

OTHERO, G. A.; MENUZZI, S. M. *Linguística Computacional: teoria & prática*. São Paulo: Parábola, 2005. 126 p.

letrônica

Aline Lorandi

Entre as poucas certezas que se tem sobre o futuro, umas delas resulta indubitável: a evolução assombrosa da informática. Sob essa perspectiva, o linguista se encontra frente a uma importante missão, que é a de promover a interação natural – ou o mais natural possível – entre máquina e ser humano. Para que isso seja possível, as áreas pertinentes à Linguística e à Computação são articuladas e belos trabalhos surgem dessa articulação.

Gabriel de Ávila Othero e Sérgio de Moura Menuzzi estabelecem grandes parcerias que permeiam as 126 páginas de seu livro “Linguística Computacional: teoria & prática”. Os autores, além de fornecerem um rico e atualizado sustento teórico para o desenvolvimento da prática sugerida no livro, de forma notavelmente didática e clara, combinam: Linguística e Computação, sintaxe e semântica, história e atualidade, atualidade e futuro, humanos e robôs.

Em uma conversa com o leitor, Othero e Menuzzi conduzem o linguista, assim como os demais interessados, a percorrerem a interface entre Linguística e Computação, mostrando que, em um futuro próximo, a Inteligência Artificial (IA) estará ainda mais “humanizada”. Este é o trabalho da Linguística Computacional: explorar “as relações entre linguística e informática, tornando possível a construção de sistemas com capacidade de reconhecer e produzir informação apresentada em linguagem natural” (cf. Vieira e Lima 2001:¹).

Uma breve introdução leva o leitor ao ano de 2025, numa simulação de conversa entre a máquina e o seu usuário, em que a comunicação com o computador é estabelecida oralmente, em

¹ Citado por Othero e Menuzzi, p. 22.

linguagem natural. Esta é uma realidade sugerida pela velocidade incrível com que a tecnologia tem avançado e pela “tendência cada vez maior de buscar uma melhor interação entre homens e máquinas nos avanços tecnológicos”².

Tecnologia e Linguagem dão o título ao primeiro capítulo, que visa a mostrar ao leitor que, frente ao comprovável avanço tecnológico, a tendência à implementação de meios que possibilitem e facilitem a interação entre humanos e máquinas também cresce e origina, através do estudo aprofundado da linguagem humana e da informática, a necessidade da “humanização” do computador. Os autores ainda antecipam, nesse capítulo, que a Lingüística Computacional pode ser dividida em duas subáreas: a lingüística de *corpus* e o processamento de linguagem natural (PLN), sendo que a primeira é dedicada à busca de dados lingüísticos, extraídas por meio de computador, que servirão para a pesquisa de uma língua ou variedade lingüística (cf. Berber Sardinha, 2000:2³), e a segunda, voltada à “construção de softwares e sistemas computacionais específicos”⁴, como tradutores automáticos, *parsers*, *chatterbots*⁵, etc.

No capítulo 2, os autores apresentam “Trabalhos em Lingüística Computacional”, divididos em 3 subcapítulos, os quais fornecem suporte para a prática que se encontra no capítulo 3. Aplicações da Lingüística à Computação, nas áreas da fonética e da fonologia, e a importância da sintaxe e da semântica no desenvolvimento de sistemas como os chamados “*chatterbots*”, são tema do primeiro subcapítulo, intitulado “Desenvolvimentos em Lingüística Computacional”. Os autores ilustram com uma conversa o funcionamento do primeiro *chatterbot*, criado em 1966 para ser uma espécie “psicólogo virtual”, além de discorrerem sobre o “Teste de Turing”, responsável pela postulação de um parâmetro para a avaliação de programas de IA. Ademais, a importância de outras áreas da Lingüística, além das supracitadas, para a sofisticação de tais programas, também é salientada.

No segundo subcapítulo, Othero e Menuzzi explicam o que é um *parser* e qual é a sua utilidade. Os autores fazem uma explanação histórica do desenvolvimento de *parsers* e elucidam o papel da sintaxe no funcionamento de aplicativos em PLN. O *parser* – ou analisador gramatical – é desenvolvido em uma linguagem de programação denominada Prolog (PROgramming in

² Othero e Menuzzi, p. 11.

³ Em Othero e Menuzzi, p. 23.

⁴ Othero e Menuzzi, p. 24.

⁵ Para a consulta do significado de termos específicos, os autores disponibilizam um glossário. É importante salientar também que há várias sugestões de sites em que o leitor encontra material teórico e prático em Lingüística Computacional.

LOGic). A construção de um *parser* da Linguística Computacional é o objetivo da parte prática desse livro.

O Prolog, “linguagem de programação baseada na lógica⁶”, foi desenvolvido visando ao trabalho de processamento de línguas naturais e ao desenvolvimento de *parsers* automáticos. O terceiro subcapítulo se dedica ao Prolog.

É interessante observar que, apesar de aparentemente complicado, o estudo da Linguística Computacional, tal como conduzido por Othero e Menuzzi, torna-se não só simples, mas instigante.

O terceiro capítulo, “Desenvolvendo um parser sintático I”, é dividido em duas partes: teoria lingüística e teoria computacional. Na primeira, são oferecidas algumas noções básicas de sintaxe gerativa, tais como os constituintes de uma sentença gramatical – nódulos não-terminais (sintagmas) e nódulos terminais (itens lexicais) – e a sua organização hierárquica (em árvores sintáticas ou por meio de colchetes rotulados). A segunda parte, de teoria computacional, explica como funciona a linguagem de programação Prolog e como a sintaxe é aplicada a essa linguagem.

“Desenvolvendo um parser sintático II” é o quarto capítulo, em que o leitor é convidado a “colocar a mão na massa”, como dizem os autores. O leitor é, então, introduzido ao “mundo da Linguística Computacional”⁷. O primeiro subcapítulo instrumentaliza o leitor a utilizar os conceitos da sintaxe, vistos no capítulo anterior, mostrando de que é formado cada sintagma, já que todas essas informações têm de ficar muito claras e explícitas no Prolog. Além disso, os autores discutem alguns problemas de ordem prática nas limitações do programa. O mesmo é feito com relação ao léxico no segundo subcapítulo.

O quinto e último capítulo intitula-se “Nossa Gramática” e destina-se a fazer com que o leitor entenda e utilize esse analisador gramatical. Os autores ensinam como implementar regras de concordância e de regência e incentivam o leitor a testar o programa e aventurar-se cada vez mais na seara da Linguística Computacional.

Ao término dessas páginas, os autores expressam as suas expectativas com relação ao leitor. São elas:

- (i) que tenha conseguido implementar uma primeira versão da gramática que apresentamos aqui e possa, a partir de agora,

⁶ Othero e Menuzzi, p. 42.

⁷ Othero e Menuzzi, p. 83.

- desenvolver novas e melhores versões dela, explorando cada vez mais o Prolog;
- (ii) que tenha conseguido acompanhar o ritmo do texto e possa partir em busca de novos conhecimentos na área da lingüística computacional com base no que aprendeu durante a leitura.

Dessa forma determinam sua relação de parceria também com o leitor que, nesse momento, já se sente um pouco autor e se mostra mais confiante para investir nessa área ainda tão nova para tantos lingüistas, professores, pesquisadores e outros interessados.