativo de um problema" (KÖCHE, 2002, p. 122).

No que concerne à abordagem geral para investigar e avaliar o aplicativo móvel em questão neste estudo, tem-se como base uma abordagem qualitativa, a partir da análise do conteúdo do material, sua criação e desenvolvimento, bem como sua relevância para fins de aprendizagem de língua inglesa.

Já no que diz respeito ao desenvolvimento do aplicativo e do *chatbot* propriamente dito, houve a preocupação inicial com a seleção do vocabulário da língua inglesa que iria compor as atividades. Nesse caso, após diversas pesquisas sobre o assunto, optou-se por utilizar a base de dados da *Cambridge*, o *Cambridge English Corpus* (CEC, 2018), que consiste em um *corpus* linguístico reconhecido internacionalmente, composto por bilhões de palavras da língua inglesa, e por dados de várias fontes, incluindo inglês escrito e falado, britânico e americano (CAMBRIDGE UNIVERSITY PRESS, 2018). A partir desse *corpus*, foi feito um recorte das 10.000 palavras mais recorrentes do idioma, com o intuito de compor o banco de dados das atividades do aplicativo.

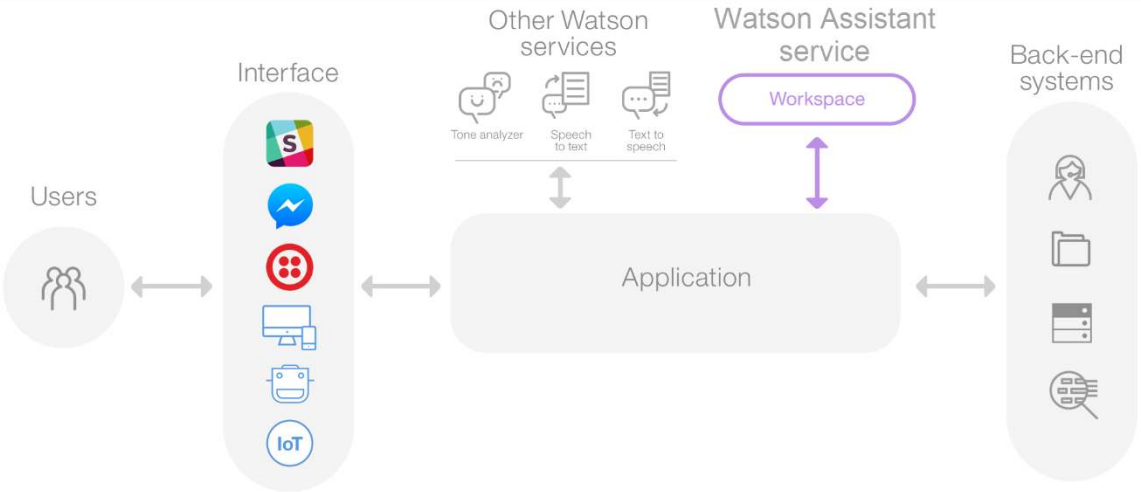
4 RESULTADOS E DISCUSSÃO

Esta seção visa apresentar a estrutura do aplicativo, bem como a ambientar o funcionamento do *IBM Watson™ Assistant*, seu funcionamento básico e aplicação para a aprendizagem do inglês.

4.1 Estrutura do aplicativo

O *IBM Watson™ Assistant* consiste em uma ferramenta que possibilita a criação de aplicações com cognição, a qual utiliza o processamento da linguagem natural, viabilizando a interpretação e interação do agente virtual (*chatbot*) com o usuário, através de mensagens de texto dentro do aplicativo, como demonstrado na figura 1 a seguir:

Figura 1 – Fluxograma básico de uma aplicação *no IBM Watson™ Assistant.*



Fonte: Azraq et al. (2017, p. 2)

Como destacado na Figura 1, o *chatbot* funciona apenas como uma *interface* com o usuário, pois todo processamento cognitivo é feito pelo *Watson*, sendo consumido através de uma *Application Programming Interface* (API), no formato *JavaScript Object Notation* (JSON).

O aplicativo faz o uso do Processamento da Linguagem Natural, do inglês, *Natural Language Processing* (NLP), que é a maneira com que as pessoas se comunicam em um determinado idioma, no caso, o inglês. Por meio do NLP, é possível identificar informações-chave no diálogo, as quais são rotuladas de duas maneiras pelo Watson: intenção (*intent)* e entidade (*entity*). A primeira, identificada pelo símbolo #, consiste na intenção do usuário e alude a ações identificadas por verbos, que indicam um objetivo, isto é, o tema sobre o qual o usuário quer discorrer no diálogo; a segunda, denominada *entity* e identificada pelo símbolo @, refere-se ao contexto da intenção. Neste sentido, considera-se entidade tudo aquilo que está relacionado à ação do usuário, como objetos e lugares, expressos por substantivos. Toda esta análise depende do *input* do usuário.

