



Universidade do Estado do Rio de Janeiro

Centro de Educação e Humanidades

Instituto de Letras

Marco Antônio Macedo Gutierrez

Navegantes e surfistas: as curvaturas do ciberespaço nas vizinhanças dos portais

Rio de Janeiro

2007

Marco Antônio Macedo Gutierrez

Navegantes e surfistas: as curvaturas do ciberespaço nas vizinhanças dos portais



Dissertação apresentada, como requisito parcial para obtenção do título de Mestre, ao Programa de Pós-Graduação em Letras da Universidade do Estado do Rio de Janeiro. Área de concentração: Linguística.

Orientadora: Prof^ª. Dra. Anna Elizabeth Balocco

Rio de Janeiro

2007

CATALOGAÇÃO NA FONTE
UERJ/REDE SIRIUS/CEHB

G984 Gutierrez, Marco Antônio Macedo.
Navegantes e surfistas: as curvaturas do ciberespaço nas
vizinhanças dos portais / Marco Antônio Macedo Gutierrez. – 2007.
208 f. : il.

Orientador : Anna Elizabeth Balocco.
Dissertação (mestrado) – Universidade do Estado do Rio de
Janeiro, Instituto de Letras.

1. Análise do discurso – Teses. 2. Ciberespaço - Teses. 3.
Internet (Redes de computação) – Teses. 4. Sistemas de hipertexto –
Teses. I. Balocco, Anna Elizabeth. II. Universidade do Estado do Rio
de Janeiro. Instituto de Letras. III. Título.

CDU 82.085

Autorizo, apenas para fins acadêmicos e científicos, a reprodução total ou parcial desta
dissertação desde que citada a fonte

Assinatura

Data

Marco Antônio Macedo Gutierrez

Navegantes e surfistas: as curvaturas do ciberespaço nas vizinhanças dos portais

Dissertação apresentada, como requisito parcial para obtenção do título de Mestre, ao Programa de Pós-Graduação em Letras da Universidade do Estado do Rio de Janeiro. Área de concentração: Linguística.

Aprovada em 29 de março de 2007.

Banca Examinadora:

Prof^a. Dra. Anna Elizabeth Balocco (Orientadora)
Instituto de Letras da UERJ

Prof. Dr. David Shepherd
Universidade Federal Fluminense

Prof^a. Dra. Gisele Carvalho
Instituto de Letras da UERJ

Rio de Janeiro

2007

*Two roads diverged in a wood, and I –
I took the one less traveled by,
And that has made all the difference.
Robert Frost*

A Marta, Maurício e Marcel. Que nossos
caminhos – diferentes – se cruzem pelas
milhas que ainda faltam antes do sono.

AGRADECIMENTOS

Esta dissertação sequer teria sido concluída sem os esforços da minha orientadora, Dr^a Anna Elizabeth Balocco, que encontrou tempo em meio às suas inúmeras tarefas acadêmicas para enfiar algumas boas idéias nessa minha cabeça dura. Seus seminários e encontros foram preciosos. Encantou-me sobretudo seu permanente exercício de tolerância intelectual para com a diferença, cada vez mais rara no nosso mundo acadêmico. Meus sinceros agradecimentos.

Os demais professores do curso foram também essenciais à pesquisa. Em especial a Dr^a Gisele de Carvalho e a Dr^a Tânia Shepherd, que aos seminários que me permitiram elaborar e amadurecer inúmeras idéias aqui apresentadas acrescentaram a honra de compor a banca examinadora. Também a Dr^a Tânia Saliés e a Dr^a Maria Teresa Abreu, em cujos cursos algumas das seções da dissertação foram elaboradas. E a Dr^a Zinda Vasconcellos, pelas críticas e sugestões feitas em várias jornadas e congressos, onde testei algumas das minhas idéias.

Agradeço igualmente ao Dr. David Shepherd e à Dr^a Solange Vereza pela honra concedida em compor a banca examinadora.

RESUMO

GUTIERREZ, Marco Antônio. *Navegantes e surfistas: as curvaturas do ciberespaço nas vizinhanças dos portais*. 2007. 208 f., il. Dissertação (Mestrado em Linguística) – Instituto de Letras, Universidade do Estado do Rio de Janeiro. Rio de Janeiro, 2007.

De um “território livre”, construído em torno dos princípios da liberdade e da colaboração – um ambiente onde pode ser exercida a liberdade de ser (ou parecer), de dizer (ou escrever) e de ouvir (ou ler), para o bem ou para o mal, a massificação do uso da Internet tem constituído um espaço social (virtual) contraditório e diversificado, onde forças cuja vocação parece ser homogeneizar (e possivelmente pasteurizar) esse “território livre da diferença” têm sido postas em cena. Dentre essas forças, lugar especial é ocupado pelo que se convencionou chamar “portais de interesse geral”. Tais *sites*, ao procurar responder a diferentes interesses, de modo a assegurar que o máximo possível de anseios de seus usuários seja satisfeito sem a necessidade de navegarem para fora dos seus domínios, são construídos como empreendimentos cujo sucesso requer o *controle* do comportamento de navegação dos usuários – o que, neste caso, não pode ser feito senão através do discurso (persuasivo). A hipótese que salta aos olhos é evidente: e se este não for apenas um controle *pelo* discurso, mas um controle *do* discurso? A premissa é, evidentemente, que tais *sites* constituem uma categoria especial de artefatos de linguagem produzidos no interior de eventos comunicativos bem determinados, isto é, um gênero emergente do discurso. Assim, o objetivo primário da presente pesquisa consistiu em determinar como o modo de funcionamento discursivo do portal contribui para afetar os hábitos de navegação dos seus usuários, curvando e distorcendo o ciberespaço nas suas vizinhanças. Para isso procuramos abordar o que chamamos *atratores retóricos*, isto é, aquelas estratégias (discursivas) destinadas a seduzir, convencer ou mesmo fraudar os usuários de modo a assegurar sua permanência na órbita dos portais, fazendo desses *websites* vetores da *ordem do discurso*, um elemento de controle do caos típico do ciberespaço, um agente de exorcismo da diferença e subjugação do discurso desenfreado.

Palavras-chave: Gêneros do discurso. Hipertexto. Internet. Portais de interesse geral.

ABSTRACT

The starting point for this research is a “free-space” built around principles of freedom and cooperation, an environment where may be exerted freedom of being (or pretend), of saying (or writting) and of listening (or reading). The massive growth in use of Internet has built, for better or worse, a social (virtual) contradictorial and varied space. Here certain strenghts have been at work, whose purpose seems to be homogenise (or possibly pasteurise) a “land-free-of-difference”. Among these, a special place is occupied by what has been called “general interests portals”. The latter websites seek to provide answers to several desires. Their aim is to ensure that the most probable wishes of it’s users are satisfied inside it’s domains. Thus these websites are built as enterprises, whose success depends upon the user’s navigational behavior control. In this particular case this cannot be achieved except by (persuasive) discourse. My research hipothesis might thus have easily seen: what if this is not just a control done *through* discourse but also a control *of* discourse? What we assume here is, evidently, that such websites form a special category of language artifacts created inside determined communicative events, e. g., an emerging discursive genre. Thus the primary objective of the present research was to determine whether the portal’s discourse functioning mode contributed in altering the navigation habits of its users, bending and distorting cyberspace in its surroundings. Therefore, an effort was made to focus on what have been defined as rhetoric atractors. These are those (discursive) strategies meant to seduce, convince or even deceive the users in order to assure their permancece in the portal’s orbit. These strategies, in turn, make the same websites carriers of the order of discourse, an element that controls the typical chaos of the cyberspace, an agent of exorcism of difference and domination of the un restrained discourse.

Keywords: Genres of discourse. Hipertext. Internet. General interests portals.

LISTA DE TABELAS

Tabela 1: Tempo de exposição relativo dos usuários da Internet	17
Tabela 2: Atividades realizadas na Internet	41
Tabela 3: Eventos comunicativos de interesse no ciberespaço.	99
Tabela 4: Distribuição de eventos comunicativos num ciberespaço bidimensional.	100
Tabela 5: Distribuição dos eventos comunicativos num ciberespaço tridimensional.	102
Tabela 6: Distribuição dos eventos comunicativos no ciberespaço multidimensional	104
Tabela 7: Os parâmetros de localização no ciberespaço das ocorrências do corpus.	108
Tabela 8: Incidência média dos parâmetros considerados na amostra.	109
Tabela 9: Resultado de exemplo da análise das propriedades de G.	122
Tabela 10: Portais ativos listados no diretório Cadê.	125
Tabela 11: Portais ativos obtidos em consulta aos buscadores Google e Yahoo.	126
Tabela 12: Portais adicionais obtidos no diretório Acesso Gratuito do Cadê.	126
Tabela 13: Distribuição de entradas por categoria no Cadê.	127
Tabela 14: Sites utilizados como grupo de controle.	128
Tabela 15: Expressões mais utilizadas nas buscas do Exploora em julho de 2004.	130
Tabela 16: Primeira hipótese de propriedades temáticas dos portais.	131
Tabela 17: Cálculo do valor lingüístico de verdade.	134
Tabela 18: Relevância das propriedades arbitradas no corpus.	135
Tabela 19: Valor lingüístico de verdade da relevância das propriedades.	136
Tabela 20: Prototipicidade das instâncias da amostra.	139
Tabela 21: Prototipicidade do grupo de controle.	139
Tabela 22: Quantidade total de documentos devolvidos em cada experimento.	150
Tabela 23: Taxa de incidência de saídas diferentes por dispositivo.	150
Tabela 24: Resultado da primeira busca do experimento Copérnico.	157
Tabela 25: Resultado de uma busca captando audiência independente.	158
Tabela 26: Resultado de busca com realimentação positiva.	159
Tabela 27: Corpus do experimento com os dispositivos de busca dos portais.	161
Tabela 28: Expressões de busca utilizadas no experimento.	162
Tabela 29: Taxa de redundância das saídas em comparação com o Yahoo.	163
Tabela 30: Taxa de redundância das saídas em comparação com o Google.	163
Tabela 31: Taxa de fechamento dos portais do grupo Yahoo.	163
Tabela 32: Taxa de fechamento dos portais do grupo Google.	164
Tabela 33: Documentos inspecionados na observação.	168
Tabela 34: Média de links externos por página em cada nível da amostra.	168
Tabela 35: Resultados brutos da inspeção	169
Tabela 36: Resultado ajustado da inspeção	170
Tabela 37: Manifestantes atacam a polícia em Hong Kong.	178
Tabela 38: Ocorrência de metáforas envolvendo os esquemas de container e força.	185
Tabela 39: As seções temáticas dos canais dedicados à mulher	192
Tabela 40: Categorias de produtos anunciados nos portais da amostra.	194
Tabela 41: O amor como patologia, segundo o Terra.	195
Tabela 42: O canal Mulher tematiza o casamento.	196
Tabela 43: Composição das notícias dedicadas à mulher em 21 novembro de 2006.	197

LISTA DE ILUSTRAÇÕES

Ilustração 1: O modelo estelar de comunicação	29
Ilustração 2: O modelo web de comunicação	31
Ilustração 3: Os quadrantes mágicos do hipertexto.	47
Ilustração 4: Domínios visitados e tempo de page view na web.	51
Ilustração 5: Domínios mais visitados pelos usuários brasileiros.	53
Ilustração 6: Exemplo de comunicação não verbal.	66
Ilustração 7: Representação visual de conjuntos nebulosos.	71
Ilustração 8: O sinal gráfico de um carro ao longo do tempo.	75
Ilustração 9: Escala de modalidade das cores na representação naturalista.	77
Ilustração 10: Fotografia de Hugo Faleiros.	77
Ilustração 11: Um possível lugar para as convenções de gêneros na concepção de Kress e Van Leeuwen.	81
Ilustração 12: Uma codificação simples.	89
Ilustração 13: Ciberespaço plano com topologia heterogênea.	97
Ilustração 14: O ciberespaço tridimensional.	101
Ilustração 15: Uma representação quadridimensional do ciberespaço.	103
Ilustração 16: Protótipo (intuitivo) de um portal de interesse geral.	115
Ilustração 17: Convenções para os valores lingüísticos de verdade.	133
Ilustração 18: Uma curva de distribuição normal..	136
Ilustração 19: Distribuição dos valores predominantes de verdade.	137
Ilustração 20: Outra distribuição dos valores predominantes de verdade..	140
Ilustração 21: A estrutura de links dos hiperdocumentos do experimento Copérnico..	156
Ilustração 22: A home page do UOL em 20 de novembro de 2006.	171
Ilustração 23: Ilustração do relato do Terra.	175
Ilustração 24: Ilustração do relato do Globo.	177
Ilustração 25: O canal Mulher do portal Terra.	187
Ilustração 26: O canal dedicado à mulher do UOL.	188
Ilustração 27: O canal do iG.	189
Ilustração 28: La Beauté, segundo o Terra.	190
Ilustração 29: Horóscopo no UOL	193

LISTA DE SÍMBOLOS

\wedge	conjunção
\rightarrow	implicação
\exists	quantificador existencial
	tal que
\in	pertence a
\cong	aproximadamente igual a
$=$	igual a
Σ	somatório
\cap	interseção

SUMÁRIO

O sonho do engenheiro	13
Os portais: a restauração da cultura do espetáculo	15
Um problema e muitas hipóteses	17
Que ciência é essa – nota metodológica	21
1. O CIBERESPAÇO: INTERAÇÃO À VELOCIDADE DA LUZ	24
1.1. Da informação como espetáculo à turbulência informacional	25
1.2. A Internet: novo meio de interação em massa	35
1.3. A interação virtual	39
1.4. A comunidade virtual	42
1.5. O hipertexto e as teias do ciberespaço	43
1.6. As curvaturas do ciberespaço	48
2. GÊNEROS DO DISCURSO – ALGUMAS ÂNCORAS TEÓRICAS	55
2.1. O problema – teórico – dos gêneros do discurso	56
2.1.1. <u>O Locutor, o Interlocutor e o Outro: gêneros do discurso em Bakhtin</u>	58
2.1.2. <u>Gêneros e eventos comunicativos: a perspectiva de Swales</u>	65
2.1.3. <u>A multimodalidade e a produção do sentido em Kress e Van Leeuwen</u>	73
2.1.4. <u>As mensagens do meio: gêneros do discurso em Maingueneau</u>	82
2.2. Canais (marcianos) e veículos de comunicação	87
2.3. O lugar dos gêneros digitais no ciberespaço	95
2.3.1. <u>Um modelo multidimensional para o ciberespaço</u>	96
2.3.2. <u>Um experimento identificador de novos gêneros digitais</u>	105
2.3.3. <u>O gênero portal de interesse geral</u>	111
3. ALGUMAS CURVATURAS NO CIBERESPAÇO: OS ATRADORES RETÓRICOS	118
3.1. Satisfação garantida ou seu dinheiro de volta: o domínio temático dos portais	119
3.1.1. <u>Uma caracterização vaga e imprecisa dos gêneros digitais</u>	120
3.1.2. <u>A seleção do corpus para a análise</u>	124
3.1.3. <u>As propriedades temáticas dos portais – uma primeira hipótese</u>	128
3.1.4. <u>A metodologia de coleta e análise dos dados</u>	131
3.1.5. <u>Primeira avaliação da amostra – a relevância das propriedades</u>	134
3.1.6. <u>Segunda avaliação – a prototipicidade da amostra</u>	138
3.2. O controle da palavra do outro nos dispositivos de busca	140
3.2.1. <u>Diferentes concepções dos dispositivos de busca</u>	143
3.2.2. <u>Primeiro experimento: o diferencial das hipóteses interpretativas</u>	148
3.2.3. <u>Segundo experimento: em busca da palavra do Mesmo</u>	152
3.2.4. <u>Terceiro experimento: o super-controle da palavra do outro nos portais de interesse geral</u>	160
3.3. O labirinto do surfista	165
3.3.1. <u>A retórica dos links: o hipertexto degenerado</u>	165
3.3.2. <u>A retórica da publicidade – surfando no UOL</u>	170
4. AS VOZES DO MESMO – INCURSÕES NOS DISCURSOS DOS PORTAIS	173
4.1. A retórica da verdade no relato jornalístico	174
4.1.1. <u>Referente, relato, objetividade</u>	174
4.1.2. <u>Pressuposição e argumentação</u>	179
4.1.3. <u>Protesto é guerra</u>	182
4.2. A construção da identidade feminina nos portais	186
4.2.1. <u>Eu, mulher jovem, magra e sensual</u>	186

4.2.2.	<u>Dize-me o que consumes e...</u>	191
4.2.3.	<u>Relacionamentos – o jogo da sedução</u>	194
4.2.4.	<u>A título de síntese</u>	196
4.3.	O discurso para todos e para o mesmo	198
5.	FECHANDO PARA BALANÇO	200
	REFERÊNCIAS	204

O SONHO DO ENGENHEIRO

O engenheiro sonha coisas
claras: superfícies, tênis, um copo de
água.
(João Cabral de Melo Neto. O
Engenheiro)

Desde o momento em que decidi retomar a carreira acadêmica em sua área de origem, após mais de uma década dedicada ao mundo da tecnologia da informação, seja como profissional de engenharia de *software*, seja como editor, escritor e professor, sempre me pareceu óbvio que o tema de pesquisa deveria envolver esse novo campo de atividade (comunicativa) humana – a Internet. E não apenas por se tratar do meu principal ramo de atividade profissional. O processo de “commoditização” dessa tecnologia não parece ter paralelo na história: entre a invenção da imprensa e sua ampla disseminação, séculos se passaram; algumas décadas foram necessárias para que a televisão se tornasse um eletrodoméstico de largo uso; no entanto, em menos de dez anos a Internet se tornou fenômeno de massas. O Comitê Gestor da Internet no Brasil (CGI.BR, 2005: 44) estima que 32,2% da população brasileira já acessou a Internet pelo menos uma vez e que aproximadamente 21% dos domicílios têm acesso regular à rede. Como o fenômeno está diretamente associado à escolaridade e à renda (o acesso à rede atende a 53,4% dos domicílios com renda superior a R\$ 1.800,00, a 87% da classe A e a 58% da classe B), é certo que “a Internet em casa vai aumentar na medida em que crescer a escolarização da população” (CGI.BR, 2005: 58). Isso é extremamente significativo num país de população pobre e pequena tradição cultural como o Brasil. E ao contrário da televisão, trata-se de mídia com ampla vocação interativa: o CGI estima também que 81,67% dos usuários utilizam a rede para o que podemos chamar “viver com os outros pela linguagem”, isto é, enviar *e-mails* e mensagens instantâneas, participar de *chats* e comunidades, criar e manter *blogs* e *websites*, entre outras atividades de interação comunicativa. É inevitável que um fenômeno com tal alcance afete de algum modo não apenas o uso da linguagem como o próprio comportamento dos envolvidos.

