

Amazônia: um corpus de blogs?

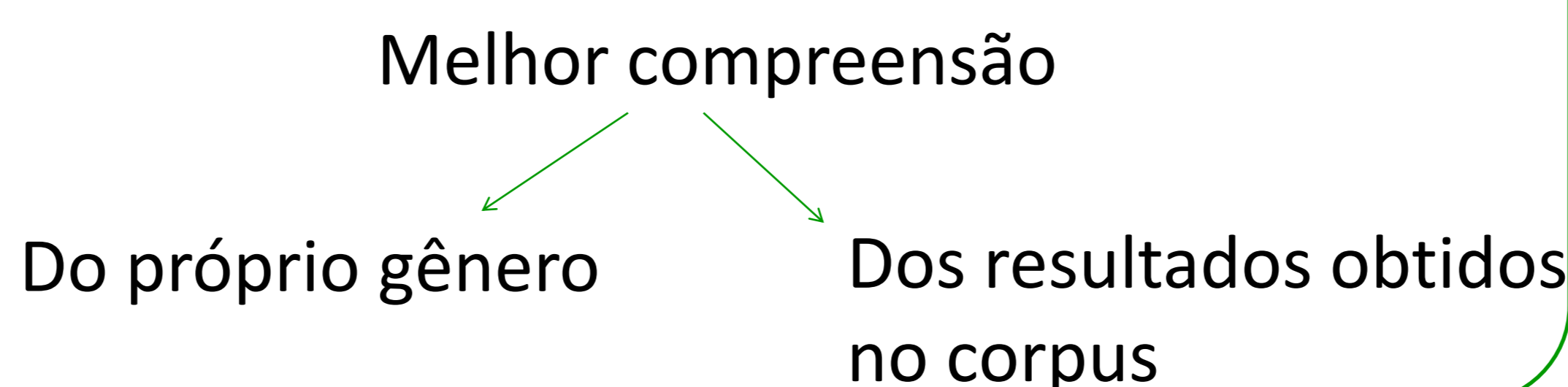
Cláudia Freitas (Linguatca-FCCN / PUC-Rio)

Motivação

Interesse crescente na descrição de blogs

- Descrição de práticas linguísticas/sociais /culturais (Herring et al 2005; Miller & Shepherd 2004; Blood, 2000)
- Detecção de opinião (Andreevskaia et al., 2007; Leshed &Kaye, 2006)

Caracterização do corpus quanto ao gênero



Gênero

Multiplicidade de perspectivas
Critérios *funcionais*, linguísticos , contextuais

- Forma de ação social (Miller, 1984);
 - maneira de realizar objetivos linguísticamente.
 - Estrutura textual

O que há na Amazônia

"Lacuna esta que, ao longo dos últimos anos, buscou-se suprimir com um conjunto de pesquisas acadêmicas, nos cursos de História, Letras e Comunicação, abordando desde a história dos festivais de música realizados na cidade (como o histórico Massafeira, realizado em 1979 com a participação de vários artistas locais, materializado em disco duplo) à riqueza lingüística das composições."

"Recentemente, o IBAMA, através do CGPEG (Coordenação Geral de Petróleo e Gás), tem intensificado a pressão para que as empresas implementem projetos ambientais consistentes na região, incluindo os de Educação Ambiental, historicamente relegados à última prioridade."

"E a gente não tem hit nenhum em rádio, haha, neguinho não vai nem conhecer"

"Podendo ajudar mó galera a perturbar o status quo, vou ficar feliz pra caralho!!!"

"acho um grande passo pelo menos trocarmos pra 'detesto', que é menos pesado que 'odeio', será que conseguimos? ;)"

"Coloque sua cidade pra aparecer!:D Seja a mídia que você quer ver no mundo! Hehehe"

"E, acreditem, essa galera ainda vai dar o que falar ... : D Vê o blog lá:"

"Hummm, não sei... ;)"

Blogs

Origem: diários pessoais

Subjetividade

Veiculação: Internet

Rapidez

Interatividade

Subjetividade

Função

Espaço alternativo de publicação

Espaço de divulgação

Espaço de interação, expressão pessoal e compartilhamento de ideias

Conteúdo

Receitas

Relatos (de viagem, de eventos, de experiências sexuais etc)

Entrevistas

Resenhas

Convites

Autoria

um autor

vários autores

Gêneros "tradicionais" ?

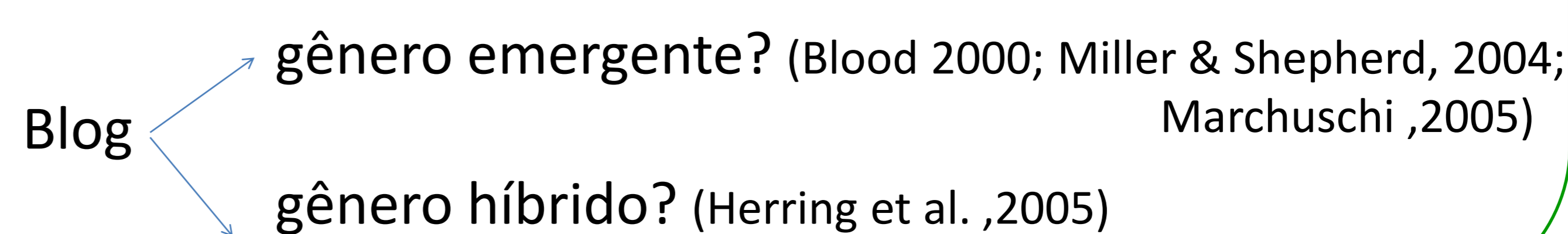
escritas digitais:
+ liberdade

+ heterogeneidade
quanto aos
gêneros /
intenções /funções

O que constitui um gênero novo?