Usuário: *How are you, Mr. Chatbot?. How's the weather today?*

#greeting *@weather*

Mr. *Chatbot*: *I'm fine. The weather is pretty good.*

Neste exemplo, pode-se perceber como o *Watson* detecta a *intent #greeting* e a *entity @weather*, podendo assim, buscar frases que podem ser usadas neste contexto:

Em resposta a *intent #greeting*, foi usada a expressão *I'm fine* e para a *entity* @weather, optou-se pela expressão *the weather is pretty good*.

Em um cenário real, o usuário sempre faria o *input* das frases de maneira correta e concisa com o que acontece em um diálogo. Neste contexto, porém, existe a funcionalidade de nó (*node*) no *Watson*, que consiste em alternativas, caminhos que o diálogo poderia seguir. Após identificar os nós, procura-se a melhor resposta para a entrada de dado.

4.1.1 Aplicação para a aprendizagem do inglês

Como visto anteriormente, o uso adequado das *intents* e *entities* possibilita a marcação de informações importantes nas frases do usuário, este é o primeiro processamento que o *Watson* faz. Neste sentido com o propósito de adquirir melhoria no nível da língua em questão, foi proposto o uso do CEC como base de conhecimento e treinamento do *chatbot*, como visto na seção Metodologia. Mas apenas identificar as palavras importantes no texto não é suficiente para que haja uma comunicação eficaz entre o emissor e o receptor, é necessário trazer um contexto para estas palavras, com isto, manter um diálogo compreensível de ambos os lados. Para tanto, foi utilizado o Corpus Estatístico, uma importante ferramenta para a elaboração deste ambiente, que consiste na frequência com que estas palavras são usadas para a formatação de frases (PARTINGTON, 1998), e com que palavras estão associadas (FIRTH, 1957). O quadro 1 apresenta uma breve amostra de como foi abordada:

Quadro 1 – Demonstração da utilização do Corpus Estatístico

Palavra	Substantivos	Diversos	Verbos	Adjetivos
good	*morning, thing, news, time, night, idea, job, evening, luck, reason, friend, man*	*very, feel, pretty, bad, enough, sound, welcome,*		

		excellent.		
weather	*condition, •service, •pattern, •report, •forecast, winter• , news•, change, •station*		*•turn, predict, •permit, check•, •warm, affect, •clear, •cooperate, fly, control•*	*bad, cold, hot, national, cool, dry, severe, wet, local, nice*

Fonte: Elaborado pelos autores com base em Davies e Gardner (2010)

As palavras em inglês *good* e *weather,* em um diálogo estão comumente ligadas as palavras destacadas no quadro 1. Com esta informação, pode-se realizar o treinamento do *chatbot*, associando de maneira mais eficiente as *intents* e *entities*, obtendo-se, assim, o contexto para diálogos. Abaixo, apresenta-se uma amostra sobre como seria o diálogo perfeito entre o usuário e o *chatbot*.

Usuário: *Good Morning, Mr. Chatbot? How are you doing?*
Mr. *Chatbot*: *I'm very good. And you?*
Usuário: *I'm OK too. How's the weather today?*
Mr. *Chatbot*: *The weather is pretty good.*

Especificando #*intent* no *Watson* para a palavra *good*: #*good*
Tipificando exemplos de frases usados na *intent #good*:
Good morning
Good night
Good evening
I'm very good
The weather is pretty good

Sempre que for usada a palavra *good* em uma conversa dentro do aplicativo, as frases proposta na *intent*, serão usadas pelo *chatbot* e também pelo usuário, pois a máquina já conhece a palavra, e as frases que podem ser usadas nesse ambiente de conversa. Isso ocorre porque o aplicativo já passou pelo treinamento, já possui cognição suficiente para fazer uso destas *intents* em um diálogo.

Especificando @*entity* para a contextualizar a *intent #weather*: @*umbrella.* As *entities* no *Watson*, usam 2 campos: *Value* (valor), e *Synonym* (Sinônimo). Neste caso, pensando em um diálogo referente a roupas, pode-se usar a entidade @*shirt*, que engloba todos os tipos de camisas e camiseta, permitindo, assim, usar o valor para especificar qual estamos identificando como, por exemplo, *t-shirt* (camiseta), e seu sinônimo *tee*, termo informal também usado para para se referir ao vocábulo *t-shirt*. Após evidenciar o *value* e o *synonym* em qualquer contexto, é possível usá-los em um diálogo, o que permite invocar a entidade a que pertencem. No caso o @*shirt*.