A implícita relevância do fenômeno, portanto, me permitiu satisfazer uma inclinação e ainda manter relativo interesse acadêmico na pesquisa. Porém, mais que a “natural” inclinação, uma inquietação me movia. Àqueles poucos usuários e profissionais (à época, a

distinção praticamente inexistia) envolvidos com a Internet no princípio da década passada a rede sempre apareceu como um “território livre”, construído tecnologicamente em torno dos princípios da liberdade e da colaboração – uma terra onde poderia ser exercida a liberdade de ser (ou parecer), de dizer (ou escrever) e de ouvir (ou ler), para o bem ou para o mal. Afinal, onde mais podíamos encontrar tanto o ponto de vista dos palestinos quanto o dos israelenses, expressos por eles próprios, embora estes últimos nunca tenham precisado da Internet para se expressar? Ainda que essa tendência tenha se mantido com a massificação do uso (até porque se trata das próprias características construídas para a rede ao longo do tempo), forças de todo opostas passaram a encontrar nela espaço para seu exercício de poder – no mínimo porque a Internet não é um fenômeno apartado da vida social, mas profundamente entranhado nela (e novamente para o bem e para o mal) –, forças com vocação para homogeneizar (e possivelmente pasteurizar) o “território livre da diferença”, de modo a assegurar sua possibilidade de controle.

Não se trata, naturalmente, de uma “conspiração” levada a cabo por “forças ocultas” (embora a ação de alguns governos pareça contradizer essa afirmação), mas da própria natureza (contraditória) da vida social, onde os atores não apenas *representam* papéis, mas disputam seu espaço em cena. Com Foucault, suponho “que em toda sociedade a produção do discurso é ao mesmo tempo controlada, selecionada, organizada e redistribuída por certo número de procedimentos que têm por função conjurar seus poderes e perigos” (FOUCAULT, 2004: 8). Ora, na “era da informação”, sua origem é poder e sua circulação descontrolada, um perigo.

Evidentemente, os controles técnicos são aqueles mais óbvios e que recebem maior atenção. Ações como as iniciativas do governo chinês de bloquear o acesso a informações “perigosas”; o patrocínio do governo francês de um dispositivo de busca gaulês para se contrapor ao domínio anglo-saxão exercido pelo Google, Yahoo e MSN; o consórcio de inúmeras empresas e governos em torno de um projeto de rede “segura” (isto é, onde as pessoas não possam realizar determinadas atividades) – estes são apenas alguns exemplos mais recentes e mais evidentes dessas tentativas de controle. Minha inquietação, porém, não se refere a essas tentativas evidentes de controle, simplesmente porque seu alcance é limitado. Interessa-me mais o “controle do discurso”, que parece mais eficaz. Eficaz porque assumido por livre consentimento; eficaz porque voluntariamente buscado por aqueles a quem é reservado o papel de objeto do controle; eficaz porque raramente notado tanto por quem o exerce quanto por aqueles para quem é exercido. Assim, mais que a Internet, meu objeto é o

seu controle – o que acabou por implicar em conceber esse projeto de pesquisa como um projeto *político*, onde a *polis* é o ciberespaço.

Por outro lado, esse projeto “micro-político” implícito na pesquisa deverá me permitir fugir ao risco de me comportar como o gato da fábula, que resolveu estudar ornitologia para melhor compreender o que devorava, quando a natureza o queria apenas como ornitófono. O risco evidente é conhecer para melhor devorar a presa! Ao empreender essa pesquisa com base naquela inquietação quero, sim, estudar o predador para melhor compreender sua presa – e se não tenho a pretensão de libertá-la, espero pelo menos esclarecer a natureza da caça...

Os portais: a restauração da cultura do espetáculo

Com esse programa definido, minha atenção se voltou rapidamente para o que se convencionou chamar “portais de interesse geral”. Trata-se de vastos empreendimentos comerciais dedicados ao fornecimento de *conteúdo*, este último termo tendo um significado todo especial quando aplicado ao problema. Ao procurar descrever o que os avatares dos empreendimentos comerciais na Internet pregavam na virada do milênio, Orlando assinala que:

Conteúdo (...) é o que possa interessar aos visitantes de um site: publicação de notícias, relatórios especializados, material produzido por usuários - em salas de bate-papo, fóruns de discussão, enquetes, material reaproveitado de produções para mídias offline, críticas de eventos esportivos e impressões sobre restaurantes, shows e filmes escritas por usuários e o que seja motivo de atenção, de acordo com a estratégia de público (ORLANDO, 2001: 83).

A definição deliberadamente engloba quase tudo o que possa interessar aos usuários da Internet... E, de fato, os portais de interesse geral mais bem sucedidos são aqueles que funcionam como uma espécie de cartola mágica de onde praticamente qualquer coisa pode ser extraída. Duas são as fontes de renda básicas para esses empreendimentos: o fornecimento de acesso à Internet e a comercialização de espaço publicitário. A primeira deve ser, para todos os efeitos práticos, irrelevante; a ser de outro modo, um dos principais portais brasileiros, o iG, já teria falido há muito, uma vez que o seu apelo original era (e ainda é) o fornecimento de acesso *gratuito*. Logo, a principal fonte de remuneração desses empreendimentos deve ser a publicidade. Portanto, o projeto comercial desses empreendimentos depende de fixar os usuários nos *sites*, de modo a favorecer a comercialização de espaço publicitário, proporcional à quantidade de usuários e ao tempo de exposição aos *banners*. Desse modo, todos se

organizam como grandes portais de conteúdo, englobando vastas constelações de *sites* na maioria dos casos respondendo no mesmo domínio¹, mas não necessariamente. Tais *sites* procuram responder a diferentes interesses, de modo a assegurar que o máximo possível de anseios de seus usuários seja satisfeito sem a necessidade de navegar para fora dos seus domínios. Esses portais procuram atrair e manter os usuários em sua órbita. Em outras palavras, o sucesso desses empreendimentos requer o *controle* do comportamento de navegação dos usuários – o que, neste caso, não pode ser feito senão através do discurso (persuasivo).

A hipótese que salta aos olhos é evidente: e se este não for apenas um controle *pelo* discurso, mas um controle *do* discurso? O interesse pela questão aumenta se avaliarmos o alcance dessa categoria de *sites*, em especial junto ao público usuário conectado à Internet a partir da rede brasileira. Em pesquisa publicada em 14 de julho de 2005 (RELATÓRIO, 2005), a organização Ibope determinou que os usuários brasileiros gastam 10,5% do seu tempo de acesso à Internet navegando nesses *sites*, percentual expressivo também quando comparado a outros países, conforme ilustrado na Tabela 1, cujos dados são referentes ao primeiro trimestre de 2005. Os dados obtidos por essa pesquisa sugerem também que o perfil do comportamento dos usuários brasileiros é diferente dos outros países pesquisados. Note-se que as cinco categorias da pesquisa com maior tempo dedicado respondem, no Brasil, por 53,2% do tempo total consumido na Internet, enquanto que as cinco categorias mais utilizadas nos Estados Unidos e Espanha respondem apenas por 30,4% e 39,8%, respectivamente, sugerindo um uso mais fortemente concentrado e menos diversificado por parte dos usuários brasileiros. Além disso, se compararmos o tempo dedicado às principais categorias de *sites* obtidas no Brasil com o tempo dedicado às mesmas categorias nos outros países, notaremos que o tempo dedicado a elas responde apenas a 21% nos EUA e 31,3% na Espanha. Isso sugere hábitos e interesses fortemente distintos.

Categoria	Brasil	EUA	Espanha
Comunidades	20.5	2.8	5.1

¹ No jargão da Internet, um domínio é definido como um conjunto de endereços de rede identificados por um nome, como, por exemplo, *www.uerj.br*. O termo é derivado da Ciência da Computação, onde tende a se referir a um conjunto de conhecimentos identificados por um nome.

E-mail	11.3	7.7	10.9
Portais de interesse geral	10.5	6.6	6.9
Instituições financeiras	6.2	-	-
Ferramentas de busca	4.7	3.9	8.4
Jogos online	4.0	7.2	3.8
Classificados e leilões	1.5	5.0	1.7
Fabricantes de software	-	-	8.5

Tabela 1: Tempo de exposição relativo dos usuários da Internet (Fonte: Ibope)

Esses portais funcionam como “portais” de acesso a outros *sites* (daí o nome) temáticos, dedicados aos diferentes categorias do “conteúdo” publicado. Todos os portais chamam esses *sites* de *canais*, utilizando-se da metáfora da televisão. Isso nos leva a uma suspeita imediata: o princípio organizador dos portais seria o mesmo dos meios de comunicação de massas convencionais, numa tentativa de importar seu modelo de comunicação para a *web*? Como veremos no próximo capítulo, nossa premissa é de que esses modelos são substancialmente distintos: o modelo dos *mass media* seria baseado no *espetáculo*, enquanto que o modelo do ciberespaço seria nitidamente *interativo*. De fato, uma simples consulta ao portal Terra (<http://www.terra.com.br>), um dos prováveis “campeões de audiência” no Brasil, nos permite confirmar que essa suspeita merece investigação. Aquele portal fornece 60 domínios de conteúdo diferentes (sob a forma tanto de *canais* quanto de *serviços*). No entanto, apenas 12% desses canais são interativos ou permitem que os usuários se comportem como fonte de informação e não apenas receptores (um comportamento típico da Internet); além disso, nenhum desses canais permite que os usuários empreendam projetos colaborativos de qualquer espécie.

Um problema e muitas hipóteses

Isso parece suficiente para justificar a principal pergunta de pesquisa deste trabalho. Por coerência com o que se veio expondo, essa pergunta deve estar relacionada ao perfil de navegação peculiar ao usuário brasileiro. A idéia foi simplesmente correlacionar os dois problemas suscitados pela pesquisa citada na seção anterior, postulando a possibilidade de que o comportamento pouco diversificado dos usuários brasileiros esteja associado às

características dos *sites* preferencialmente visitados. Como discutiremos no próximo capítulo, esse comportamento parece estar em desacordo com o comportamento que pode ser inferido das características da Internet. Assim, é possível formular o primeiro problema de pesquisa através da proposição:

- (i) Os portais de interesse geral afetam o comportamento navegacional esperado (esperado com base no modelo comunicacional da Internet) dos usuários?

É evidente que essa questão é mais abrangente que os meios à nossa disposição para sua verificação empírica. Isso provavelmente deveria envolver pesquisa de campo, acompanhando os hábitos de navegação de amostra significativa dos usuários da Internet que não acessam regularmente os portais, confrontando-os com os hábitos de amostra de usuários típicos desses *sites*. Uma pesquisa desse tipo está fora de cogitação em vista dos meios disponíveis. Desse modo, ela será assumida como *premissa* da própria pesquisa. No entanto, utilizando os meios à nossa disposição, procuraremos refletir sobre a proposição (i), de modo a tornar sua resposta afirmativa menos arbitrária.

Essa reflexão estará no centro das preocupações do capítulo 1. *O ciberespaço: interação à velocidade da luz*, que tem como objetivo primário contextualizar o leitor, situando a pesquisa numa determinada interpretação do que seja o ciberespaço. A idéia é procurar inferir das características tecnológicas e históricas da rede um determinado modelo interpretativo do que seja esse campo de atividade (comunicativa) humana que chamamos ciberespaço. Idealmente, esse modelo deveria prever um determinado tipo de comportamento navegacional, o que nos permitiria pelo menos argumentar em favor de uma resposta à proposição (i). Como se verá naquele capítulo, esse comportamento esperado é contraditório com o sugerido pela pesquisa do Ibope que acabamos de citar, o que nos permitiria justificar a resposta à proposição (i) e afirmar, pelo menos metaforicamente, que os portais de interesse geral provocam uma curvatura no ciberespaço, afetando o comportamento navegacional dos usuários, que seriam atraídos para órbitas no seu interior.

Esse artifício argumentativo nos permitirá identificar um interessante problema – a anomalia nos hábitos de navegação no ciberespaço dos usuários conectados a partir da rede brasileira – e uma (dentre muitas) possível hipótese explicativa: a popularidade de alguns *sites* cuja organização e propósito provocam uma distorção, uma curvatura no ciberespaço nas suas vizinhanças, afetando esses hábitos. Essa hipótese inicial conduz ao nosso problema de pesquisa – buscar respostas à pergunta a propósito de que mecanismos são utilizados pelos portais de interesse geral para distorcer a topologia do ciberespaço, curvando-o nas suas vizinhanças, questão que pode ser formulada do seguinte modo:

- (ii) Quais os mecanismos utilizados pelos portais de interesse geral para curvar o ciberespaço e afetar o comportamento navegacional dos usuários?

Há certamente inúmeras abordagens para esse problema. Podemos, por exemplo, focar o *empreendimento* e buscar formular as perguntas necessárias à resolução daquele problema em termos de estratégias empresariais. É essa uma das principais abordagens da única pesquisa fora dos domínios estritamente tecnológicos sobre os portais que fomos capazes de encontrar no Brasil (ORLANDO, 2001). O foco da pesquisa de Orlando referia-se ao jornalismo na *web* e o ponto de partida que utilizou foi a abordagem do *empreendimento* para daí chegar a um *artefato* em particular – o portal UOL. Propomos uma outra abordagem, mais compatível com o modelo interpretativo do ciberespaço que adotamos.

Como veremos no próximo capítulo, a Internet é um novo meio de interação (comunicativa) em massa cujas características técnicas historicamente desenvolvidas propiciam o surgimento de novas dimensões culturais – a cibercultura. Veremos também que a conexão à rede implica no bloqueio de tudo que não seja estritamente material semiótico (comunicativo). Portanto, é de se esperar que todos os artefatos encontráveis no ciberespaço sejam *artefatos de linguagem* (natural, mas não somente) e que todos os eventos ali ocorrentes sejam *eventos comunicativos* (entre seres humanos, mas não somente). Desse modo, é possível abordar o problema que nos move encarando os portais como artefatos de linguagem e formulando as questões de modo compatível com esse enfoque, passando o nosso interesse a se concentrar nos mecanismos *discursivos* que curvam o ciberespaço nas vizinhanças dos portais de interesse geral. Infelizmente, essa abordagem requer a resolução de alguns outros problemas antes de ser iniciada.

A razão é que, no estágio atual dos estudos da linguagem aplicados à Internet, os portais em si mesmos (isso não se refere, evidentemente ao seu conteúdo) não são unânime e inequivocamente considerados artefatos de linguagem. Um dos pioneiros da pesquisa lingüística da Internet no Brasil, Marcuschi, ao discutir os novos gêneros discursivos emergentes no ciberespaço, não considera como tal “a *home page* (portal, sítio, página), já que ela não passa de um ambiente específico para localizar uma série de informações, operando como um suporte e caracterizando-se cada vez mais como serviço eletrônico” (MARCUSCHI, 2004: 20). Em vista disso, é preciso resolver preliminarmente este problema. Diante disso, precisaremos adiar a indagação da proposição (ii) e formular uma nova, a saber:

- (iii) Os portais de interesse geral podem ser considerados um gênero particular do discurso?

Tal pergunta, se respondida afirmativamente, nos permitirá abordar essa categoria de *sites* como artefatos de linguagem. Isso será abordado nos capítulos 2. *Gêneros do discurso* –

algumas âncoras teóricas. Nossa estratégia será buscar construir um modelo teórico capaz de reconhecer um gênero digital e determinar o seu lugar no ciberespaço, construção que será feita simplesmente a partir da avaliação de diferentes teorias dos gêneros do discurso, em busca de insumos para o modelo. Do mesmo modo, a objeção de Marcuschi, centrada no argumento de que a *home page* (e os portais, evidentemente), são *suportes* para gêneros e não gêneros em si, deverá ser abordada, assim como outras questões correlacionadas.

Resolvido o problema, poderemos então nos dedicar à questão mais relevante (para nós) de compreender como se dá o funcionamento discursivo dos portais com referência a outros gêneros digitais, buscando responder a proposição (ii). Trata-se aqui de determinar como o modo de funcionamento discursivo do portal contribui para afetar os hábitos de navegação dos seus usuários, curvando e distorcendo o ciberespaço nas suas vizinhanças. Para isso procuraremos abordar o que chamamos *atratores retóricos*, isto é, aquelas estratégias (discursivas) destinadas a seduzir, convencer ou mesmo fraudar os usuários de modo a assegurar sua permanência na órbita dos portais. O capítulo 3. *Algumas curvaturas no ciberespaço: os atratores retóricos* será dedicado ao tema, notando-se apenas que não examinaremos *todas* as estratégias discursivas disponíveis e eventualmente utilizadas nos portais, apenas aquelas próprias a esse novo meio de interação em massa – a Internet.

