

UM CORPUS DE BLOGS: UM CORPUS DE UM NOVO GÊNERO?

CLÁUDIA FREITAS (LINGUATECA/PUC-RIO)

Tradicionalmente, gênero textual é entendido como uma maneira de realizar objetivos linguisticamente. O foco da definição de gênero está, principalmente, na sua função, nos objetivos pretendidos com uma determinada ação verbal, e a tipicidade de um gênero vem com suas características funcionais e organização retórica. No entanto, a dimensão técnica - o suporte -, embora raramente explicitada, também pode ser tomada como critério na descrição dos gêneros textuais “Os gêneros textuais (...) definidos por composições funcionais, objetivos enunciativos e estilos concretamente realizados na integração de forças históricas, sociais, institucionais e técnicas”, o que leva à identificação de diferentes gêneros, como telefonema, carta comercial, carta pessoal, romance, bilhete, receita culinária, carta eletrônica etc. Ao assumir que um bilhete pode ser “transformado” em recado caso seja veiculado em um papel ou deixado em uma secretária eletrônica, isto é, a partir do momento em que se postula que o meio é capaz de interferir de maneira tão direta na classificação de gêneros, as distinções e os próprios gêneros, até então claros e estáveis, começam a ser alvo de discussão. Nesse contexto aparecem os blogs: são um gênero novo, um gênero híbrido, apenas um suporte? Se, inicialmente, blogs surgiram com o propósito de relatar experiências pessoais, em uma atualização dos diários pessoais, cada vez mais eles exercem diferentes funções: para jornalistas, podem servir como espaços alternativos de publicação; para empresas (mas não somente), mais um espaço de divulgação; para boa parte dos blogueiros, um espaço de interação, expressão pessoal e compartilhamento de ideias. Tratam dos conteúdos mais diversos, podem ser escritos nos mais diversos registros, e compreender textos em gêneros variados - de ensaios a receitas culinárias, de entrevistas a relatos de viagens, de eventos e até mesmo de experiências sexuais. Podem ser de um único autor ou escritos coletivamente, e seus escritores têm uma formação diversa, o que contribui para a diversidade dos estilos. Há, no entanto, pouca investigação até o momento sobre como de fato são ou quais as funções mais comuns dos blogs. Além disso, para sistemas que lidam com o processamento computacional da língua, blogs são uma rica fonte para a área de análise de subjetividade/ de opinião e de mineração de textos. Acreditamos que toda a discussão seja suficiente para ilustrar a relevância de um corpus de blogs. E, embora seja crescente o interesse na compilação de novos corpora para a língua portuguesa, notamos que, em geral, há pouca variação quanto ao gênero, com predominância de corpora compostos por textos de jornal. Com relação a textos criados especificamente para ambientes digitais temos conhecimento apenas do Corpus ANCIB, que corresponde a mensagens de correio eletrônico na lista ANCIB, e da Coleção Dourada do Segundo HAREM que inclui, dentre outros tipos de textos, blogs e textos da Wikipédia. Nesse contexto, apresentamos a Amazônia, um corpus publicamente disponível criado pela [omitido] no âmbito do projeto [omitido], que contém 4.6 milhões de palavras da variante brasileira do Português, analisadas morfossintaticamente pelo analisador PALAVRAS. A Amazônia contém os textos da seção OverBlog do site Overmundo, disponíveis em 30 de Setembro de 2008, perfazendo um total de

4070 textos. O Overmundo é um site colaborativo - qualquer um pode escrever - voltado para a cultura brasileira. Como qualquer um pode colaborar, há textos de todas as regiões do Brasil, e uma menor presença (teoricamente) do jargão jornalístico, o que também o diferencia dos demais corpus. A seção OverBlog, especificamente, contém “reportagens, entrevistas e críticas sobre cultura do Brasil”. O objetivo deste trabalho, portanto, é duplo: (i) contribuir para uma descrição deste tipo de texto/gênero eletrônico, que, devido à crescente produção, facilidade de obtenção e liberdade de escrita, é uma amostra rica da língua; (ii) descrever o tipo de texto do corpus Amazônia. Nesse ponto, se impõe a seguinte questão: até que ponto a descrição do material compilado não é exatamente um dos resultados esperados a partir de sua análise? Em uma análise preliminar, quanto ao gênero, podemos caracterizar os textos da Amazônia/Overblog em entrevistas, resenhas, reportagens, artigos e narrações/relatos de eventos. Por isso, em termos de gênero, usar blog pode não fazer sentido, visto que todo o conteúdo parece poder ser enquadrado em um gênero tradicional. No entanto, há nítidas influências do suporte no texto - em termos não apenas de informalidade, mas principalmente de hibridismo entre diferentes tipos de textos e de quantidade de interação, características que são mais evidenciadas se tratarmos dos artigos e das resenhas. Gêneros textuais são fenômenos sociais e históricos. Quanto aos blogs, especificamente, questionar o seu enquadramento em termos de um novo gênero não significa questionar a influência do suporte nos gêneros textuais. Enquanto objetos históricos, é natural que os gêneros sofram mudanças/ ampliação - o que se observa, por exemplo, nos bate-papos e páginas pessoais. No entanto, nem todo texto apoiado em um novo suporte talvez mereça ser designado um novo gênero. Leis já foram escritas em paredes, em tábuas, pergaminhos, papel e editores de texto, e nem por isso deixam de ser leis. Com o reconhecimento de que não se trata de um novo gênero, mas de gêneros tradicionais com diferenças decorrentes das características do suporte, abre-se espaço para uma descrição comparativa assentada nas diferenças que podem, inclusive, ser tão gritantes que forcem de fato a inclusão de blogs como um novo gênero. Mas tal só poderá ser feito se partirmos do muito de comum - em termos de movimento de texto, de função, de objetivos - que existe entre os diversos textos abarcados/veiculados pelo blog. A disponibilização de um corpus com tais características é o primeiro passo em direção a uma descrição mais sistemática desses “novos” textos.