Value: t-shirt
Synonym: tee

Com base na aplicação do *corpus* do inglês para uso no treinamento de *chatbot*, visando a melhoria no nível do inglês dos usuários, viabiliza-se o desenvolvimento de um aplicativo híbrido móvel para o aprimoramento do idioma, pois o *corpus* do inglês em questão identifica as palavras comumente usadas em diálogos e seus contextos, sendo mais eficiente para o funcionamento do *chatbot*. Dessa forma, consegue-se assim, pelo que foi pesquisado, uma maior probabilidade de acertos em simulação de diálogos com humanos.

5 CONSIDERAÇÕES FINAIS

Por se tratar de uma pesquisa em andamento, os resultados preliminares obtidos permitem analisar e avaliar parcialmente a aplicação do referido *corpus* linguístico através de um *chatbot* para fins de aprendizagem de línguas. Sabe-se que aplicativos móveis podem significar oportunidades interessantes para ambientes personalizados e centrados no aluno, com acesso flexível à aprendizagem, afinal, são materiais que ficam disponíveis a qualquer hora e em qualquer lugar. Novas formas de se aprender um idioma tem um grande potencial para aumentar a motivação de aprendizagem e encorajar o usuário a avançar nos níveis, buscando fluência e aperfeiçoamento linguístico.

REFERÊNCIAS

AZRAQ, A.; AZIZ, H.; NAPPE, N.; BRAVO, C.R.; SRI, L. **Building Cognitive Applications with IBM Watson Services**: Volume 2 Conversation, Redbooks, 2017. Disponível em <http://www.redbooks.ibm.com/redbooks/pdfs/sg248394.pdf> Acesso em: 20 abr. 2018.

CAMBRIDGE ENGLISH CORPUS. Disponível em < http://www.cambridge.org/about-us/what-we-do/cambridge-english-corpus > Acesso em: 10 fev. 2018.

CAMBRIDGE UNIVERSITY PRESS. Disponível em < http://www.cambridge.org/ > Acesso em: 10 fev. 2018.

CRYSTAL, D. **English as a global language**. 2nd Edition, New York, Cambridge University Press, 2003.

DAVIES, M.; GARDNER, D. **A Frequency Dictionary of Contemporary American English**. Nova Iorque, Nova Iorque, Routledge, 2010.

FIRTH, J. R. **Papers in linguistics – 1934-1951**. Oxford, Oxford University Press, 1957.

GUENVEUR, Lauren. **Apple é principal marca de smartphone nos EUA e Reino Unido**. 2017. Disponível em: <http://br.kantar.com/tecnologia/m%C3%B3vel/2017-/fevereiro-comtech-smartphone-dados-ios-android-market-share/>. Acesso em: 11 abr. 2018.

GUO, H. **Analysing and Evaluating Current Mobile Applications for Learning English Speaking.** Birkbeck, University of London, 2016. Disponível em <https://englishagenda.britishcouncil.org/sites/default/files/filefield_paths/analysing_and_evaluating_current_mobile_applications_v2.pdf> Acesso em: 02 abr. 2018.

KENSKI, V. M. **Educação e Tecnologias: O novo ritmo da informação**. 2ª edição. Campinas – SP: Papirus, 2007.

KÖCHE, J. C. **Fundamentos de metodologia científica: Teoria da ciência e iniciação à pesquisa**. Petrópolis; Vozes, 2002.

LAVEN, Simon. **The Simon Laven Page**. 2001. Disponível em: < http://www.simonlaven.com >. Acesso em: 30 abr. 2018.

MARTINS, C. B. J.; MOREIRA, H. O campo CALL (Computer Assisted Language Learning): definições, escopo e abrangência. In: **Unisinos**, vol. 10, n. 3, p. 247-255, set/dez 2012. Disponível em < http://revistas.unisinos.br/index.php/calidoscopio/article/view/3254/1280> Acesso em: 02 mai. 2018.

PARTINGTON, A. **Patterns and meanings: using corpora for English language research and teaching**. Amsterdã/Filadélfia, John Benjamins, 1998.

REIS, S. C. dos. As tendências teóricas em estudos de CALL no Brasil: identificando o estado da arte. In: **ENCONTRO DO CELSUL**, 8, Porto Alegre, 2008. Porto Alegre, 2008, p. 1-21. Disponível em < www.celsul.org.br/Encontros/08/estudos_de_call.pdf> Acesso em: 03 mai. 2018.

SCHWARTSMAN, H. Não utilizar o inglês é arremessar-se para fora do mundo. **Folha de São Paulo**, São Paulo, 07 fev 2011. Caderno 9, p. 2.

TEIXEIRA, S.; RAMIRO, T. B.; OLIVEIRA, E.; MENEZES, C. S. Chatterbots em ambientes de aprendizagem – uma proposta para a construção de bases de conhecimento. In: **Congresso da Sociedade Brasileira de Computação**, XXV, São Leopoldo, 2005. Disponível em <http://www.br-ie.org/pub/index.php/wie/article/view/863/849> Acesso em: 08 abr. 2018.