Um último problema precisará ser abordado, para mantermos coerência com o modelo interpretativo do ciberespaço que descreveremos brevemente no próximo capítulo. Como se verá, as características do ciberespaço afetam (ou pelo menos podem afetar) significativamente o modo como construímos nossas identidades sociais e culturais, permitindo a elaboração de identidades virtuais. Alertaremos também para o fato de que os portais apenas marginalmente admitem a interação discursiva dos usuários, buscando restaurar o modelo comunicativo dos *mass media*. Ora, nos *mass media* o poder de representar (e, portanto, de determinar e definir a identidade) não pertence aos espectadores. Se os portais buscam reproduzir o modelo dos *mass media*, então é pelo menos possível que eles, nos seus discursos, projetem e imponham identidades aos usuários, *sujeitandos-os* e restringindo a liberdade de ser, de dizer e de ouvir que assinalamos como típica do ciberespaço. Em outras palavras, os portais podem ser vetores da *ordem do discurso* de que falava Foucault, um elemento de controle do *caos* típico do ciberespaço.

Se for um dispositivo discursivo, o portal deve igualmente ser um *dispositivo identitário*; e se é um dispositivo discursivo fechado e não interativo, então deve ser igualmente um dispositivo de *imposição* de identidades às quais os usuários devem aderir para

participar do “jogo” do discurso dos portais. Desse modo, nossa pesquisa deverá abordar uma nova questão, bastante ampla, a saber:

- (iv) Que imagem de sujeito se constrói a partir do funcionamento discursivo dos portais de interesse geral?

Nesse ponto nossa abordagem não poderá ser abrangente, isto é, não tentaremos elaborar proposições aplicáveis a *todos* os portais, simplesmente porque procuraremos analisar alguns *textos* produzidos neles, buscando inferir dessa análise os *discursos* que os articulam. Assim, adotaremos a esse ponto a postura de “estudos de caso”, avaliando alguns artefatos textuais “interessantes”, o que será feito no capítulo 4. *As vozes do Mesmo – incursões nos discursos dos portais*. Em vista das evidentes limitações de tempo a que se sujeita a pesquisa, abordaremos o problema de forma igualmente limitada, bastando-nos com apenas dois estudos de caso. O objetivo aqui é menos dar uma visão abrangente do problema e mais chamar atenção para ele.

Que ciência é essa – nota metodológica

Pode-se objetar facilmente que as questões apresentadas nos parágrafos anteriores são excessivamente vagas e abrangentes para permitir a formulação de hipóteses facilmente verificáveis num *corpus* em particular. De fato – e isso foi deliberado. O problema é que a conversão dessas questões genéricas em *proposições passíveis de teste empírico* requer, muitas vezes, raciocínio longo e amplas justificativas. Isso transcende o escopo de uma simples introdução ao projeto de pesquisa. Desse modo, a abordagem de cada uma das questões nas seções apropriadas da dissertação será precedida pela formulação dessas proposições passíveis de teste, assumindo que a solução (ou corroboração) de tais proposições nos permitirá responder (provisoriamente) as questões genéricas aqui formuladas.

Tais proposições deverão, obrigatoriamente, ser limitadas e precisas, de modo a satisfazermos os requisitos de abrangência e profundidade requeridas por uma dissertação de mestrado. Se bem elaboradas, tais proposições nos permitirão responder a uma gama vasta de perguntas de pesquisa de modo completo, pelo menos no que concerne à completude requerida pela proposição. Daí termos corrido o risco de conferir certo aspecto “panorâmico” a esta dissertação. De fato, isso ocorreu. Porém, a diversidade de questões colocadas sem

prejuízo da abrangência e completude foi possível por força da precisão e limitação das proposições formuladas para respondê-las de forma satisfatória.

Trata-se, então, de definir o que sejam essas proposições e o que entendemos por teste. De uma forma geral, nossa metodologia consistirá em propor modelos interpretativos (assumidos como mais ou menos arbitrários) para os vários problemas, isto é, *teorias* em sentido estrito. Esses modelos devem ser construídos de tal forma que possamos *deduzir* (ou pelo menos *inferir*) deles *previsões* passíveis de teste empírico. Testes, para nós, deverão (sempre que possível) ser estruturados como *experimentos* destinados a *negar* as previsões feitas, de tal modo que possamos *rejeitar* o modelo interpretativo proposto. Por esse método, um bom modelo, uma boa teoria é simplesmente aquela que *ainda* não conseguiu ser rejeitada (por um teste empírico decisivo) e, conseqüentemente, substituída por uma nova (e provisória) teoria.

É preciso observar inequivocamente que, para nós, experimentos não são “provas”, mas corroborações (ou falseamentos) de teorias. Isso significa que admitimos que os experimentos sejam elaborados (e interpretados) com base nas “previsões” de uma teoria, sendo sempre possível encontrar explicações diferentes daquelas previstas, bastando para isso supor um conjunto diferente de premissas para um mesmo experimento. Isso quer dizer que não depositamos fé cega nos experimentos, necessariamente limitados pelo que queremos corroborar. Incidentalmente, assinalamos que cada um dos “experimentos” descritos nesta dissertação deve ser avaliado com base nessa concepção.

É fácil concluir que estamos assumindo a atitude (e pelo menos parte do método) filosófica de Popper com relação à ciência. Estamos, por exemplo, assumindo que “a ciência não é um sistema de enunciados certos ou bem estabelecidos” que avança continuamente do não conhecimento para o conhecimento; para aquele filósofo da ciência, “ela jamais pode proclamar haver atingido a verdade ou um substituto da verdade, como a probabilidade” (POPPER, 1975: 305). Popper assume que “não sabemos: só podemos conjecturar” – e que essas conjecturas assumem a forma de “antecipações, de intentos temerários e prematuros e de preconceitos”, no entanto “cuidadosamente controlados por testes sistemáticos” (POPPER, 1975: 306). Para o filósofo, o “método” científico

não se orienta no sentido de defendê-las [aquelas “antecipações”] para provar que tínhamos razão. Pelo contrário, procuramos contestar essas antecipações. Recorrendo a todos os meios lógicos, matemáticos e técnicos de que dispomos, procuramos demonstrar que nossas antecipações são falsas – a fim de colocar, no lugar delas, novas antecipações injustificadas e injustificáveis, novos “preconceitos temerários e prematuros”, como Bacon pejorativamente as denominou (POPPER, 195: 306).

Em outras palavras, o melhor que se pode esperar daquilo que chamamos ciência não é a verdade, mas simplesmente um processo contínuo e interminável de eliminação dos erros –

através de experimentos. Estamos cientes que esse método não é aquele normalmente aplicado nas ditas Ciências Humanas; estamos igualmente cientes de que ele traz implicações filosóficas no mínimo polêmicas para inúmeras correntes da Lingüística; estamos principalmente cientes de que essa não é a concepção de ciência de muitos daqueles pesquisadores cujas idéias nortearam nosso trabalho. Não é sequer certo de que o objeto da Lingüística se deixe estudar por este método. Daí não sermos estritos na sua aplicação: nem todos os enunciados que propomos falsear ao longo do nosso esforço têm a forma lógica de proposições falseáveis por *modus tolens*; nem todas as análises que realizamos se baseiam em *experimentos* (mesmo que imaginários); nem todos os modelos que propomos têm a forma de um conjunto rigoroso de proposições lógicas elaboradas a partir de inferências dedutivas. Na verdade, as concepções de Popper transparecem no nosso esforço acadêmico menos como um método e mais como uma *atitude* perante a pesquisa. Por que, então, insistir nela? Porque ela justifica, torna mais compreensível e possivelmente tolerável (para atitudes distintas) o modo particular como construímos nossa argumentação, nossa insistência em enfatizar experimentos e não a análise de textos. Mas também para melhor definir o lugar daquele projeto “micro-político” de que falávamos logo nos primeiros parágrafos deste capítulo – o lugar reservado aos preconceitos reconhecidos como temerários e prematuros, mas deliberadamente submetidos ao crivo empírico e ao teste intersubjetivo.

1. O CIBERESPAÇO: INTERAÇÃO À VELOCIDADE DA LUZ

Ao introduzir uma das pesquisas pioneiras sobre os gêneros digitais do discurso, Marcuschi assinala que “o impacto da Internet é menor como revolução tecnológica do que como revolução dos modos sociais de interagir lingüisticamente” (MARCUSCHI, 2004: 19). É possível que tal avaliação – conquanto ousada para a média dos pesquisadores no domínio dos usos da linguagem – acabe por vir a ser considerada cautelosa e conservadora em se tratando das possíveis alterações na cultura e sociedade balizadas por invenção de uma ordem não vista desde Gutenberg e Marconi. É provável que não possamos exigir mais da avaliação de um fenômeno que mal se inaugura, do mesmo modo que não devemos ser capazes de avaliar em sua totalidade a fúria de uma tempestade estando no olho do furacão. No entanto, a experiência passada na avaliação da agregação de novos meios de comunicação de massa nos permite algumas suposições razoáveis.

O já clássico McLuhan enfatiza que “a 'mensagem' de qualquer meio ou tecnologia é a mudança de escala, cadência ou padrão que esse meio ou tecnologia introduz nas coisas humanas” (MCLUHAN, 1979: 22). A invenção de Gutenberg ampliou a capacidade de leitura de alguns poucos monges e leigos letrados para as enormes massas humanas, embora essa ampliação dos sentidos só tenha se efetivado séculos depois da sua inauguração, com a universalização da alfabetização. O rádio nos permitiu escutar pessoas localizadas a centenas e milhares de quilômetros – e o radiotelescópio nos permite mesmo realizar o sonho do poeta e... ouvir estrelas! A televisão ampliou visão e audição (ou o sentido do *tato*, se preferirmos as proposições de McLuhan) para o planeta inteiro – e mesmo fora dele. Os *mass media* elevaram aquilo a que comumente nos referimos por informação à escala planetária, capacitando-a a circular à velocidade da luz; essas novas *extensões* da capacidade humana globalizaram a cultura num sentido bem concreto: a expressão “aldeia global” é um lugar comum que ocorre naturalmente nesse contexto. Essa *globalização da informação* se desenvolve sob o signo da velocidade (da luz): “hoje invadimos culturas inteiras com informação acondicionada, diversão e idéias. Em vista do alcance global instantâneo dos novos meios de visão e som, até mesmo o jornal é vagaroso” (MCLUHAN, 1982: 147).

Trata-se de uma avaliação anterior à globalização da Internet – e se ela é correta, podemos supor que a introdução da rede mundial de computadores tenha no mínimo acelerado essa tendência. Independente de como avaliamos o “impacto” da Internet na

sociedade, sua massificação crescente, aliada às suas características técnicas pode ter criado um novo campo para a atividade humana. Não são poucos os pesquisadores se referem a esse novo campo como *ciberespaço*, termo que, como o vocábulo “robô”, ultrapassou sua fronteira original na literatura, disseminando-se para o universo acadêmico e, de lá, para o uso cotidiano. Tais pesquisadores têm se utilizado igualmente do termo *cibercultura* para se referir ao produto humano das interações sociais nesse novo campo de atividade. Só isso é o bastante para indicar a possível emergência de um novo campo de atividade humana.

O objetivo do presente capítulo é, portanto, conceituar o *ciberespaço* e possivelmente delinear as linhas gerais da *cibercultura* relevantes para a presente pesquisa. A primeira pergunta que cabe responder é simples: é possível falarmos mais do que apenas metaforicamente no *ciberespaço*? Existe realmente uma *cibercultura*? Por que não podemos encarar a Internet simplesmente como uma nova versão dos já conhecidos *mass media*? Afinal, a introdução da TV e a universalização da imprensa não produziram nenhuma cultura típica, ainda que tenham alterado a cultura existente. É, portanto, legítimo abordarmos algo como a *cibercultura*? Assim, precisaremos aqui discutir em detalhes aquelas características que tornam possível falarmos de um novo campo de atividade humana e não simplesmente de um novo meio de comunicação.

1.1. Da informação como espetáculo à turbulência informacional

Um evento emblemático da era dos *mass media* pode facilmente ser situado há mais de trinta anos, na transmissão global da missão Apolo 11 para praticamente cada aparelho de TV disponível no planeta. Uma escada, alguns passos; um astronauta perfilado refletindo as estrelas no seu capacete; uma única frase – “we came in peace for all mankind”. Essa única cena resume todo o programa dos *mass media*: informação condensada e plana, em somente duas dimensões, sem profundidade alguma; informação que exige participação (entendida aqui no sentido próprio de McLuhan), mas não reflexão. Toda a indústria dos *mass media* se condensou metaforicamente nessa única cena, tão *encenada* quanto um *take* cinematográfico: baseada num *script* detalhado; ensaiada à exaustão; refletindo a concepção do diretor e não dos atores, cujo desempenho assegura apenas a sensação de “realismo”.

A comunicação como espetáculo (global) – tal o programa dos *mass media*. A televisão é um dispositivo comunicacional unidirecional modelado como uma estrela (ver LÉVY, 1999: 63 e seguintes), onde um único centro ativo dissemina radialmente suas mensagens em todas as direções. O “receptor” nesse dispositivo não é, evidentemente, passivo. Aliás, partimos da premissa de que na comunicação ele *nunca* é passivo: o sentido não é dado de antemão, constituindo-se *durante* o processo de comunicação, em maior ou menor grau, com a participação ativa do “receptor”. No entanto, o dispositivo estelar não é diretamente interativo: nele, as respostas do receptor não realimentam o circuito, salvo indiretamente. Exemplos típicos dessa realimentação indireta do circuito na televisão são as modificações introduzidas nos programas a partir da interpretação de resultados de pesquisas de audiência. Nesse tipo de dispositivo, as respostas do receptor não retornam ao circuito da comunicação senão como respostas médias a perguntas apenas indiretamente relacionadas às mensagens que efetivamente trafegam no “meio” (as pesquisas de audiência). Elas são principalmente respostas *interpretativas* e, como tais, não realimentam diretamente o circuito de comunicação. Nesse tipo de dispositivo, seu *feedback* é sempre indireto, realizado *fora* dos limites do circuito, através de outros eventos comunicativos.

No entanto, aqui o “receptor” não se limita a *fruir e interpretar*, embora essas sejam provavelmente as respostas mais comuns ao estímulo comunicacional: novas mensagens podem ser produzidas em resposta direta e imediata à mensagem recebida. O evento comunicacional sobrevive sob outras formas durante e após sua realização, da conversa cotidiana às manchetes dos jornais. É preciso, porém, notar que os números de jornais e revistas inteiramente dedicados à missão Apolo, os comentários, as conversas e debates – enfim, todos esses novos eventos comunicativos foram respostas a um outro evento comunicativo (a transmissão em si) e não eventos independentes, dedicados à missão: somente a câmera de Aldrin testemunhou o ocorrido durante a missão – todos os demais “testemunhos” jornalísticos foram *testemunhas* apenas da transmissão. Portanto, todas as reportagens sobre a missão foram, na verdade, um relato sobre a transmissão de Aldrin, um relato secundário sobre como Aldrin *narrou* as ações de Armstrong. Em todos os casos, porém, as respostas produzidas não realimentaram (nem podiam fazê-lo) o circuito comunicacional da TV, permanecendo como eventos secundários e relativamente independentes.

Nesse modelo comunicacional a informação tem centro: tudo o que pode ser dito sobre a missão está disponível somente para o transmissor original – os próprios astronautas da missão e a equipe de terra, responsável pelo planejamento de cada evento dela. Esse modelo

espera que todos os demais participantes do circuito se comportem apenas como *espectadores*, como receptores do espetáculo. A capacidade desses espectadores se comportarem também como fontes de informação é reduzida ao mínimo pelas características do modelo – eles só podem fazê-lo nas fronteiras do sistema.

O circuito de comunicação da TV – aliás, como todos os circuitos estelares, o rádio comercial, o jornal, a revista, etc. – prevê um único centro transmissor e isso impõe que as respostas dos receptores não possam se dar senão na periferia do sistema, principalmente como respostas interpretativas: eles são aqui *espectadores* e não *interlocutores*. O diálogo direto implícito em *toda* comunicação humana é reduzido ao mínimo. A comunicação torna-se espetáculo.

No circuito do espetáculo, onde apenas uma única fonte transmissora é admitida, a informação² se perde na direção da periferia. Isso é evidente: se ela é a medida da liberdade de escolha no processo de seleção de mensagens, a probabilidade de escolha de uma mensagem em particular deve ter um valor menor no centro do sistema, o que corresponde à nossa incerteza (maior) quanto a que mensagem será selecionada. A medida que nos movemos em direção à periferia do circuito – isto é, a medida que as mensagens são escolhidas e a informação é transmitida – essa incerteza só pode decair para valores de máxima probabilidade, até a certeza. Em outras palavras, a *entropia* sempre aumenta na direção da periferia, como esperado.

As implicações filosóficas disso são imediatas. Em primeiro lugar, a informação (entendida no sentido que lhe empresta Shannon) só pode ser perdida, nunca recuperada. Isso indica que, como qualquer sistema físico, a comunicação humana, como portadora de informação, deve estar igualmente sujeita à desorganização e indiferenciação crescentes.

² Ao procurar responder ao problema de se é possível medir a informação, Shannon (1948) supôs que, do ponto de vista da engenharia do circuito comunicacional, tudo se passa como se o emissor de uma mensagem (a fonte de informação) selecionasse o que será transmitido de um conjunto finito de mensagens possíveis. Nesse contexto, a informação mede a *liberdade de escolha* da fonte no processo de seleção de mensagens. No caso mais simples possível, onde apenas duas mensagens igualmente prováveis estão disponíveis para a fonte, essa medida é definida como um *bit*, isto é, o logaritmo na base 2 da quantidade de mensagens disponíveis. Portanto, do ponto de vista de um observador (incluindo o destinatário), a informação é a medida da nossa *incerteza* sobre que mensagem será selecionada pela fonte. Isso significa que, para Shannon, a informação não está relacionada ao que é dito (o significado das mensagens), mas ao que *poderia* ser dito. Esse modo de medir a informação, baseado na *probabilidade* associada à escolha de cada mensagem do conjunto, permite que Shannon associe a informação à noção física de *entropia*, a tendência inerente aos sistemas físicos de se tornarem cada vez mais desorganizados.

