

# Idéia ou ideia?

Ferramentas compatíveis com as versões atuais do pacote Office são licenciadas pela Techno para a Microsoft



TEXTO **Marili Ribeiro** ILUSTRAÇÃO **Guilherme Lepca**

A última grande mexida se deu em 2009. Foi por conta da introdução da reforma ortográfica que, por exemplo, deixou a escrita da língua portuguesa sem o trema. Mas isso já foi feito por uma empresa privada, longe do *campus* universitário. Consolidado e bem-sucedido, o *Revisor Gramatical Automático para Português* – desenvolvido em parceria entre o Núcleo Interinstitucional de Linguística Computacional da Universidade de São Paulo (USP), do *campus* de São Carlos, e a Itaotec-Philco S.A., com apoio da FAPESP – hoje é operado pela Techno Software.

“O convênio com a Itaotec terminou em 2008 e, desde então, não temos investido no revisor”, conta Maria das Graças Volpe Nunes, que foi a coordenadora do projeto desde o seu início, em 1993, e que também é professora do Departamento de Ciências da Computação e Estatística da USP em São Carlos. “Como o revisor já estava estabilizado, dentro de um ótimo padrão de funcionamento, após os ajustes da revisão ortográfica, feitos pela equipe da empresa Techno Software (parceira da Itaotec), acredito que o produto não evoluiu mais”, completa Maria das Graças.

Já o grupo de pesquisa do Núcleo da USP, que participou ativamente do projeto, segue investindo em vários estudos ligados ao processamento computacional da língua portuguesa, como informa a professora. “O projeto do Revisor Gramatical foi, sem dúvida, precursor de todo o desenvolvimento dessa área no Brasil.”

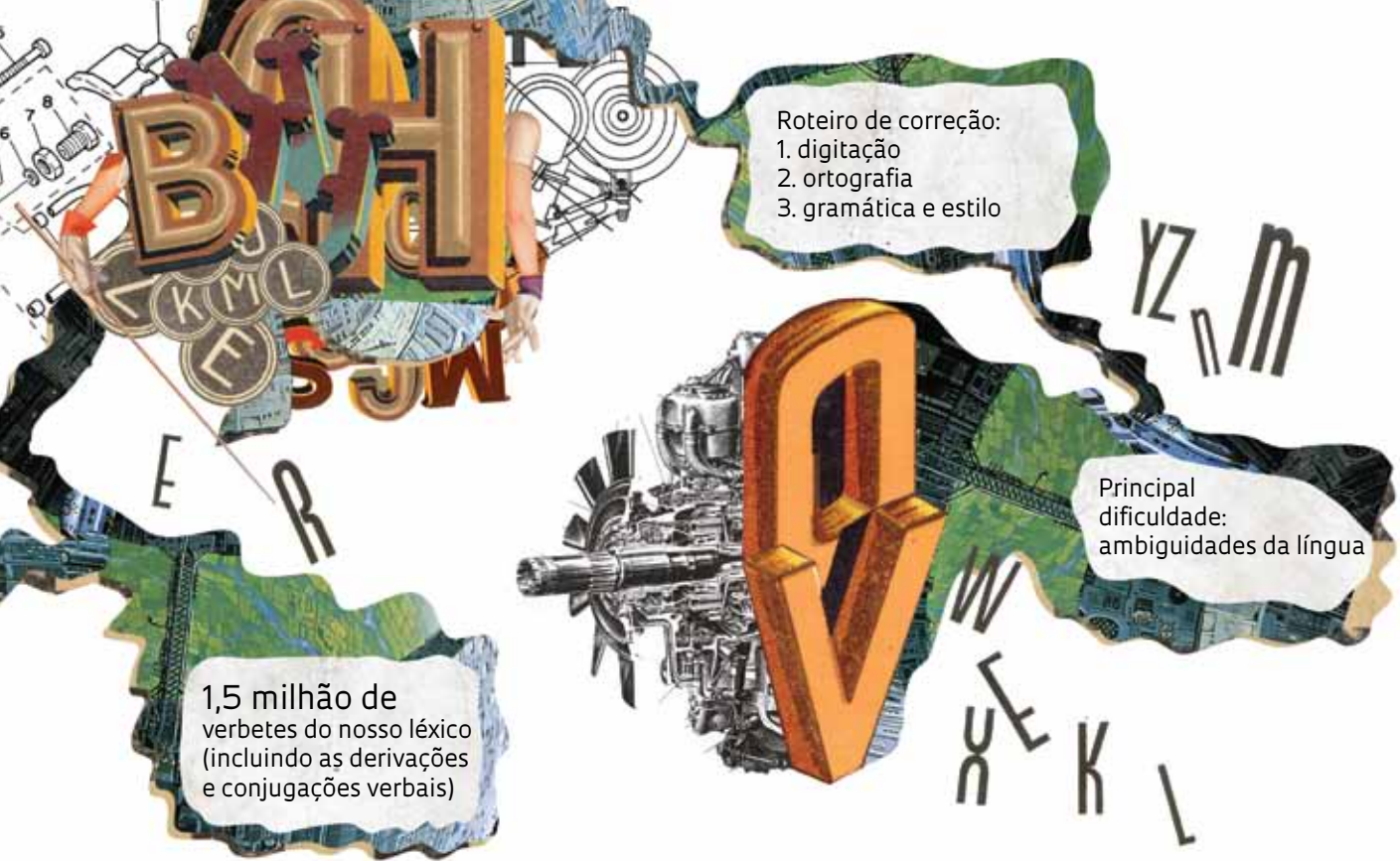
As ferramentas de revisão (ortográfica, gramatical, hifenizador, dicionário de sinônimos) são de propriedade da Itaotec e foram trabalhadas em conjunto com a USP. Após o término do convênio com a USP, a Techno Software assumiu todos os desdobramentos do produto a pedido da Itaotec. “As últimas implementações foram para adequá-las à reforma ortográfica, e isso foi feito no final de 2009. A partir de então, temos dado manutenção nas ferramentas e complementando e corrigindo o léxico quando necessário”, informa, por e-mail, Carlos Henrique Ferreira, da Techno Software. “Como estas ferramentas são licenciadas para a Microsoft, também fazemos adequações para compatibilizar com as versões atuais do pacote Office”, acrescenta. O *software* produzido e comercializado pela Itaotec teve seus direitos adquiridos pela Microsoft, que o incorporou ao programa Office 2000.