- ✓ Nascido por novas necessidades / funções (em novo suporte) -- homepage
- ? Repetição de necessidades / funções em (em novo suporte) – e-mail?

Qual o peso do suporte?



A Amazônia

- textos da seção OverBlog, do site Overmundo, disponíveis em 30/09/08.
- Overmundo: site colaborativo voltado para a cultura, "em especial as práticas, manifestações e a produção cultural que não têm a devida expressão nos meios de comunicação tradicionais."
- OverBlog: "reportagens, entrevistas e críticas sobre cultura do Brasil".

AMAZÔNIA

Parte da Floresta Sintá(c)tica

(Afonso et al., 2002; Freitas et al., 2008)

Acessível pelo AC/DC (Costa et al., 2009)

Anotada pelo PALAVRAS (Bick, 2000)

Palavras 4.6 milhões

textos 4070

Publicamente disponível a partir de [www. Linguatca.pt/Floresta](http://www.Linguatca.pt/Floresta)

Breve exploração da Amazônia no AC/DC

	Amazônia (blog)	Corpo CHAVE (jornal impresso)
Bem,	0,003%	0,0009%
Bom,	0,003%	0,0007%
né	0,006%	0,0007% (apenas BR)
você/voce/vc	0,09%	0,04% (apenas BR)
verbos na 1a pes. sing.	7%	1%

CAMINHOS FUTUROS

Como caracterizar a Amazônia:

Blogs?

⊗ perde-se a especificidade (já descrita) dos gêneros tradicionais

😊 acrescentam-se as dimensões de liberdade (quanto ao tipo de texto), interação, subjetividade

? A descrição do material é um dos resultados esperados da sua análise?

Pesquisadores da Floresta Sintá(c)tica responsáveis pela criação da Amazônia

Cláudia Freitas
Eckhard Bick
Diana Santos

Referências

Susana Afonso, Eckhard Bick, Renato Haber & Diana Santos. "Floresta sintá(c)tica: um treebank para o português". In Anabela Gonçalves & Clara Nunes Correia (eds.), *Actas do XVII Encontro Nacional da Associação Portuguesa de Linguística (APL 2001)*, Lisboa, Portugal : APL, pp. 533-545. 2002.

Alina Andreevskaia, Sabine Bergler, e Monica Urseanu. "All blogs are not made equal: Exploring genre differences in sentiment tagging of blogs". In *International Conference on Weblogs and Social Media*. Hilton Seattle Downtown, Seattle, Washington, U.S.A. 2007

Eckhard Bick. *The Parsing System "Palavras": Automatic Grammatical Analysis of Portuguese in a Constraint Grammar Framework*. Dr.phil. thesis. Aarhus University. Aarhus, Denmark: Aarhus University Press. Novembro de 2000.

Rebecca Blood. "Weblogs: A History and Perspective" *Rebecca's Pocket, Sept. 7, 2000*; disponível em www.rebeccablood.net/essays/weblog_history.html.

Guimar Elena Ciapuscio. "Famílias de gêneros e novas formas comunicativas para a ciência" *Calidoscópio* Vol. 7, n. 3, p. 243-252, set/dez 2009.

Luís Costa, Diana Santos & Paulo Alexandre Rocha. "Estudando o português tal como é usado: o serviço AC/DC". In *The 7th Brazilian Symposium in Information and Human Language Technology (STIL 2009)* (São Carlos, SP, Brasil, 8-11 de setembro 2009).

Cláudia Freitas, Paulo Rocha & Eckhard Bick. "Um mundo novo na Floresta Sintá(c)tica - o treebank para Português". *Calidoscópio*, vol. 6.3, pp. 142-148, 2008.

Susan C. Herring, Louis A. Scheidt, Sabrina Bonus. "Weblogs as a bridging genre". *Information Technology & People*; vol 18, n. 2, 2005. pp 142-171.

Gilly Leshed & Joseph Kaye. Understanding How Bloggers Feel: Recognizing Affect in Blog Posts. In Proc. of CHI-2006, Montreal, Canada.

Luiz Antônio Marcusch. "Gêneros textuais emergentes no contexto da tecnologia digital". In Luiz Antônio Marcusch e Antonio Carlos Xavier (Eds.), *Hipertexto e Gêneros Digitais*. Rio de Janeiro: Editora Lucerna, 2005.

Carolyn R. Miller, & Dawn Sheperd. "Blogging as social action: A genre analysis of the weblog". In Laura Gurak, Smiljana Antonijevic, Laurie Johnson, Clancy Ratliff & Jessica Reyman (Eds.), *Into the Blogosphere. Rhetoric, Community, and Culture of Weblogs*. 2004.

Carolyn R. Miller. *Genre as a social action*. Quarterly Journal of Speech, Vol 70, pp.151-167. 1984

A Amazônia e o projeto Floresta Sintá(c)tica são desenvolvidos no âmbito da Linguatca, co-financiada pelo governo português, pela União Europeia (FEDER e FSE), sob o contrato POSC/339/1.3/C/NAC, pela UMIC e pela FCCN.