Podemos imaginar a comunicação unidirecional como o acoplamento de duas máquinas – a fonte de informação (o emissor) e o destinatário (o receptor). O sistema resultante pode ser analisado como uma máquina de estado discreto³. Se esta visão é correta, a trajetória desse sistema, isto é, a sucessão dos seus estados ao longo do tempo, deve obedecer à tendência geral à entropia crescente. Isso é fácil de se observar.

Suponhamos que estamos observando um circuito de comunicação unidirecional ao longo do tempo, onde uma fonte transmite informação para um receptor. No instante t_0 , que pode corresponder, por exemplo, ao estado α do sistema, nossa incerteza é máxima: não temos como saber qual mensagem, das muitas possíveis, será enviada, já que elas devem ter probabilidades de ocorrência muito próximas entre si. Qual a primeira frase numa conversação sobre a qual não sabemos o tópico? O estado α é o estado de máxima incerteza, onde, portanto, a entropia é menor. No entanto, no instante t_1 – digamos, no estado β do sistema – nossa incerteza deverá, necessariamente, ser menor. Por exemplo, numa aula de Linguística gerativa, é altamente provável que à seqüência de símbolos “colorless green ideas sleep” se siga o símbolo “furiously”. A incerteza, portanto, decai ao longo da trajetória desse sistema – e a entropia, portanto, sempre aumenta.

Naturalmente, na comunicação humana real, mesmo quando analisada numa única direção (isto é, para um receptor humano inteiramente calado), a curva de entropia crescente não é linear. Fórmulas de etiqueta, regras de conversação, frases feitas, tópicos rituais, etc. – numa palavra, os gêneros do discurso (embora não apenas eles) contribuem para a redução da incerteza na comunicação, produzindo flutuações locais na entropia, mesmo naquele instante t_0 de que falávamos. Por exemplo, é muito provável (embora não uma certeza) iniciarmos uma conversa telefônica enviando a mensagem “alô”. Além disso, ao longo da trajetória dessa conversação, é muito provável que o interlocutor pontue uma locução longa do emissor com mensagens de *handshake* como “hum-hum”, “certo”, “sei”, entre outras fórmulas típicas desse “gênero” de conversação. No entanto, a *tendência* geral é sempre em direção à certeza e à entropia máxima. Vista dessa forma, a comunicação unidirecional tende a transitar “de um

³ Uma máquina de estado discreto é simplesmente um dispositivo que transita de um conjunto de estados bem definidos para outro. Nesse dispositivo, cada estado é definido por um conjunto de variáveis discretas, cuja alteração de valor implica na mudança do estado corrente para um novo.

estado de organização e diferenciação, em que existem formas e distinções, a um estado de caos e mesmice” (WIENER, 1993: 14).

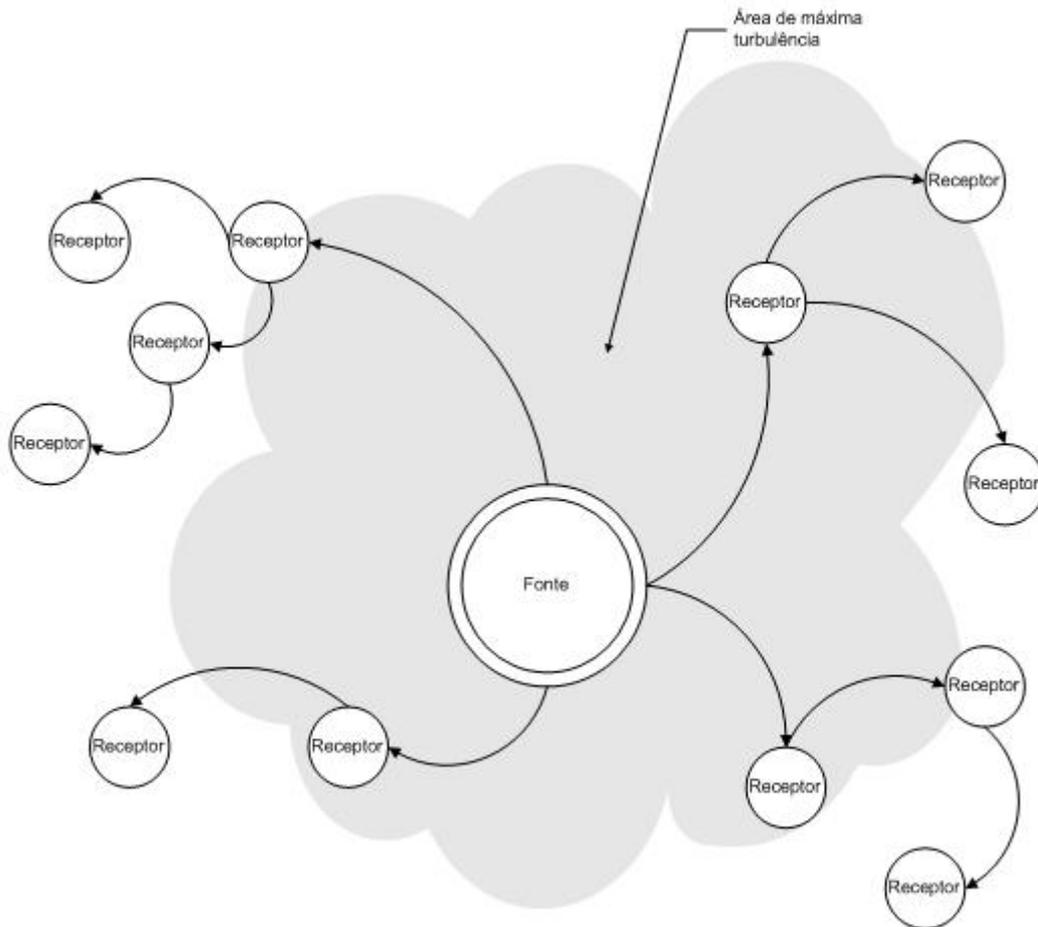


Ilustração 1: O modelo estelar de comunicação

Essa longa digressão nos serviu principalmente para demonstrar que o modelo de comunicação estelar típico dos *mass media* é inerentemente entrópico. Nele, a informação se perde rapidamente do centro para a periferia, em que pesem algumas poucas flutuações locais, correspondentes às respostas interpretativas dos receptores. A circulação da informação nesses circuitos pode ser concebida como na Ilustração 1, onde o centro do sistema é a área de máxima incerteza, que decai a medida que progredimos para sua periferia. O espetáculo da informação é, pois, apenas a perda da informação...

Qualquer circuito de comunicação não interativo, mesmo aqueles de topologia diferente da estelar, deve se comportar de forma bem semelhante ao da TV e dos *mass media*. Nesses circuitos, a possibilidade de controle (do modo como a informação circula no circuito) é bastante alta, simplesmente porque eles têm um único centro emissor. Daí pesquisadores como Lemos (1973: 73), à esteira de filósofos como Deleuze, utilizarem a metáfora da árvore

(em oposição à metáfora do rizoma) para descrevê-los, enfatizando a hierarquia de controle implícita no modelo dos *mass media*. O modo mais simples desse controle é exercê-lo sobre a fonte, já que no modelo estelar não é possível controlar diretamente o modo como os destinatários recebem, interpretam e, eventualmente, reagem às mensagens. Esse controle, quando ocorre, se exercita principalmente através da redução das mensagens possíveis de serem selecionadas pela fonte – por exemplo, pela proibição, mais ou menos rigorosa, imposta ou auto-imposta, de determinados assuntos, expressões, idéias, atitudes, etc. Pelo enfoque que vimos dando ao problema, esse controle implica, portanto, na redução da incerteza global do sistema, isto é, no aumento da sua entropia. Emblema desse controle é a concessão, em países como o Brasil, pelo poder público dos direitos de transmissão nos *mass media* mais abrangentes, como o rádio e a TV, na ação de agências reguladoras externas ou de “comissões de ética” internas, nas eventuais sentenças de tribunais, etc. O que nos importa aqui é que este controle só é eficaz por conta da arquitetura (centralizada) do modelo de comunicação próprio dos *mass media*.

No contexto da presente pesquisa, essa longa abordagem do modelo de comunicação dos *mass media* importa apenas porque nos auxilia a compreender melhor a radical inovação do modelo de comunicação típico da Internet. Nossa abordagem utilizará um outro evento emblemático para descrever esse modelo, no caso a difusão do atentado terrorista ao World Trade Center, em 11 de setembro de 2001. Tal como a missão Apollo, o evento foi transmitido em tempo real pelos principais meios de comunicação de massas. No entanto, grande parte do seu impacto social não decorreu da transmissão por esses meios, mas pela sua difusão na nova *media*. O interessante aqui foi a eclosão de inúmeros outros eventos comunicativos diretamente relacionados às transmissões realizadas pelos órgãos de imprensa e publicadas na Internet. Em tempo real, os usuários da rede, onde quer que um computador estivesse conectado, não apenas *assistiam* ao evento como *participavam* dele, comentando-o, polemizando, divulgando novas informações (reais ou imaginárias), cada um deles servindo igualmente como nova fonte de informação (expressão aqui utilizada tanto no sentido dicionarizado quanto no técnico). E isso ocorria, ao contrário da transmissão da missão Apollo, no *interior* do próprio circuito de comunicação original. *Sites* de notícias, *chats*, correio eletrônico, *blogs*, *newsgroups*, comunidades virtuais – um sem número de recursos típicos da Internet participaram desse processo de difusão.

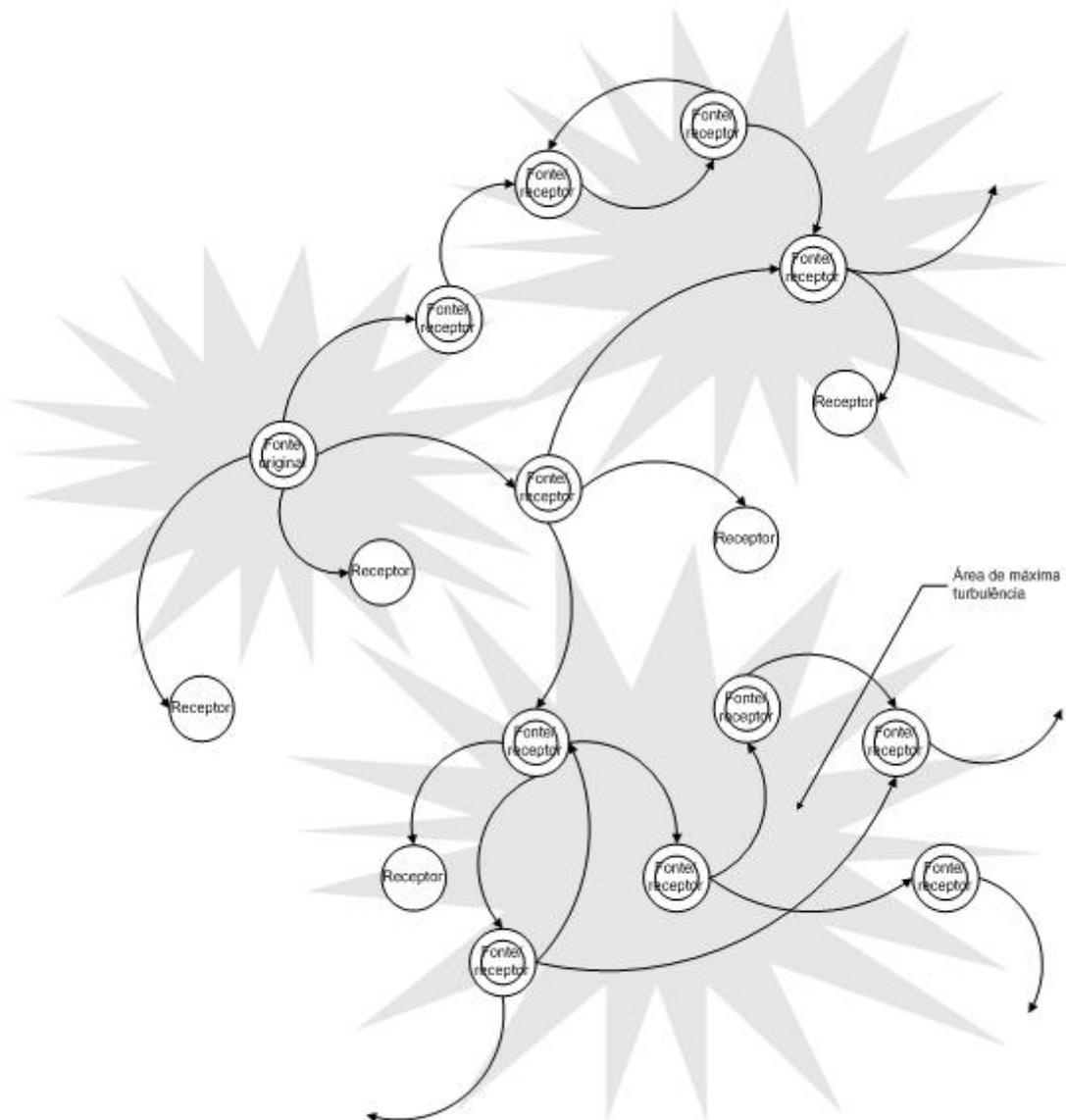


Ilustração 2: O modelo *web* de comunicação

Talvez a melhor expressão para esse caso não seja *difusão*, já que não se tratou apenas de *transmitir* as mensagens originais, mas de produzir e difundir milhares de novas mensagens, algumas inteiramente contraditórias entre si. Em cada nó do circuito, as mensagens recebidas não eram apenas retransmitidas para outros nós: novas mensagens eram criadas e agregadas ao fluxo global de circulação, já que cada usuário da rede pode funcionar não apenas como receptor (de informação), mas como *fonte*. Os prédios ainda não haviam desabado e piadas (de humor negro, é claro) já circulavam pelos correios eletrônicos e grupos de distribuição. Nesse cenário, a melhor expressão que nos ocorre a ser utilizada em lugar de *difusão* é a *expansão turbulenta* da informação. Tudo se passa como se a informação explodisse de um (ou provavelmente vários) ponto inicial, o processo de expansão ocasionando novas explosões de informação, algumas delas possivelmente maiores que as

primeiras. Uma representação gráfica desse processo turbulento de circulação da informação é exibido na Ilustração 2.

O que queremos enfatizar com essa representação é a existência de inúmeros *pólos* de transmissão de informação, nenhum deles mais relevante do que os demais. Essa rede *não tem centro!* Os receptores são igualmente fontes de informação. As novas mensagens não são produzidas fora das fronteiras do sistema, como no modelo estelar, mas no seu próprio interior – e essa é a natureza desse modelo! Mais do que “rizoma”, expressão forjada por Deleuze e Guattari (1995), é a *teia* (a *web*) a expressão que melhor caracteriza esse modelo, expressão que pode ser facilmente utilizada a propósito da representação gráfica da Ilustração 2.

Como nessa teia cada um dos nós é não apenas destino como origem da informação e se analisarmos a trajetória de cada uma das “regiões” do circuito, concluiremos que entropia pode não aumentar simplesmente a medida que o tempo passa e os estados do sistema se sucedem. Ao contrário, notaremos a ocorrência de “ilhas” de entropia negativa nas diversas trajetórias, picos de incerteza que correspondem à entrada em linha das novas fontes de informação. Naturalmente, do ponto de vista do sistema como um todo, a tendência à entropia crescente, inerente a todos os sistemas físicos, permanece. No entanto, pelo menos em nível local, atuam contra-tendências em direção ao aumento da incerteza e à redução da entropia. O universo em volta pode agonizar; porém, a vida floresce nessas ilhas de entropia negativa, para utilizarmos a imagem poética de Norbert Wiener.

Essa contra-tendência é possível porque o modelo comunicativo *web* é eminentemente *interativo*, permitindo que qualquer receptor (da informação) seja igualmente fonte no mesmo processo comunicativo. Para dar conta dessa característica, hoje trivial nos problemas da comunicação humana, passaremos a utilizar, em lugar de fonte e receptor, as expressões *locutor* e *interlocutor*, salvo quando aquelas forem preferíveis em vista de contexto específico (em especial a teoria da informação). Uma característica evidente desse modelo é ele ser *ex-cêntrico* e, como tal, não ter periferia, apenas *fronteiras*, que o distinguem de outros sistemas comunicativos. É por esta razão que aqueles controles, fáceis de serem exercidos sobre os *mass media*, parecem inaplicáveis à Internet. Trata-se de uma caixa de Pandora tecnológica: uma vez solta na rede, a informação não pode ser controlada.

Essa impossibilidade de controle ligada à própria arquitetura tecnológica em torno da qual a rede foi fundada pode ser facilmente ilustrada pelos fracassos seguidos do governo chinês em restringir a navegação dos seus cidadãos. Para isso, aquele governo elaborou diversas estratégias, que têm sido objeto de reportagens na imprensa ocidental. As mais

relevantes (com maior – embora ínfima – chance de sucesso) envolvem a tentativa de criar centros (parciais) na rede. Se não pode controlar que computadores estão conectados à rede, um país autocrático, que restringe o empreendedorismo privado, pode, no entanto, controlar (e restringir) as conexões da sua rede nacional à rede internacional. Na maioria dos países, qualquer empresa ou instituição com recursos suficientes pode estabelecer seus próprios canais de acesso à rede internacional (*backbones*), utilizando-os privadamente ou liberando-os (mediante remuneração) ao público através dos provedores de acesso. São esses os canais que ligam as redes nacionais à rede internacional. É o caso, por exemplo, do Brasil, que conta com vários canais de acesso. Na China, no entanto, apenas o estado tem esse direito. Como esses canais de acesso são necessariamente limitados (pelo seu alto custo), é possível estabelecer neles mecanismos de controle. É o que fez (ou tentou fazer) a criatividade governamental chinesa.

