

Apresentação

Léxico, Texto e Tecnologias intitula uma das linhas de pesquisa do Programa de Pós-Graduação em Linguística Aplicada, da UNISINOS. São campos de investigação cujas temáticas, a despeito de suas identidades, teorias e propósitos próprios, alinham-se na mesma direção, tendo em vista a produtiva interface entre questões de léxico, de texto e de recursos tecnológicos. Dessa inter-relação decorrem inúmeras aplicações a distintos objetos de investigação, bem como se abrem novas perspectivas de diálogo com outras áreas afins, como ocorre com a lexicografia, a tradução e estudos textuais voltados à temática das tecnologias.

Este número da *Calidoscópio*, dedicado à linha de pesquisa em pauta, traduz muito dessas várias possibilidades que aqui estão representadas por pesquisas de léxico, semântica, texto e tecnologias informatizadas. Nesse elenco, agregam-se artigos que abordam desde questões de léxico terminológico, chegando à problemática do surgimento de novos gêneros textuais, influenciados pelas novas tecnologias informáticas. Todos os artigos expressam os desafios que a Linguística teórica, bem como a Linguística Aplicada passaram a enfrentar na época atual.

Maria Margarida Martins Salomão abre este número, apresentando o Projeto FrameNet Brasil, de pesquisa lexicográfica e vinculado ao Projeto FrameNet, liderado pelo linguista Charles Fillmore, do *International Computer Science Institute* (ICSI), em Berkeley, Califórnia. No artigo, a autora chama a atenção para o aporte teórico que fundamenta o projeto – a semântica de frames – e destaca também sua aplicação à lexicografia a partir do estudo de Fillmore e Atkins no princípio da década de noventa. É dada ênfase também às condições de implantação do projeto, abrangendo a escolha do corpus e a explicitação das categorias teóricas empregadas nesse empreendimento de pesquisa inédito no Brasil.

O artigo intitulado *O processamento automático de línguas naturais enquanto engenharia do conhecimento linguístico*, de Ariani Di Felippo e Bento Carlos Dias-da-Silva, aborda o domínio de pesquisa multidisciplinar denominado Processamento Automático de Línguas Naturais (PLN), que visa a capacitar um computador para lidar com a língua. Assumindo “uma concepção

linguisticamente motivada do PLN”, os autores, de forma didática, trazem à baila uma série de questões centrais concernentes à natureza heterogênea do PLN. Destacam-se, por um lado, a revisão do percurso histórico da área, assim como a reflexão sobre a relevância do trabalho colaborativo entre os cientistas da linguagem ou linguistas e os engenheiros da linguagem.

A complementariedade entre Linguística e Tecnologia também é tema do artigo intitulado *Compilação de corpus: representatividade e o CORPOBRAS*, de Lúcia Pacheco de Oliveira e Maria Carmelita Padua Dias. Vinculado à área da Linguística de Corpus, o artigo trata da compilação de um corpus de português do Brasil, tendo como objetivo principal na coleta dos dados a representatividade para estudos de gêneros discursivos. Tendo como foco principal esse parâmetro, as autoras discorrem sobre aspectos como extensão, objetivos e adequação aos interesses de pesquisa. Três questões são apontadas como centrais, quando o foco de atenção é a representatividade: *Do quê? Para quê? Para quem?*. No que concerne à compilação do CORPOBRAS, destacam que uma de suas características marcantes é apresentar uma ampla variedade de modalidades, de gêneros discursivos e de regiões, assuntos e autores. Outro ponto de destaque nesse artigo é a reflexão acerca das aplicações que podem se beneficiar de um corpus representativo do português do Brasil com gêneros do discurso oral e escrito. O estudo de gêneros do discurso pedagógico, profissional e espontâneo, assim como os estudos da variação e lexicografia são exemplos de aplicações práticas.

No artigo *Produtividade morfológica e tecnologia do texto: aspectos da construção de um transdutor lexical do português capaz de analisar neologismos*, Leonel Figueiredo de Alencar apresenta o LEXPOR, um protótipo de um componente morfológico do português capaz de segmentar e classificar os constituintes de derivados por meio da sufixação de *-ismo*, *-iano*, *-ês* e *-mente*, bem como de derivados desses por prefixação com elementos de origem grega ou latina do tipo de *neo-*, *pseudo-*, *anti-* ou *ultra-*. O protótipo consiste de um transdutor lexical que modela o conjunto de palavras que se podem potencialmente construir usando esses afixos derivacionais. Alencar

mostra que o LEXPOR, capaz de ser aplicado na anotação automática de corpora, pode realizar análises plausíveis para um número potencialmente infinito de palavras completamente novas, como as que se encontram em corpora eletrônicos abrangentes como o CETENFolha e em textos disponíveis na Internet.