## O PRINCÍPIO

O *software* de revisão gramatical para a língua portuguesa nasceu como parte dos projetos de pesquisas em que a FAPESP financia de 20% a 70% da iniciativa, no âmbito do Programa de Apoio à Pesquisa em Parceria para Inovação Tecnológica (Pite), sempre com a contrapartida da empresa interessada. Quanto maior o risco do projeto, maior é a parte da FAPESP. Essa linha de investimento é consequência da crescente consciência de que o ritmo de inovação tecnológica mundial é tão acelerado que o Brasil precisa alavancar sua capacidade para atuar de forma equivalente aos países desenvolvidos. Não há dúvida de que quem não tiver tal capacidade de inovar fica marginalizado.

Em sintonia com essa demanda, empresas brasileiras experimentam os efeitos de parcerias com universidades e institutos de pesquisa. Novos produtos e processos de produção industrial vêm surgindo dessa convivência e têm garantido retorno financeiro às que apostam nessa interação. Um dos caminhos de impulso à arrancada do processo, que cresce em diferentes setores econômicos, se dá graças ao apoio do PITE.

O projeto do revisor gramatical foi aprovado no PITE em 1996. Nesse mo-



mento, ganhou impulso. Antes disso, a proposta de um revisor estava presente nas pesquisas realizadas desde 1993 através de convênio entre a Itautec-Philco e a Fundação de Apoio à Física e à Química de São Carlos. O trabalho era executado por uma equipe multidisciplinar de linguistas e profissionais da área de computação, com a participação de docentes do Departamento de Ciência da Computação e Estatística, do Instituto de Ciências Matemáticas e de Computação e do Instituto de Física de São Carlos, sempre coordenado pela professora Maria das Graças.

A entrada da FAPESP facilitou a ampliação do escopo da pesquisa, que passou a contar com a colaboração dos professores Claudio Lucchesi, Tomas Kowaltowski e Jorge Stolfi, do Instituto de Computação da Unicamp. Em São Carlos, sob a coordenação de Maria das Graças, foram desenhados os algoritmos e formado o banco de base de palavras e, em Campinas, desenvolveram-se a compactação do sistema e a diminuição do tempo de resposta do programa.

#### O PRODUTO

Em caixas próprias, a Itautec começou a vender a primeira versão do revisor gramatical em 1997. O produto nas prateleiras do varejo atraiu atenção e, já no final desse mesmo ano, ganhou o interesse da

gigante Microsoft, que procurou a empresa para incorporar o revisor no programa Office, o mais vendido no Brasil e em todo o mundo. Para a língua portuguesa falada no Brasil, a multinacional americana usava o antigo revisor criado em Portugal, que comportava 200 mil palavras. O da Itautec já dispunha de 1,5 milhão de palavras. O revisor foi incorporado ao Office 2000, com a Itautec licenciando o produto por um período de três anos pelo valor de US\$ 421 mil. Novos acertos foram feitos depois na renovação de uso.

No desenvolvimento do produto, a Itautec gastou R\$ 78 mil, enquanto a FAPESP investiu R\$ 17,9 mil, além de US\$ 9,2 mil, utilizados na compra de máquinas e equipamentos para a USP. Entre os acadêmicos envolvidos com o projeto não havia qualquer projeção sobre a dimensão que ganharia, principalmente com o contrato com a Microsoft.

O revisor detecta um grande número de erros comuns cometidos por usuários de nível de segundo grau. É capaz de indicar e sugerir alternativas para erros ortográficos, mecânicos (colocação irregular ou falta de pontuação no final de frases, por exemplo) e erros gramaticais relativos ao uso da crase, regência, concordância verbal e nominal, colocação pronominal e inadequações lexicais e outros.

O software não interfere no estilo, não padroniza a estrutura do texto, mas percorre sentença por sentença para verificar a estrutura sintática e oferecer opções gramaticalmente corretas de construção. O programa arrisca a função pela proximidade de um verbo ou adjetivo e sugere correções para as diferentes estruturas sintáticas decorrentes de cada um dos diferentes significados. ■

#### O PROJETO

Projeto e implementação de um revisor gramatical automático para o português – nº 1997/02608-1 (1997-1998)

#### COORDENADORA

Maria das Graças Volpe Nunes – Instituto de Ciências Matemáticas e de Computação, USP

#### MODALIDADE

Programa de Apoio à Pesquisa em Parceria para Inovação Tecnológica (Pite)

#### INVESTIMENTO

R\$ 17.900,00 e US\$ 9.200,00 (FAPESP) e R\$ 78.000,00 (Itautec Philco S.A.)

#### DE NOSSO ARQUIVO

*Benefícios de uma parceria*  
Edição nº 58 – outubro de 2000

*Os novos rumos da pesquisa tecnológica*  
Edição nº 47 – outubro de 1999

*Revisor Gramatical Automático em nova versão*  
Edição nº 35 – setembro de 1998