O controle envolveu a utilização não convencional de um número limitado de *firewalls*, introduzindo artificialmente um centro numa rede ex-cêntrica. Os *firewalls* são dispositivos computacionais projetados para analisar (e eventualmente bloquear) o tráfego de rede. Tais dispositivos foram construídos para localizar tráfego que possa representar ameaças aos usuários (ataques de *crackers*, invasões à privacidade, vírus, etc.) com base em padrões já razoavelmente conhecidos e bloqueá-lo. No entanto, como qualquer arma, essa também aponta para dois lados. Como a maior parte do tráfego na Internet é legível por seres humanos, a maior parte da informação trafegando em língua natural, é possível utilizar os *firewalls* para localizar requisições a sites específicos e/ou contendo palavras “suspeitas”, bloqueando-as. Que palavras podem ser essas deixo à imaginação e às convicções do leitor. O preço a pagar, no entanto, é elevado: o processamento computacional envolvido torna *todo* o tráfego de rede proibitivamente lento, não apenas aquele considerado suspeito. Isso prejudica mesmo o uso “legítimo” da rede, incluindo o próprio uso governamental. A solução encontrada envolveu a restauração da ex-centricidade da rede: o *firewall* apenas analisaria o tráfego, que seria bloqueado por outros recursos computacionais. Os detalhes técnicos não nos importam aqui; importa o fato de que o controle, ao envolver a restauração da ex-centricidade da rede, poderia ser vencido. Assim é que pesquisadores (ocidentais, naturalmente) demonstraram rapidamente *como* o controle poderia ser vencido, publicando seus resultados na Internet e condenando ao fracasso a iniciativa chinesa.

Um outro controle pode ser estabelecido, este com interesse mais direto para a nossa pesquisa. Sabemos que a arquitetura tecnológica da Internet impede a existência de um centro, fazendo com que qualquer nó da rede possa ser tanto receptor quanto fonte de informação.

Isso torna possível que todo o conhecimento registrado em meio magnético no planeta possa vir a estar disponível em algum lugar da rede. Na medida em que a Internet se globaliza, esta possibilidade está se tornando cada vez mais próxima. No entanto, essa mesma arquitetura tecnológica cria um outro problema: *como* encontrar o conhecimento? Ele está acessível, sim – mas somente se conhecermos sua localização! E como a rede foi projetada sem centros, esses “endereços” não estão disponíveis num único repositório. A solução encontrada pela própria comunidade tecnológica foram os dispositivos de busca.

Os *search engines* são dispositivos computacionais destinados a localizar e indexar qualquer documento publicado na rede, através de autômatos que varrem sistematicamente todos os seus nós. Havendo poder computacional bastante e dado um tempo suficientemente longo é pelo menos teoricamente possível a esses dispositivos catalogar *todos* os documentos jamais publicados na Internet. Assim, os *search engines* podem funcionar como imensos “catálogos telefônicos” da Internet, fornecendo – pelo menos teoricamente – a localização de qualquer conhecimento disponível na rede. Essa necessidade dos usuários é de tal ordem que o *site* Search Engine Watch estima que sejam realizadas na rede 213 milhões de buscas (ver SULLIVAN, 2006) por dia em todo o planeta nas poucas centenas de dispositivos existentes! Como são relativamente poucos, esses dispositivos funcionam como “centros de informação” numa rede sem centro – e, portanto, podem ser controlados de algum modo. Se tivermos o “controle” desses dispositivos, podemos, por exemplo, fornecer, como resultado de uma busca, determinados *sites* de interesse nosso e não quaisquer outros, estes se tornando inacessíveis na prática. Podemos também escolher *não* responder a determinadas buscas.

Esse controle, uma vez mais, ainda que possível, é limitado na prática e novamente pela arquitetura tecnológica da rede: afinal, qualquer um com relativamente poucos recursos pode instalar o seu próprio dispositivo de busca e funcionar como um centro de informação. Na verdade, existem centenas de dispositivos funcionais já catalogados. Na prática, porém, os dispositivos realmente utilizados pelo público são muito poucos. A Nielsen//Net Ratings, em pesquisa permanentemente disponível no Search Engine Watch, estima que apenas 3 *sites* (o Google, o Yahoo e o MSN) são responsáveis por 81,1% de todas as buscas realizadas na Internet. Isso torna viável o controle – e foi dessa possibilidade que se aproveitou o governo chinês na sua eterna ânsia de controle.

Para permitir que as empresas responsáveis por aqueles *sites* realizassem negócios na China, o governo daquele país exigiu (e foi atendido) que seus dispositivos não respondessem a buscas realizadas utilizando-se de expressões consideradas “suspeitas”, nem que retornassem determinados *sites* independentemente do tipo de busca realizada. Esse controle é

pelo menos temporariamente eficaz, dada a barreira da língua. Além disso, mesmo para aqueles poucos chineses com conhecimento de inglês as versões americanas dos dispositivos não são acessíveis, já que a estrutura centralizada da rede chinesa permite que eles sejam facilmente bloqueados. Embora qualquer cidadão chinês possa acessar qualquer *site* em qualquer lugar do planeta, dada as características técnicas da rede, e maioria não será capaz de fazê-lo simplesmente porque não sabe onde estão localizados.

Não nos alongaremos demais no problema nesse momento. No capítulo a propósito dos dispositivos de busca discutiremos em mais detalhes os mecanismos de controle que podem ser aplicados com algum sucesso aos dispositivos de busca. Neste ponto, queremos apenas assinalar que os dispositivos de busca, por se tratarem de poucos centros de informação indispensáveis a uma rede ex-cêntrica, representam (por enquanto) o único flanco acessível ao controle dessa caixa de Pandora chamada Internet. Mesmo assim, o controle total é impossível, como assinalamos, o que torna a Internet o único meio de comunicação onde a “liberdade de expressão” (e de acesso à informação) pode ser total e integralmente praticada. Essa é, talvez, a característica mais relevante (pelo menos para o meu ponto de vista político particular) dessa nova mídia.

1.2. A Internet: novo meio de interação em massa

Vimos que é a arquitetura tecnológica da Internet, ao constituir uma rede ex-cêntrica, que a torna um espaço de (quase) irrestrita liberdade de expressão, o que faz com que a informação circule de modo turbulento e explosivo, de tal modo que Lévy, em seu prefácio à obra de Lemos, se refere a ela como “um espaço de comunicação propriamente surrealista, do qual ‘nada é excluído’, nem o bem, nem o mal, nem suas múltiplas definições, nem a discussão que tende a separá-los sem jamais conseguir” (LEMOS, 2002: 14). Deve-se notar apenas que esta não é uma característica acidental, mas o *objetivo primário* da rede. Não é necessário nos alongarmos aqui na história da rede, de resto disponível em Castells (2003), em especial no seu primeiro capítulo; basta-nos assinalar que ela nasceu como um projeto militar destinado a encontrar alternativas de comunicação para a eventual destruição dos centros de comando dos Estados Unidos por um hipotético ataque nuclear. Ora, a alternativa encontrada foi uma rede cujo centro de comando não pudesse ser destruído simplesmente

porque inexistia... Em outras palavras, a rede nasceu com características tais que visavam impedir o seu controle por potências hostis. Como decorrência inevitável desse objetivo, ela se tornou avessa a seu controle por *qualquer* entidade.

Essa característica foi reforçada pela sua evolução posterior, levada a cabo por uma comunidade acadêmica fortemente influenciada pelos movimentos de base típicos da sua época, fazendo com que “a cultura da liberdade individual que floresceu nos *campi* universitários nas décadas de 1960 e 1970 [usassem] a interconexão dos computadores para seus próprios fins” (CASTELLS, 2003: 25). Por conta disso, o mecanismo de agregação de novas tecnologias e padrões à rede sempre foi fortemente participativo, aberto à discussão e à intervenção de *toda* a comunidade técnica interessada, mecanismo utilizado até hoje. Era, portanto, inevitável que essa evolução fosse acompanhada pela inclusão de tecnologias que facilitassem a *participação* e a *interação* ativa dos usuários da rede. Ao contrário da maioria das novas tecnologias computacionais, originadas nas pesquisas da indústria e dos governos, as tecnologias da Internet se originaram a partir da própria comunidade de usuários, colaborando em projetos tecnológicos. Todos os grandes padrões e tecnologias utilizados na Internet nasceram desse modo: a *web*, o *chat*, o *e-mail*, etc. – a lista simplesmente abrange *todos* os recursos de *software* utilizados na rede. Neste caso, a indústria simplesmente seguiu a comunidade de usuários, bem depois das tecnologias serem inventadas.

Essas tecnologias e padrões se originaram de projetos colaborativos de iniciativa espontânea das comunidades tecnológicas. Era inevitável que dessa interação surgisse “um conjunto de valores e crenças”, isto é, uma cultura. E é essa cultura que Castells e Lévy chamam de *cultura hacker*, um subconjunto da cibercultura, que “emergiu das redes de programadores de computador que interagiam on-line em torno de sua colaboração em projetos autonomamente definidos de programação criativa” (CASTELLS, 2003: 38). Essa cultura autônoma permanece ativa e forte até hoje, mesmo com a “invasão” pela indústria dos mecanismos colaborativos clássicos da Internet (como o IETF – *Internet Engineering Task Force*, responsável por todos os padrões da Internet). Pensemos na – imensa – comunidade de *software* livre que continua a florescer.

Trata-se de um movimento de base na Internet iniciado por Richard Stallman, pesquisador do MIT, na segunda metade da década de 1980. Esse pesquisador (e agora as centenas de milhares de participantes dos mais de 180 mil projetos colaborativos de *software* disponíveis no *site* Source Forge – <http://www.sourceforge.net>) concebe o *software* como conhecimento humano materializado e, como tal, não deveria ser patenteado e comercializado, mas *compartilhado* por todos. Ele se propunha a produzir *software* em

colaboração com quaisquer voluntários, compartilhando com eles seus conhecimentos e divulgando-os livremente para todos os usuários interessados. Assim, fundou a *Free Software Foundation*, mais conhecida como Projeto GNU, cujo lema principal é “*Free as in Freedom*”. Para essas iniciativas – tão bem sucedidas que a própria indústria aderiu – o “*software* livre se refere à liberdade dos usuários executarem, copiarem, distribuírem, estudarem, modificarem e melhorarem o *software*” (THE FREE, 2006). O sucesso é tal que mais de 80% do tráfego Internet se baseia em aplicativos desenvolvidos sob o regime de *software* livre, pela própria comunidade especializada de usuários.

Portanto, colaboração e interação na busca e no exercício da liberdade – tais são as palavras-chave da cultura *hacker*, expressões que acabam por se materializar nos produtos desenvolvidos por essa comunidade. É a encarnação desses princípios no *software* produzido por essa tribo que permite que eles possam migrar para *toda* a comunidade, simplesmente através do seu exercício. Talvez o melhor exemplo da migração desses princípios para além da tribo tecnológica seja a Wikipedia (<http://www.wikipedia.org>). Trata-se de uma “enciclopédia livre”, um projeto colaborativo para o registro em regime de compartilhamento e liberdade dos conhecimentos disponíveis, contendo hoje mais de 3 milhões de verbetes em 10 idiomas, incluindo o Português. Naturalmente, o *site* e os aplicativos que suportam a iniciativa foram produzidos em regime de *software* livre; porém, o que nos importa é o fato de essa enciclopédia ser produzida por *qualquer* usuário da rede que queira participar voluntariamente. Os colaboradores simplesmente acessam funções públicas do *site* e editam os verbetes existentes, produzem artigos novos, criticam o material disponível – enfim, interagem colaborativamente para registrar o conhecimento disponível na comunidade. A participação é inteiramente livre e a avaliação dos produtos é igualmente livre, baseada no escrutínio dos próprios usuários.

É provável que se registre ali muita tolice; aliás, os próprios mantenedores são os primeiros a reconhecê-lo e convidam os usuários a formularem suas críticas e proporem correções. Não é certo, porém, que a iniciativa privada equivalente mais consagrada – a Enciclopedia Britannica – esteja totalmente livre de tolices; afinal, a variedade e a especialidade dos conhecimentos ali registrados por pessoal remunerado (e, portanto, necessariamente limitado tanto em abrangência quanto em qualidade) devem ser fortes dificultadores para assegurar a precisão do material. Aliás, qualquer especialista em qualquer assunto pode contribuir com a Wikipedia (e não poucos o fazem), o que não é verdadeiro para a Enciclopedia Britannica. Do ponto de vista do conhecimento ali registrado, porém, é mais importante o fato de se tratar do *conhecimento da comunidade* ali comunitariamente

partilhado. Mais do que a “sociedade do conhecimento” o que a Wikipedia anuncia é a *comunidade do conhecimento*.

O que queremos assinalar com estes exemplos é o fato de a arquitetura tecnológica da rede, por suas características históricas, ter fornecido a infra-estrutura material necessária para que a atividade humana ali realizada se baseasse fortemente na interação e na colaboração, livremente definidas e adotadas entre os agentes. Interação e colaboração em massa – creio serem estes os subprodutos mais relevantes da arquitetura tecnológica da Internet. E é este subproduto que torna possível à rede ser um palco para o exercício da liberdade (pelo menos por seus usuários). E é esta infra-estrutura material que permite que a atividade humana na rede acabe por produzir um conjunto de valores e crenças específico – a cibercultura. É certo que, tomados isoladamente, esses valores e crenças estão presentes em outras “culturas”; afinal, a Internet não é uma nova sociedade, mas uma extensão das sociedades humanas atuais. No entanto, é a velocidade com que tais valores se disseminam, a liberdade com que podem ser exercidos em conjunto, a facilidade com que podem ser partilhados, a impossibilidade de se estabelecerem fronteiras políticas e nacionais que tornam o ciberespaço um novo campo de atividade humana, capaz de produzir uma cultura própria.

São essas características que podem ser inferidas facilmente da avaliação da difusão turbulenta dos eventos de 11 de setembro de 2001. Comparando esses eventos com o espetáculo televisivo da missão Apolo, podemos concluir que são essas as características mais relevantes que fazem com que a Internet não possa ser definida simplesmente como um “meio de comunicação de massa”. Se queremos uma expressão equivalente, é melhor defini-la como um “meio de interação em massa” (embora ainda não *das* massas, já que a maior parte da população mundial ainda se encontra para lá da fronteira da exclusão digital). Nos *mass media* não temos escolha: a nós é reservado o papel de espectadores. Na Internet, a escolha é nossa: podemos ser os atores – e a maioria dos usuários escolhe esse papel, mesmo que não o faça conscientemente. Basta lembrarmos que 81,67% dos usuários da Internet no Brasil (ver Tabela 2, mais adiante) usam a rede para “viver com os outros”, para interagir através de *e-mails*, mensagens instantâneas, *chats*, *blogs* e comunidades virtuais como o Orkut – enfim, para *compartilhar* conhecimentos e informações, crenças e valores, sentimentos e atitudes.

Ao contemplarmos esses vastos empreendimentos colaborativos, acabamos por repensar a avaliação clássica de McLuhan, de que os meios de comunicação de massa “frios” (como a televisão) exigem a participação do espectador. Essa participação de fato pode ser verificada, mas se trata de “participação interpretativa”, onde o destinatário ainda é um *espectador*, mesmo que chamado a incluir seus próprios significados aos sinais recebidos. Na Internet, a

participação não é uma metáfora, é parte integrante do ambiente: os usuários são chamados a participar desde a construção da infra-estrutura técnica da rede até à produção dos conhecimentos ali difundidos. Daí enfatizarmos a *colaboração* como marca definidora do ciberespaço. Na televisão (e no rádio) apenas a massa “participa”, interferindo apenas indiretamente na produção das mensagens. Nos *mass media* é a totalidade da audiência, a multidão quem interage com o enunciador; são as expectativas pressupostas do público, tal como delimitadas pelos índices de audiência e pesquisas de opinião, que determinam a enunciação. Na Internet é o indivíduo que colabora; ele é responsável direto (na medida em que queira sê-lo) pela produção dos enunciados. Isso pode ocorrer simplesmente como um ato participativo elementar como o envio de um *e-mail*; pode, porém, tornar-se uma atividade discursiva para milhões, através da criação e manutenção de um *website* (bem sucedido).

1.3. A interação virtual

Vimos que as características técnicas construídas ao longo do tempo fazem da Internet um ambiente que propicia fortemente a *interação colaborativa* entre seus usuários – e essa é provavelmente a característica mais relevante dessa mídia, se a compararmos aos meios de comunicação de massa em geral. Ela, portanto, pode ser facilmente considerada um novo modo de ocorrência das interações sociais. No entanto, ela traz uma outra característica técnica bastante óbvia, mas que merece avaliação. Trata-se de uma *rede de computadores* de âmbito mundial; nela, portanto, toda interação é *mediada* pela máquina. Isso, certamente, traz conseqüências para a interação social – e precisa ser avaliado.

A conexão com a Internet age como um *firewall* filtrando para fora do ciberespaço um conjunto de elementos presentes nas interações sociais “convencionais”. Numa interação na Internet, o fato é que não temos como saber *quem* se apresenta do outro lado da conexão: *a priori*, não temos como ter certeza sequer de que se trata de pessoas já conhecidas. É tão fácil alterar o endereço do remetente de um *e-mail* para um outro que não o nosso próprio, que normalmente não temos como ter nenhuma garantia a propósito de quem enviou determinada mensagem. Numa sala de *chat* ou numa seção de *messaging*, os interlocutores são apenas o que *dizem* ser. Não temos sequer como assegurar que nossos interlocutores não sejam, na

verdade, homenzinhos verdes oriundos de Alpha Centauri! Apenas nossa *crença* na sinceridade dos outros permite que afirmemos *quem* são nossos interlocutores.