Mônica Maria Pereira da Silva e Regina Celi Mendes Pereira, em *A linguagem dos manuais de aparelho celular: desafios de leitura a serviço da tecnologia*, apresentam os resultados de uma pesquisa desenvolvida com diferentes usuários de telefone celular, focalizando os efeitos e impactos da leitura do manual sobre esses leitores e o consequente uso do aparelho. No estudo dos manuais, as autoras destacam a utilização de manifestações multimodais, que visam a exemplificar ou a produzir uma melhor compreensão de uma determinada instrução de uso. Com base em parâmetros de análise sociointeracionistas, Mônica e Regina verificam, em sua pesquisa, que a conjugação do texto escrito com a imagem minimiza as dificuldades de leitura que remetem a interpretações equivocadas.

Reunindo também aspectos de tecnologia e de interesse de profissionais, no caso, tradutores, Adriane Orenha em seu artigo, *A compilação de corpora comparáveis na área de negócios e sua relevância para a tradução e terminologia*, ressalta a importância da Linguística de Corpus para a área da Tradução. Enfoca, em particular, a compilação de corpora comparáveis como instrumentos essenciais para o processo de tradução. Trata-se de corpora que são denominados de comparáveis, porque possuem algumas características em comum. A autora descreve a constituição de alguns corpora que relacionam textos em português e inglês na área de Negócios, cujo objetivo maior é disponibilizar para tradutores um material de pesquisa que retrate a linguagem em uso. Salienta, conclusivamente, que esse tipo de corpora é útil também para a identificação da terminologia técnico-científica das áreas contempladas pelos textos compilados.

No artigo *Terminologia a serviço da informação: rede de palavras-chave para artigos de divulgação científica da Medicina*, Márcio Sales Santiago e Maria da Graça Krieger observam a co-ocorrência entre termos técnico-científicos e denominações de caráter popular presentes em artigos de divulgação científica da Medicina, publicados em um *site* destinado ao público leigo, cujo objetivo é a informação, divulgação e educação sobre temas relacionados à saúde. Tal fato

motivou a proposição de bases teórico-metodológicas para a elaboração de mecanismos de busca, recuperação e aumento de informações especializadas, os quais se denominaram redes de palavras-chave. Essas, por sua vez, têm como principal objetivo auxiliar o leitor dos artigos de divulgação a entender melhor os termos e conceitos advindos do discurso médico. Com essa proposição, a área de Terminologia dialoga com a de Documentação, em especial, no plano dos princípios de elaboração da linguagem documentária, objeto precípua da organização da informação sobre ciências, técnicas e tecnologias.

Fecha este número o texto *Famílias de gêneros e novas formas comunicativas para a ciência*, de Guiomar Ciapuscio. Como uma espécie de afirmação das relações entre tecnologia e linguagem, a autora invoca a necessidade de que a Linguística encontre princípios para responder a questões basilares relacionadas a práticas comunicativas do atual mundo tecnológico. Propõe, de maneira inovadora, o conceito de famílias de gêneros, ilustrando seu pensamento com a realidade da ciência que, hoje, é divulgada e validada através de muitas práticas comunicativas, estruturadas e divulgadas com apoio de recursos informatizados a exemplo de *facebooks*, *blogs* ou *microblogging*, *e-books* e *interação comunicativa*. Entende que tais práticas têm vínculos e características comuns, típicas de um mesmo gênero textual, o que justifica seu postulado sobre famílias de gêneros. Inscreve seu trabalho no âmbito da Linguística dos Gêneros, de tradição alemã, mas, diferente de uma concepção estática de tipos de gênero, adota uma visão dinâmica das tipologias textuais relacionadas através do conceito de família genérica. Esta, por sua vez, é fundamentada na ideia de que cada âmbito de atividade desenvolve suas próprias formas genéricas como estratégias de solução pré-formadas para as tarefas próprias a esse âmbito. Guiomar Ciapuscio considera ainda que a noção de família permite melhor entender o surgimento de novos “membros” da família genérica, que respondem às mudanças socioculturais nos modos de entender a relação entre a ciência e a cidadania, bem como às possibilidades brindadas pelas novas tecnologias.

Maria da Graça Krieger
Maria Eduarda Giering
Rove Chishman