Isso deve afetar de algum modo a construção das identidades sociais. O *firewall* representado pela conexão à Internet filtra, de saída, alguns dos atributos mais relevantes na construção dessas identidades, como cor da pele, gênero civil, idade, etc. Tais atributos, evidentemente, não são eliminados; no entanto, ao não estarem disponíveis para o interlocutor, podem reduzir seu peso na construção das identidades sociais no ciberespaço. Além disso, algumas semioses não verbais significativas para a construção dessas identidades são igualmente filtradas para fora do ciberespaço: convenções de vestuário e aparência, gestos, tons e timbres de voz, sotaques – são alguns dos elementos que, ao serem utilizados diretamente ou indiretamente na comunicação não mediada pela rede de computadores, contribuem de algum modo para a construção das várias identidades sociais e não podem estar presentes (senão marginalmente, já que os recursos de vídeo-conferência ainda não são largamente utilizados) no ciberespaço.

Este *firewall* aceita somente a codificação escrita da linguagem natural e as semioses típicas da televisão – sons de apoio, imagem, sinais gráficos e possivelmente animação, embora o peso dessas semioses nas interações sociais no ciberespaço ainda não seja predominante. É importante assinalar que, ao contrário das interações sociais fora do ciberespaço, essas semioses não-verbais estão inteiramente sob controle consciente dos locutores, o que deve afetar de algum modo seu comportamento observável. É certo que nossa imagem corporal é *construída*; no entanto, essa construção é feita, em grande medida, *fora* do nosso controle consciente. Isso pode não ocorrer na Internet. Por exemplo, numa comunidade como o Orkut, onde os participantes devem previamente se apresentar por seus atributos, incluindo fotografia, nossa imagem é conscientemente projetada através de escolhas (normalmente) criteriosas daqueles atributos que desejamos ver conhecidos, incluindo nosso próprio rosto. Literalmente falando, “na Internet você é o que diz ser” (CASTELLS, 2003: 109). Isto não significa, porém, que falsificamos deliberadamente nossa “identidade” na Internet (não há evidências de que isso seja feito na Internet com frequência maior que fora dela); significa apenas que temos um controle consciente maior sobre a construção das identidades que assumimos nas interações na rede. E mais: tudo o que está disponível para a construção dessas identidades é aquilo que somos capazes de *comunicar* por escrito e, secundariamente, através de algumas poucas semioses áudio-visuais. Se “quem tem o poder de representar tem o poder de definir e determinar a identidade” (SILVA, 2000: 91), na

Internet, na medida em que tudo o que temos à disposição são *representações*, nossa capacidade de determinar nossa identidade é bem maior que fora dela.

É importante assinalar que esse *firewall* de que falávamos implica não apenas num novo modo de interação lingüística, num novo modo de comunicação, mas num novo modo de interação social mediada exclusivamente pela linguagem. E como tal esse tipo de interação não substitui de modo algum as demais interações sociais. Avaliando um conjunto de pesquisas destinadas a verificar essa possibilidade, Castells assinala que “o corpo de dados não sustenta a tese de que o uso da Internet leva a menor interação social e maior isolamento social” (CASTELLS, 2003: 104). Embora o pesquisador não assinale o fato, é possível dizermos que, ao contrário, as características da rede possibilitam o aparecimento de todo um novo modo de sociabilidade inteiramente baseado na linguagem. Essa é uma inferência razoavelmente evidente dos argumentos apresentados ao longo deste capítulo. Ela pode ser corroborada pelos dados obtidos pelo Comitê Gestor da Internet no Brasil, conforme Tabela 2

Atividade	% de usuários
Comunicação	81,67
Enviar e-mail	70,51
Enviar mensagens instantâneas	34,74
Participar de chats	28,21
Participar de comunidades (Orkut)	21,97
Participar de listas de discussão	7,82
Criar blogs	10,20
Utilizar telefone/conferência	6,78
Criar e manter websites	6,66
Treinamento e educação	56,45
Busca de informações e serviços online	80,96
Lazer	71,25
Interação com autoridades públicas	28,22
Banking, compra e venda de bens e serviços	17,68

Tabela 2: Atividades realizadas na Internet (Fonte: CGI.BR, 2006)

Como se pode notar, a Internet no Brasil é utilizada principalmente para a interação comunicativa e a busca de informações. O detalhamento daqueles dados indica que parte substancial dessa interação comunicativa *não* é individual, mas coletiva, como pode se notar no detalhamento do item Comunicação da Tabela (participação em *chats*, comunidades e listas de discussão). Isso significa que devemos avaliar cuidadosamente a interpretação de Castells de que “o novo padrão de sociabilidade em nossas sociedades é caracterizado pelo

individualismo em rede” (CASTELLS, 2003: 109). Deve-se notar que “individualismo” nesse contexto não implica em “isolamento”. A rede permite que o indivíduo se afirme enquanto tal, em sua liberdade – liberdade de expressão e liberdade de identidade. Isso decorre da conexão isolada a uma rede ex-cêntrica. No entanto, ao fornecer um ambiente apropriado à interação comunicativa e à colaboração, a rede permite que o indivíduo se integre aos demais em torno de um propósito (comunicativo ou colaborativo) comum. A interação virtual pode ser mais que uma rede de indivíduos – e os dados disponíveis o corroboram.

1.4. A comunidade virtual

Individualismo em rede e interação colaborativa – essas características contraditórias, ao se ancorarem nas tecnologias e nas tradições históricas da rede, fornecem um ambiente para a atividade humana onde uma cultura própria pode surgir, fazendo com que novos hábitos, valores e atitudes possam se incorporar ao acervo social. Essas características, porém, não são todas que podemos elencar a propósito do ciberespaço e que podem afetar o comportamento dos seus usuários. O fato de se tratar de uma rede mundial de computadores sem centro e inteiramente avessa a controles externos, que fornece um ambiente propício à interação social e à projeção de múltiplas e diferentes identidades sociais e culturais, permite que a atividade humana ali desenvolvida transcenda fronteiras de todos os tipos, nacionais e sociais. Ao propiciar o surgimento de elementos culturais novos e possivelmente uma nova cultura (a cibercultura), mesmo as fronteiras culturais extra-ciberespaço podem ser igualmente transcendidas. Posso estar em qualquer lugar, interagindo com qualquer um, assumindo qualquer papel “cibercultural” que deseje a qualquer tempo (ou ao mesmo tempo).

Isso torna concreta a compreensão do “virtual”, pelo menos na visão de Lévy. Segundo o filósofo, “é virtual toda entidade ‘desterritorializada’, capaz de gerar diversas manifestações concretas em diferentes momentos e locais determinados, sem contudo estar ela mesma presa a um lugar ou tempo em particular” (LÉVY, 1996: 47). As aspas modalizam a “desterritorialidade” típica do virtual (e, portanto, da Internet); preferimos, porém, o neologismo *trans-territorialidade* para nos referirmos a essa característica do ciberespaço, na medida em que, com isso, enfatizamos a *telepresença* como atributo do virtual próprio da Internet. Para nós essa expressão é preferível na medida em que podemos identificar o

surgimento de “territórios virtuais” na Internet (ou ambientes “virtualmente territoriais”). Lemos (2002), em sua tese de doutorado, elenca um sem-número de tribos e ambientes distintos para a interação que, embora fortemente datados e referentes ao momento anterior à massificação da rede, servem de exemplo desses “territórios” virtuais; Castells (2003) descreve experiências européias bem sucedidas na construção de cidades on-line. Uma manifestação igualmente típica da construção (voluntária) desses territórios é o *site* Orkut, amplo sucesso principalmente entre usuários de língua portuguesa, mais particularmente aqueles conectados a partir da rede brasileira de computadores. Trata-se de um ambiente virtual destinado à criação de comunidades de interesse fortemente “territorializado”: no Orkut só se entra à convite, o que impõe uma *fronteira* virtual às comunidades ali residentes.

A “desterritorialização” a que se refere Lévy deve-se ao fato de os usuários não estarem “presos” a uns poucos territórios, como no mundo não virtual. Uma implicação evidente disso é que a Internet, ainda que seja ela mesma um “território” (o ciberespaço) e admita a construção nele de um sem-número de territórios virtuais, permite uma ampla *mobilidade territorial* a seus usuários. Ao se referir às comunidades virtuais, Castells admite que “a melhor maneira de compreendê-las é vê-las como redes de sociabilidade com geometria variável e composição cambiante, segundo a evolução dos interesses dos atores sociais” (CASTELLS, 2003: 109). Se for um fato, isso coloca em causa a noção de estabilidade geralmente associada ao conceito de “comunidade”, pelo menos se aplicada ao ciberespaço.

Infelizmente, não encontramos dados empíricos que corroborem essa visão. Um modo simples de fazê-lo seria, por exemplo, investigar a composição das salas de *chat* e das comunidades de interesse no Orkut (e equivalentes) ao longo do tempo, procurando determinar a frequência com que os mesmos usuários retornam aos mesmos ambientes de interação. Até que este tipo de pesquisa seja feita, devemos encarar a “mobilidade” dos usuários e a “efemeridade” dos territórios da Internet com *tendências* que operam na rede e que podem contribuir de algum modo para as características da cibercultura, mesmo que não possamos “provar” tais tendências.

1.5. O hipertexto e as teias do ciberespaço

Esses diversos elementos que vimos discutindo são suficientes para caracterizar a *topologia* do ciberespaço. Trata-se de um espaço aberto, curvado para fora, em constante

expansão, limitado certamente (pela quantidade finita de computadores conectados a ele) mas onde a informação pode circular infinitamente. As tendências que assinalamos operam no sentido de fazer com que a trajetória dos usuários conectados ao ciberespaço tenha alta mobilidade, caracterizando um espaço de curvatura negativa (aberto). Essa topologia foi acentuada com a incorporação da *world wide web* aos recursos da rede, no início dos anos 1990.

Estritamente falando, a *web* é a rede de computadores capazes de se interligar através do protocolo HTTP (*Hyper-Text Transfer Protocol*). Este protocolo é a viabilização prática do conceito de *hipertexto*, cunhado bem antes da Internet por Theodor Nelson e definido como “informações legíveis aos seres humanos interconectadas de modo irrestrito” (NELSON, 1965). Tal conceito foi proposto pelo pesquisador Tim Berners-Lee, em 1989, como o núcleo conceitual de um sistema distribuído de informações, modelado como uma rede de nós conectados por referências cruzadas – a *web* (ver BERNERS-LEE, 1989). Esta rede, onde os nós são documentos e as conexões, *links* para outros documentos, é constituída explicitamente com o objetivo de realizar o sonho antigo de “estender seu [dos seres humanos] intelecto através da disponibilização do seu conhecimento coletivo para cada indivíduo através dos computadores” (BERNERS-LEE & all, 2006). A idéia básica desse protocolo de comunicação é fornecer acesso fácil e descentralizado a quaisquer documentos a partir de quaisquer outros documentos. Não vamos nos alongar aqui na descrição do protocolo, já que supomos que os leitores dessa dissertação sejam em alguma medida usuários dele. Basta-nos assinalar que o modelo WWW (*World Wide Web*) representa um conjunto de documentos como uma rede de nós conectados por ligações virtuais (os *links*). Num dado documento um *link* é um índice para um outro documento relacionado arbitrariamente (pelo autor do documento) àquele de algum modo. Os projetistas concebem o *hipertexto* como um conjunto de conexões virtuais que apontam para um determinado documento. Nesse modelo, “ler” um *hiperdocumento* implica em percorrer tanto os *links* que apontam para um documento inicial arbitrário (a *home page*) quanto aqueles apontados por ele.

Embora Berners-Lee não esteja preocupado com a clareza conceitual da descrição do seu modelo, algumas distinções são aconselháveis. Enquanto um nó nessa rede virtual, um documento deve conter (ou ser apontado por) *links* para outros documentos. Chamemos um documento com *links* saindo dele ou chegando a ele um *hiperdocumento*. Nesse contexto, um *hipertexto* pode ser concebido como a rede de todos os hiperdocumentos interligados através de conexões virtuais (*links*) recíprocas. Ainda que seja teoricamente possível conectar todos os hiperdocumentos do ciberespaço num único hipertexto, na prática a rede de *links* abrange

apenas um subconjunto de todos os hiperdocumentos disponíveis na *web*. Por esse enfoque, a WWW é composta por ilhas de hipertextos não relacionados.

Do mesmo modo, um navegante deveria percorrer todos os *links* a partir de um dado hiperdocumento para considerar “lido” o hipertexto associado. Isso, evidentemente, não ocorre na prática. Diferentes usuários percorrem diferentes conexões virtuais a partir de um mesmo hiperdocumento; portanto, na navegação, eles constroem seus próprios hipertextos num sentido bem concreto. E embora os *websites* tenham pontos de entrada bem definidos (a *home page*), os hipertextos não são hierárquicos; assim, eles não “começam” numa *home page*: qualquer hiperdocumento pode ser um ponto de entrada no hipertexto. Numa navegação a partir de uma busca num *search engine*, por exemplo, o acesso aos hiperdocumentos raramente se dá a partir de uma *home page*. Isso significa que os hipertextos inauguram um novo modo de “leitura” – a navegação na *web*.

É evidente que podemos focar o hipertexto tanto do ângulo da sua produção (ou enunciação) quanto do seu consumo (ou leitura). A questão mais relevante é, porém, unificar essas visões, procurando determinar as relações entre os dois campos. Os hipertextos não se dispõem ao acaso na *web*: eles são constituídos por hiperdocumentos armazenados em *websites*. O que chamamos *website* pode ser concebido como um conjunto de hiperdocumentos produzidos e organizados sob uma identidade comum: o *site* da UERJ, o *ex-blog* do alcaide carioca, a página pessoal do usuário agora não mais anônimo, o *fanzine* dos admiradores de um conjunto de *rock*, o *site* corporativo de uma empresa, etc. – onde quer que encontremos hiperdocumentos produzidos com um propósito determinado e organizados de tal modo que formem uma identidade comum, aí encontramos algo que podemos chamar um *website*.

Ora, um hipertexto pode, ao menos teoricamente, estar inteiramente contido num único *website*, no caso de um conjunto de hiperdocumentos conectados virtualmente apenas entre si, sem nenhuma referência a hiperdocumentos de outros *websites* e para o qual nenhuma referência é feita. No entanto, no caso mais comum, os hipertextos atravessam vários *websites*, até porque foi esse o objetivo primário dos projetistas da *web* – conectar documentos mantidos em qualquer lugar. Podemos, portanto, conceber um hipertexto contido inteiramente num único *website* como um caso “degenerado” (e raro) da *web*, ainda que teoricamente possível. Até porque mesmo um hipertexto atômico como este pode, subitamente e sem o consentimento do seu responsável, atravessar a fronteira do seu *website* original: basta que seja devolvido a um usuário como resultado de uma busca. Assim, os

resultados das buscas nos chamados *search engines* configuram hipertextos temporários, criados dinamicamente como resposta a uma interação do dispositivo com um usuário.

Ora, se somos capazes de reconhecer uma identidade particular num *website*, esse reconhecimento não pode ser senão uma função discursiva, social e culturalmente inscrita. Afinal, os discursos não são produzidos no vácuo, fora das relações sociais e culturais. Assim, partimos da premissa que deve ser possível reconhecer o lugar social de onde são produzidos enunciados relacionados. Por exemplo, dados os enunciados “whites only”, “black is beautiful” e “affirmative action”, somos capazes de dizer que eles expressam diferentes pontos de vista sobre um mesmo tema; eles são enunciados de diferentes (e conflituosos) lugares sociais e culturais e estão relacionados a diferentes discursos sobre a cor da pele e suas implicações numa cultura particular. Trata-se de “vozes” discursivas, falando de pontos de vista distintos e afirmando coisas distintas.

Os hiperdocumentos podem não somente atravessar as fronteiras das identidades culturais construídas em torno dos *websites* como também as fronteiras sociais inerentes aos discursos: como qualquer texto, os hiperdocumentos podem conter enunciados produzidos por diferentes “vozes” discursivas, ainda que seu locutor seja o “mesmo”. Por outro lado, um mesmo hipertexto pode ter vários locutores (os “responsáveis” pelos *websites*). Isso significa que os hipertextos, consideradas apenas essas duas variáveis, podem ser concebidos dentro de uma escala contínua de crescente heterogeneidade discursiva. Num extremo da escala estariam os hipertextos monológicos (um único locutor, uma única voz discursiva); no outro extremo, os hipertextos polifônicos (vários locutores, várias vozes).

Há uma relação definida, ainda que não necessária, entre as duas variáveis: é de se esperar que um hipertexto *multi-site* seja também polifônico, ainda que isso não ocorra sempre. A razão é que o *link* não é apenas um *elo* físico entre os hiperdocumentos: ele pode ser visto como a *presença* material do outro hiperdocumento – e, como tal, sua ausência ou presença pode ter uma função *retórica*. Pelo que vimos sugerindo, supomos que os discursos vigentes nos hipertextos tendem a ser privilegiadamente abertos e heterogêneos: na medida em que anunciam a presença material da voz do outro, os *links* abrem os textos a outros textos; do mesmo modo, os hipertextos tendem a não ser entidades homogêneas num grau muito maior que os demais textos.

Supomos que *link* seja mais que uma citação, seja mais que uma referência: na medida em que permite (ou quase que *convida*) a presença material da voz do outro nos seus próprios termos, no seu próprio contexto e com suas próprias palavras, o *link* é, num sentido bem concreto, a palavra do outro incorporada à palavra corrente do locutor. Numa citação ou

referência, a palavra do outro é digerida e reelaborada pela minha própria palavra. Na citação com *link*, não apenas digerimos e reelaboramos a palavra do outro: nós a trazemos para dentro da nossa própria palavra, nós a incorporamos ao nosso próprio texto como uma presença material. Portanto, sugerimos que o hipertexto possa ser a heterogeneidade discursiva inerente a todo texto exponenciada a seu grau máximo.

Isso nos permite colocar em cena a função *retórica* do *link*: uma citação sem *link* no hipertexto pode ser vista como uma réplica que nega ao outro o direito à tréplica. Se nosso raciocínio é correto, no hipertexto uma citação sem *link* é simplesmente um modo de calar a voz do outro e não um meio de trazê-la à cena. Um exemplo típico pode ser visto no *site* pessoal de Olavo de Carvalho (<http://www.olavodecarvalho.org>), professor de filosofia e articulista, notável principalmente pelo seu texto raivosamente anti-comunista e pela expansão semântica que empreende com a noção de “comunista”, que parece abranger qualquer um que não pense como ele. O texto do professor é invariavelmente polêmico e acusatório, e como tal é construído em torno da palavra de outrem. No entanto, no *website* desse Emile Zola às avessas *nunca* encontramos um *link* para os *sites* dos acusados: apenas seus “gurus” e “amigos”, alunos e jornais que publicam seus artigos recebem *links*. Portanto, a região *olavodecarvalho.org* desse hipertexto, ainda que incorporando inúmeros locutores, dá a palavra somente à mesma voz discursiva.

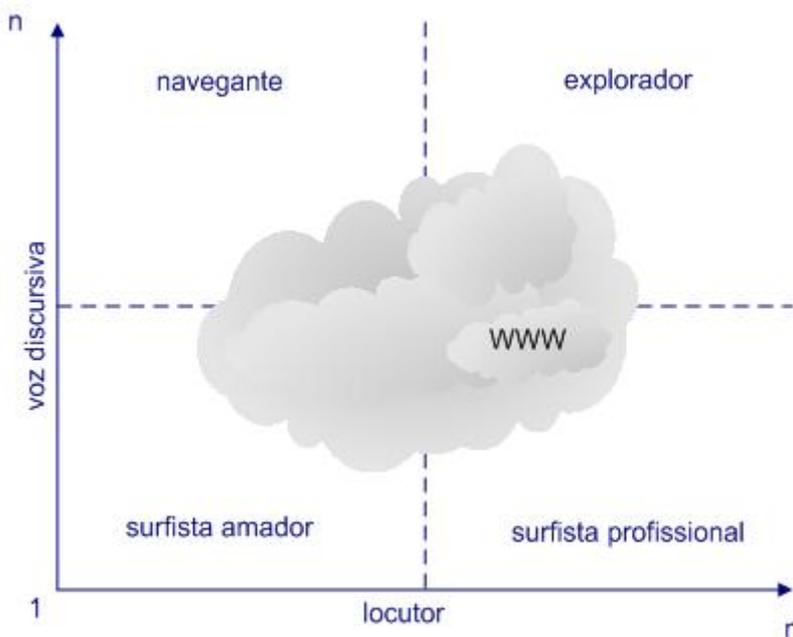


Ilustração 3: Os quadrantes mágicos do hipertexto.

A questão que se põe agora é relacionar esses tipos de hipertextos a diferentes tipos de usuários, buscando elaborar um modelo de “leitura” dos hipertextos. Suponhamos que os usuários da *web* tenham preferências bem definidas, isto é, que tomem decisões de navegação de modo racional (com um objetivo claro) e consistente (as mesmas decisões são tomadas diante dos mesmos dilemas). Suponhamos agora que um usuário determinado diante de dois hipertextos de tipos distintos (segundo o critério sugerido nos parágrafos anteriores) tenda a escolher um deles em detrimento do outro. As preferências dos usuários, portanto, tenderiam a se distribuir pelos diferentes tipos de hipertexto. Se dermos um nome a cada um desses tipos de usuários distribuídos segundo suas preferências, podemos obter uma distribuição em quadrantes como a exibida na Ilustração 3. Naturalmente, trata-se de um modelo teórico. Muito provavelmente, os usuários “reais” circulam no ciberespaço naquela região difusa que, na ilustração, chamamos WWW, suas preferências exprimindo-se mais como uma probabilidade abaixo da certeza.

1.6. As curvaturas do ciberespaço

Esse modelo de comunicação imposto pela simples existência dos *links* tem suas próprias implicações para o comportamento dos usuários. O *link* tem sua própria sedução, seu próprio convite ao clique e à navegação. Como nenhum documento pode ser considerado inteiramente consumido antes de todos os seus *links* terem sido visitados, isto é, sem que o hiperdocumento virtual associado a ele seja consumido (e isso é visualmente marcado pelo *software* de navegação, que confere cores distintas para *links* visitados e não visitados), a trajetória percorrida pelos usuários no ciberespaço tende a ser centrífuga, para fora e para longe do *site* de origem. Utilizando o vocabulário da física, podemos dizer que a “velocidade de escape” dos *sites* deve ser bastante baixa, sendo muito fácil ao usuário deixar os domínios de um *site*. A expressão *navegação* utilizada para traduzir a expressão *browsing* é bastante apropriada, exprimindo metafóricamente essa trajetória para fora e para longe da origem. Esse comportamento pode ser previsto igualmente pelas características do ciberespaço que abordamos anteriormente; no entanto, na *web* esse comportamento está sustentado diretamente na própria tecnologia que a constitui.

Nessa teia, o navegante exerce sua liberdade de um modo inteiramente novo no universo midiático: mais do que apenas selecionar que textos serão lidos, ele determina o modo pelo qual serão lidos, definindo arbitrariamente seu próprio percurso pela rede de *links*. E se pensarmos na concepção original dos projetistas, então concluiremos que na *web*, simplesmente ao escolher que *links* visitará e em que ordem, o navegante *constrói o seu próprio hipertexto*. Portanto, parece perfeitamente plausível afirmarmos, como Lemos (2002: 130), que o hipertexto só se concretiza ao interagirmos com ele, conferindo papel ativo ao leitor. Ao contrário de outros tipos de mídia, onde podemos falar inequivocamente do consumidor, na *web* é mais apropriado falarmos do co-produtor, enfatizando o papel ativo na produção de sentido do texto desempenhado pelo navegante.

A escolha do navegante pode ser simplesmente a do “bom senso”, consumindo os *links* a medida em que surgem e retornando imediatamente ao texto de origem, que funciona aqui como um condutor da navegação. Pode, contudo, estilhaçar-se em inúmeras e contraditórias direções, saltando entre os *links* em qualquer ordem, não retornando mesmo ao texto condutor, agora apenas *pretexto* para a navegação. A liberdade de escolha do navegante pode – e frequentemente o faz – contradizer aquilo que Deleuze chama as propriedades do bom senso: “a afirmação de uma só direção; a determinação desta direção como indo do mais diferenciado ao menos diferenciado, do singular ao regular, do notável ao ordinário; a orientação da flecha do tempo do passado ao futuro, de acordo com esta determinação” (DELEUZE, 1974: 79). Essa nos parece uma decorrência bastante plausível do *efeito web*.

Ora, o que chamamos *efeito web* deve ser facilmente observável; se ele é mais que uma especulação divertida, deve afetar os hábitos de navegação dos usuários na Internet de algum modo suficientemente evidente para que possamos corroborar a hipótese da sua existência. Ora, é possível prever que uma trajetória centrífuga de navegação se reflita nos hábitos dos usuários como uma variedade grande de *sites* visitados e em pouco tempo de permanência por página navegada. A primeira “previsão” é bastante óbvia; já a segunda requer alguns esclarecimentos. Deve-se considerar que o tempo de permanência na Internet é, necessariamente, limitado. Segundo o Comitê Gestor, 83% dos usuários da Internet no Brasil passam conectados no máximo 10 horas por semana. Isto significa que a maioria dos usuários dedica apenas 9% do seu tempo útil à Internet. Ora, se o consumo de hiperdocumentos requer que a maioria senão todos os seus *links* sejam percorridos e cada *link* aponta para um hiperdocumento, segue-se que uma navegação padrão deve obter uma grande quantidade de documentos; se o tempo dedicado a esta atividade é limitado, segue-se igualmente que o

tempo dedicado a cada documento deve ser necessariamente pequeno; caso contrário, os hipertextos não são consumidos.

A proposição que elaboramos a propósito dos hábitos de navegação dos usuários é ainda muito imprecisa para verificação empírica; afinal, qual o significado preciso de “variedade grande” e “pouco”? Tais conceitos, além de vagos, são relativos; o que poderia então ser um padrão preciso de medida? Podemos fazer algumas suposições simples. Por exemplo, uma página do texto desta dissertação tem em média 350 palavras e pode ser lida por um leitor suficientemente familiarizado com seu vocabulário em 90 segundos em média (pelo menos o autor da dissertação é capaz de fazê-lo). Embora não haja uma relação direta entre uma página desta dissertação e uma página qualquer da Internet, este é um ponto de partida tão bom quanto qualquer outro. Assim, diremos que qualquer tempo de *consumo* de uma página inferior a esse tempo pode ser considerado “pouco”. Estamos utilizando aqui a expressão “consumo” em lugar de “leitura” porque o simples fato de carregar um documento obtido na Internet no seu navegador não implica sua leitura pelo usuário – e é mais fácil medir a simples obtenção de um documento que sua leitura.

Por outro lado, as atividades realizadas na Internet no Brasil ilustradas na Tabela 2 podem nos servir como ponto de partida para definirmos mais concretamente o que seja uma “grande variedade” de *sites* visitados. Tomemos as classes de atividades mais expressivamente realizadas (com mais 70% do universo de usuários) e suponhamos que cada atividade seja realizada a partir da visita a um e somente um único *site* diferente; suponhamos também que, para cada classe de atividade, os usuários visitem a mesma quantidade de *sites* e que todos os usuários realizem sempre *todas* as atividades. Novamente, trata-se de suposições arbitrárias e provavelmente irrealistas, mas é necessário um ponto de partida qualquer para o nosso raciocínio. Isto significa que um usuário fiel aos *sites* que visita (centrípeto) visitaria apenas 24 domínios⁴ diferentes durante seu tempo de permanência no ciberespaço. Assim, qualquer quantidade de domínios visitados maior do que esta pode ser considerada “uma

⁴ Como vimos em nota anterior, o conceito de “domínio” pode ser equivalente, grosso modo, ao de *site*, mas não exatamente. Um mesmo domínio pode abrigar vários *sites* (por exemplo, um serviço que permite aos usuários criar e manter *websites* seria um caso desses). Para nosso raciocínio, porém, a equivalência é suficiente – e novamente porque é mais fácil medir as visitas a diferentes domínios que a diferentes *sites* (num mesmo domínio).

grande variedade”. Com isso, a previsão que fizemos pode ser precisamente aferida por algumas medições bastante simples.

Felizmente, não precisamos gastar muito esforço para efetuar essas medições: elas já são realizadas em dez países, dentre eles o Brasil, pela empresa de pesquisas Nielsen//NetRatings (<http://www.netratings.com>; ver WEB, 2006). Segundo a instituição, os usuários desses países visitam, em média, 69 domínios diferentes por mês; além disso, eles consomem em média 42 segundos em cada documento obtido durante sua navegação, valores compatíveis com as previsões feitas, desde que admitidas convenientemente nossas suposições sobre o significado desses valores relativos. Os valores individuais de cada parâmetro podem ser vistos na Ilustração . Note-se que, por simplicidade, utilizamos um único gráfico de linhas, ainda que os valores do eixo y para cada linha tenham significados diferentes em cada caso (domínios num caso e segundos em outro). Um outro comentário preliminar refere-se ao fato de, ainda uma vez, a utilização de um único gráfico para dois valores distintos não teve como intenção correlacionar tais valores.

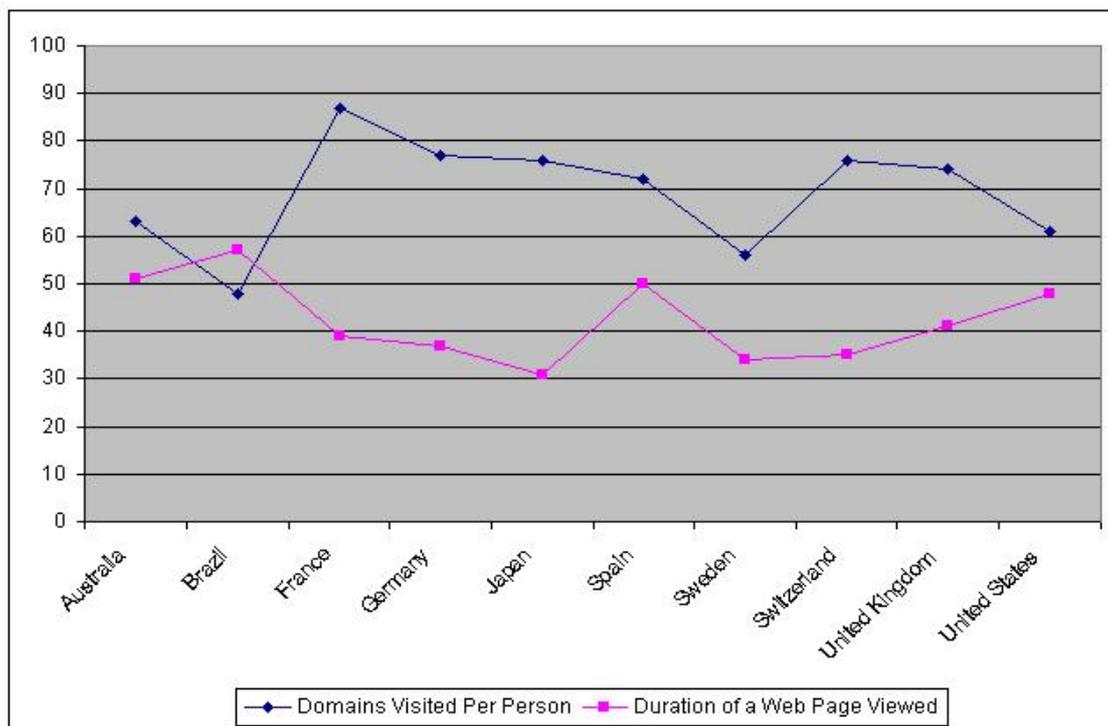


Ilustração 4: Domínios visitados e tempo de *page view* na web (Fonte: Net//Ratings).

A observação mais interessante que podemos fazer sobre o gráfico refere-se aos dados relativos ao público brasileiro. Os valores obtidos para este público estão contidos no *pico* da linha de duração do *page view* e no *vale* da linha de domínios diferentes visitados. Isto quer dizer que, dos países pesquisados, o Brasil é notável pelo fato de seu público visitar a *menor*

quantidade de domínios diferentes e gastar a *maior* quantidade de tempo em cada página visitada. Isto significa que o usuário brasileiro realiza a navegação menos centrífuga dentre os países pesquisados. Parece certo que agem aqui forças centrípetas, contraditórias com a tendência básica esperada para o *efeito web* descrito anteriormente. Aparentemente, a topologia da região do ciberespaço freqüentada pelos usuários conectados a partir da rede brasileira é *curvada positivamente* (para dentro), o que nos leva a concluir que aquela topologia não é uniforme em todas as direções.

Essa hipótese é corroborada pela análise sumária da Tabela 1, citada no capítulo anterior e referente a pesquisa publicada pela organização Ibope (RELATÓRIO, 2005). Como assinalamos ali, a pesquisa sugeria que o tempo consumido na Internet é fortemente concentrado entre os usuários brasileiros quando comparados aos americanos e espanhóis: as cinco principais categorias pesquisadas correspondem a 53,2% do tempo total de consumo dos usuários brasileiros, enquanto as cinco principais categorias nos Estados Unidos e Espanha correspondem a apenas 30,4% e 39,8%, respectivamente. O uso concentrado encontrado na pesquisa sugere hábitos de navegação centrípetos

É possível elaborarmos hipótese de que pelo menos parte da resposta a esse problema possa ser obtida pela análise do conteúdo dos domínios mais visitados pelos usuários brasileiros. Infelizmente, a NetRatings não publica gratuitamente a pesquisa de audiência para o Brasil, provavelmente por força da sua associação aqui com o Ibope, que vende essas pesquisas. Ao longo da nossa investigação, tivemos acesso somente à avaliação de audiência de setembro de 2003 (TOP 10, 2004). Supondo que a tendência básica ali expressa não tenha se modificado significativamente, encontramos listados os principais gigantes no fornecimento de acesso e conteúdo (iG, UOL, Terra, MSN e Globo), conhecidos aqui como portais de interesse geral. A pesquisa, cujo resultado é visto na Ilustração , indica que, entre os dez *sites* de maior audiência, 67% dos acessos eram feitos aos portais de interesse geral. É provável que a tendência tenha se mantido razoavelmente inalterada. Como vimos no capítulo anterior, a pesquisa 2005 já citada determinou que 10,5% do tempo gasto na Internet pelos usuários brasileiros é consumido nos portais de interesse geral.

Esses dados sugerem que a análise e a compreensão dos portais de interesse geral é decisiva para a correta avaliação dos hábitos de navegação dos usuários conectados ao ciberespaço a partir da rede brasileira. Como vimos no capítulo passado, ao sumarmos os portais de interesse geral, esses empreendimentos, para serem bem sucedidos, devem ser capazes de neutralizar aquelas tendências centrífugas que assinalamos ao discutir as características do ciberespaço. Eles devem ser capazes de alterar os hábitos de navegação

“naturais” (no sentido de hábitos mais coerentes com as características do ciberespaço) dos usuários da Internet, curvando para dentro de si mesmo um ciberespaço originalmente aberto e em expansão. Nesse sentido, sob o ponto de vista das características do ciberespaço que vimos enfatizando, os portais representam uma *anomalia* que requer interpretação. Nossa pesquisa nasceu da constatação dessas anomalias no ciberespaço e se desenvolveu buscando interpretar e explicar alguns de seus mecanismos (discursivos). A pergunta central que nos move pode, então, ser resumida da seguinte forma: *como* os portais de interesse geral distorcem a topologia do ciberespaço, curvando-o nas suas vizinhanças?

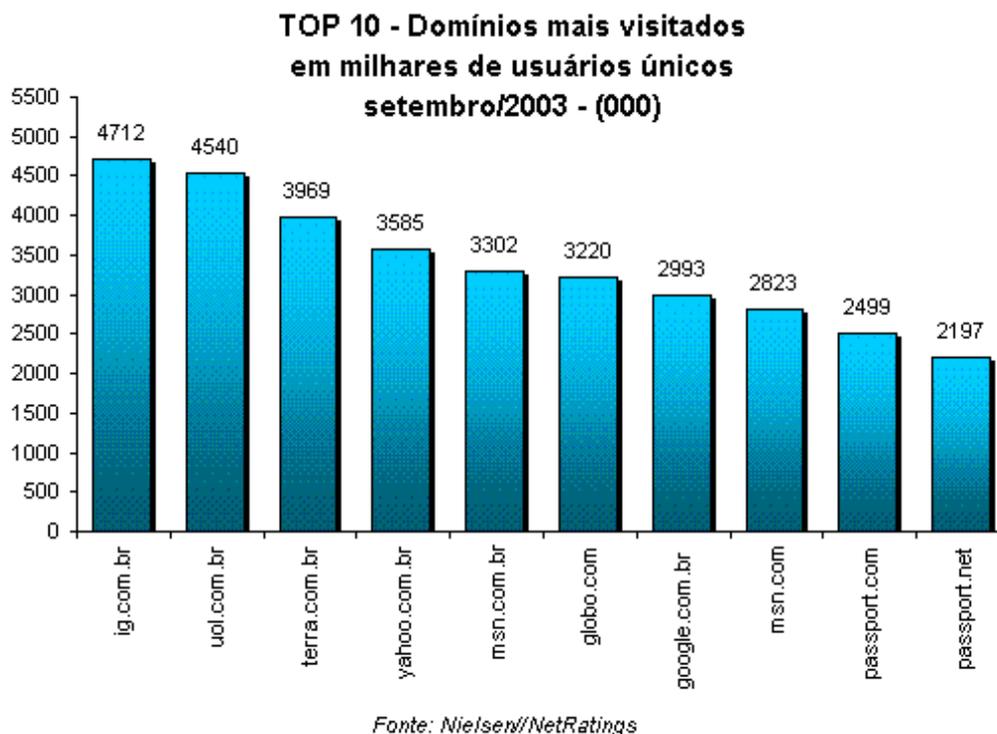


Ilustração 5: Domínios mais visitados pelos usuários brasileiros.

Como vimos, a metáfora da curvatura do ciberespaço refere-se ao comportamento navegacional dos usuários, aparentemente anômalo em face do comportamento esperado em vista das características da rede. No entanto, é bem possível que essa tendência ao fechamento da navegação se expresse, ao nível dos discursos que se articulam em seus documentos, em equivalente fechamento. Na seção anterior discutimos em parte esse aspecto. Hábitos centrípetos de navegação sugerem um hiperleitor surfista, ou pelo menos um hiperleitor que tende a restringir-se aos quadrantes a oeste do gráfico exibido na Ilustração 3. É possível, portanto, que o fechamento no hipertexto produzido por esse esforço em manter os usuários

orbitando num mesmo domínio conduza a equivalente fechamento dos múltiplos discursos que circulam na rede: se não são necessariamente caladas, as diferentes vozes desses discursos devem retornar às profundezas do texto. As fraturas do hipertexto são costuradas nos portais de conteúdo – tal a premissa na qual baseamos nossas pesquisas.

A nosso ver, esse efeito, se de fato se produzir nos portais, decorreria do próprio projeto comercial desses portais, que se traduz no que Abraham Moles, a propósito do rádio em meados do século passado, se referiu como a doutrina demagógica dos publicistas: “como alcançar e manter o maior índice de audiência durante o maior período possível? (...) [Através da] maior satisfação do maior número [possível de pessoas]” (MOLES, 1982: 73 ss.). Para que esse projeto seja viável, o meio deve fornecer apenas satisfação, resposta à busca efetuada e não problemas e questionamentos, prazer e não dor, lazer e não aventura. Satisfação garantida (como os portais não cobram pelo fornecimento do conteúdo, a insatisfação não implica na devolução do dinheiro) – tal o lema dos portais. Ou no dizer de Orlando: “um portal procura ser exatamente um ponto de partida para os ‘surfistas’ da web, reunindo conteúdo em quantidade para tentar agradar à diversidade de consumidores de informação” (ORLANDO, 2001: 86).

É conveniente que a expressão *surfing*, regularmente utilizada em lugar de “navegação”, seja utilizada no contexto dos portais... Em língua portuguesa, num acidente de tradução conveniente para nós, as duas expressões podem se opor. Enquanto o “navegante” atravessa os mares, o “surfista” limita-se a uma única praia. Um desbrava o desconhecido, arrisca-se para além dos domínios familiares; o outro recusa o perigo da novidade; num a aventura ousada em direção a mares desconhecidos, no outro o gratificante passeio por praias conhecidas. Duas atitudes de utilização da Internet podem ser aqui reconhecidas – e os portais, para serem bem sucedidos, privilegiam uma. As curvaturas no ciberespaço produzidas por eventuais estratégias destinadas a obter a atitude surfista do usuário devem ter implicações generalizadas, afetando desde os hábitos de navegação até a natureza e características dos hiperdocumentos produzidos e consumidos no espaço desses portais. É nesse ponto que se inicia a nossa investigação.

2. GÊNEROS DO DISCURSO – ALGUMAS ÂNCORAS TEÓRICAS

Ao delimitarmos o nosso espaço de pesquisa à retórica dos portais de interesse geral, estamos implicitamente admitindo que é possível tratá-los como uma categoria. E se isso é feito no interior de uma disciplina como a Lingüística, estamos igualmente assumindo que se trata de categoria de alguma forma relevante para os problemas da(s) linguagem(ns) em uso. Mas o que nos leva a assumir que tais *sites* possam ser legitimamente tratados como uma categoria especial de artefatos de linguagem? Posto de outra forma: ao categorizar implicitamente os portais de interesse geral como um artefato de linguagem distinto, estamos assumindo que é possível relacioná-los de alguma forma a outras categorias de artefatos comunicativos passíveis de ser encontrados no ciberespaço, como o *e-mail*, os *chats*, os *blogs*, entre tantos. Como estes últimos artefatos têm sido abordados diretamente nas pesquisas recentes sobre gêneros do discurso, estamos assumindo igualmente que os portais seriam uma categoria deste tipo.

Tal assunção, se posta sem maiores considerações, pode ser considerada ilegítima, no mínimo porque é polêmica. Já assinalamos a objeção de Marcuschi, que não admite caracterizar a *home page* como um exemplar de um dos novos gêneros digitais. Ainda que o pesquisador esteja aberto a contra-argumentos em vista de se tratar de território largamente inexplorado, não incorpora essa categoria no seu inventário de gêneros digitais. Ora, um portal é uma especialização da categoria dos *sites* (ou *home pages*) Internet, estando, portanto, incluído na rejeição de Marcuschi. Isto significa que o problema requer resolução preliminar.

Ser ou não ser um *gênero* do discurso – por que essa questão deve ser considerada importante? Afinal, todos concordam que a classificação é uma operação de conhecimento notoriamente superficial, prestando-se mais como ponto de partida que como de chegada. Ao encararmos objetos como membros de uma categoria, estamos assumindo que todo o conjunto possui pelo menos algumas propriedades comuns. Isto nos proporciona uma conveniente economia descritiva, ao nos permitir estudar apenas alguns exemplares e, com isso, ampliar nosso conhecimento sobre o próprio conjunto do qual fazem parte. No entanto, como se verá mais adiante, quando questionarmos a existência dessas “propriedades comuns” a um gênero, essa premissa é frágil, não sendo aconselhável utilizá-la como justificativa de relevância.

O mesmo não ocorre ao refletirmos sobre as vantagens de abordarmos os portais como gêneros *do discurso*, isto é, como uma categoria especial de *artefatos de linguagem*. É essa abordagem que nos permite falar não metaforicamente de uma retórica dos portais. Se essa operação conceitual não for possível, ao abordarmos os portais só seremos capazes de reconhecer um empreendimento comercial, os serviços que presta aos seus usuários, as estratégias de *marketing*, etc. – questões certamente relevantes aos estudos sociais, mas apenas de interesse marginal para a Linguística (do discurso). Portanto, a questão de que falávamos precisa ser respondida – no mínimo para que justifiquemos abordar os portais do ponto de vista dos discursos que nele operam. Desse modo, as próximas seções serão dedicadas a determinar se podemos legitimamente tratar os portais como um gênero (emergente) do discurso.

2.1. O problema – teórico – dos gêneros do discurso

Gênero, hipergênero ou simplesmente suporte para gêneros discursivos – qualquer que seja a resposta que busquemos, ela só pode ser dada no interior de uma teoria dos gêneros do discurso. Afinal, uma teoria dessas deve incluir no mínimo uma definição ou pelo menos uma concepção abrangente do que sejam gêneros do discurso, de tal modo que nos permita reconhecer novos artefatos e eventos como pertencentes a essa categoria geral. Nos últimos anos temos assistido um amplo interesse no tema no interior dos estudos da Linguagem, a ponto de já podermos identificar diferentes tradições na pesquisa. Assim, é de se esperar que o problema possa ser satisfatoriamente resolvido com o instrumental teórico disponível.

O problema é que a existência em si da polêmica, mesmo que marginal nas pesquisas sobre gêneros digitais, sinaliza que a questão talvez não seja tão simples assim. Ela nos sugere duas possibilidades. A primeira é decorrente de podermos identificar diferentes tradições nessas pesquisas: isso talvez signifique que diferentes teorias estejam em jogo – e diferentes teorias podem simplesmente nos conduzir a diferentes respostas para o problema. A segunda é decorrente do fato de todas as tradições estarem estritamente ancoradas no estudo das práticas de linguagem verbal – e tais práticas talvez sejam insuficientes para, sozinhas, darem conta de modo abrangente das práticas comunicativas do ciberespaço.

Em sua clássica proposição ao CERN por um sistema global de gerenciamento de informações distribuídas, Berners-Lee (1989) assinala que o termo “hipertexto”, utilizado para se referir aos esforços por sistemas capazes de disponibilizar informações relacionadas a documentos publicados, envolvia duas idéias diferentes. A primeira delas, o próprio conceito de “hipertexto”, já discutido no capítulo anterior. A outra idéia, cuja implementação técnica é “independente e largamente uma questão de tecnologia e tempo, é a de documentos multimídia, que incluem gráficos, voz e vídeo” (BERNERS-LEE, 1989). Tais tipos de documentos receberam o nome de batismo de *hipermídia*, dado pelo autor.

Se naquele momento a idéia poderia carecer de viabilidade prática, a evolução da tecnologia e seu conseqüente barateamento tem feito crescer exponencialmente a largura de banda e o poder computacional disponível para os usuários domésticos da Internet. Embora o texto em língua natural (escrita) sustente a comunicação na Internet – no mínimo porque é a língua natural que sustenta a maior parte da comunicação humana, – a incidência de material semiótico que incorpora recursos de hipermídia – que só deve crescer nos próximos anos, com o barateamento e a popularização da chamada “banda larga” – é tal que nos leva a concluir que o fenômeno pode ser indispensável a uma compreensão abrangente do ciberespaço. Se pudermos aplicar essas ferramentas conceituais ao problema dos gêneros (do discurso) digitais, possivelmente a questão de conceituar os portais de interesse geral no interior da teoria (dos gêneros) seja satisfatoriamente resolvido. Trata-se, então, de avaliar as várias tradições de pesquisa dos gêneros do discurso e seus referenciais teóricos em busca dessas ferramentas conceituais.

É conveniente assinalar, por fim, que buscaremos essa teoria dos gêneros (ou essa fundamentação teórica para o problema) não num único autor ou tradição de pesquisa, mas na costura de idéias – possivelmente distintas – que se possam complementar. O risco evidente é tentar conciliar o inconciliável (e possivelmente fracassar), particularmente em disciplina acadêmica dividida em tendências filosóficas muitas vezes antagônicas, onde as idéias se contradizem mais que se complementam. Essa tarefa, conquanto arriscada, pode ser necessária, no mínimo por força da suspeita que nos assalta, a saber: que as várias e diferentes abordagens do problema dos gêneros do discurso, se tomadas isoladamente, talvez sejam insuficientes para uma compreensão abrangente do nosso *corpus* em particular.

Nosso esforço aqui será menos conciliar as diversas teorias originadas nas diferentes tradições assinaladas e mais extrair delas algumas idéias fundamentais, de tal forma que possam ser aplicadas produtivamente ao problema geral dos gêneros digitais e à melhor compreensão do nosso *corpus* em particular. Tais idéias serão buscadas, em primeiro lugar, na

teoria dos gêneros do discurso esboçada por Bakhtin (e não na de seus intérpretes) e, em segundo lugar, na abordagem de Swales. Por fim, abordaremos um conjunto de idéias que *não* se originam numa teoria dos gêneros, mas que incluem e podem iluminar o problema dos textos de hipermídia: as idéias de Kress e Van Leeuwen. Ao discutir essas idéias, nossa ênfase recairá menos na concepção de gêneros e mais na noção de multimodalidade, que consideramos particularmente relevante para a compreensão do nosso *corpus* em particular.

2.1.1. O Locutor, o Interlocutor e o Outro: gêneros do discurso em Bakhtin

Para Mikhail Bakhtin, “cada campo de utilização da língua elabora seus tipos relativamente estáveis de enunciados, os quais denominamos gêneros do discurso” (BAKHTIN, 2003: 282). Isto significa que os gêneros do discurso aparecem para os locutores como um repertório de formas enunciativas prontas, disponíveis para uso nos diferentes contextos da atividade social. Tais *protótipos* determinariam o conteúdo temático (sobre o que se fala), a construção composicional (como se fala) e o estilo (como o locutor se posiciona em face do seu objeto e frente a seu interlocutor) dos enunciados concretos; em outras palavras, os gêneros do discurso afetam diretamente o processo da enunciação. Dito de outro modo, “falamos apenas através de determinados gêneros do discurso, isto é, todos os nossos enunciados possuem formas relativamente estáveis e típicas de construção do todo” (BAKHTIN, 2003: 282).

Definição simples e direta, ela suscita, porém, inúmeras indagações, a começar pela noção de enunciado. Rodrigues (2004) assinala que “é preciso observar que a noção de enunciado de Bakhtin difere da de outras teorias, como a da Lingüística Textual (o texto como conjunto coerente de enunciados) e a da Teoria da Enunciação (enunciado como a manifestação concreta da frase)”. Ao nos indicar a heterogeneidade do conceito, essa simples observação nos obriga a iniciar a avaliação da concepção bakhtiniana pela discussão dessa noção.

Bakhtin assinala que é o enunciado – e não a frase – a unidade real da comunicação discursiva, visto que “o discurso só pode existir na forma de enunciações concretas de determinados falantes, sujeitos do discurso” (BAKHTIN, 2003: 274). Nessa concepção o enunciado começa quando o locutor toma a palavra de outrem e termina quando ele cede a

palavra a outrem; o enunciado é, portanto, delimitado pela alternância dos sujeitos do discurso. À esquerda, o enunciado é sempre limitado por todos os enunciados precedentes produzidos no interior de um determinado campo da atividade humana – porque, insiste Bakhtin, o falante não é um Adão que toma a palavra pela primeira vez num mundo sem discurso e por isso seu “enunciado está voltado não só para seu objeto mas também para os discursos do outro sobre ele” (BAKHTIN, 2003: 300). Isso significa que um enunciado é sempre uma resposta, um diálogo, uma interlocução com outros enunciados que o antecedem. À sua direita, o enunciado está igualmente limitado pelo enunciado do outro, porque todo enunciado exige resposta, mesmo que mediata, mesmo que silenciosa, na forma de fruição estética ou apreciação; ele é construído levando em consideração essa atitude responsiva do interlocutor. Portanto, um enunciado é concluído quando o interlocutor estiver apto a tomar a palavra e responder a ele.

Portanto, a primeira observação que podemos fazer sobre a concepção de gênero do discurso em Bakhtin é que ela está diretamente ancorada na comunicação. Cabe observar apenas que a comunicação em Bakhtin não envolve apenas a transmissão de mensagens entre um emissor ativo e um receptor passivo. A emissão e a recepção da mensagem dão conta apenas de um aspecto da comunicação humana. Seria mais preciso utilizar os termos locutor e interlocutor em lugar de emissor e receptor, no entendimento que o interlocutor é ele mesmo um outro locutor, pronto a tomar a palavra e enunciar seu próprio discurso. Em Bakhtin, ambos os extremos da comunicação são locutores ativos. Daí muitos pesquisadores da tradição bakhtiniana preferirem utilizar interação em lugar de comunicação.

A concepção bakhtiniana de gênero do discurso retoma, em alguma medida, a melhor tradição aristotélica ao concebê-lo como um princípio organizador do discurso orientado a uma audiência específica. Vai mais longe, é claro, entendendo que todo discurso é orientado à audiência, da conversa cotidiana ao poema lírico. Com isso a noção se alarga para fora do domínio exclusivo da literatura e da retórica, abarcando a totalidade das atividades – discursivas – humanas, onde as manifestações artísticas fazem parte de um domínio particular. Retomaremos a questão mais adiante, precisando melhor o papel da audiência (do interlocutor) na concepção bakhtiniana.

O segundo aspecto que queremos ressaltar na concepção de gênero do discurso em Bakhtin é que ela rejeita a fala (a língua em uso) como atividade puramente individual, fruto do livre arbítrio do falante, concepção típica da tradição saussuriana. São os “campos de utilização da língua” que produzem tais formas prontas e são elas que determinam os tipos de enunciados estilísticos, temáticos e composicionais através dos quais os falantes produzem

seus próprios enunciados. Embora o texto de Bakhtin não esclareça diretamente o que sejam tais “campos” – salvo indicar que são compostos por uma determinada função e implicar em determinadas condições de comunicação – sua própria concepção de uso da linguagem como diálogo nos autoriza a entender tais campos como produto histórico das interações discursivas dos falantes.

Ao propor os gêneros do discurso como determinantes da enunciação, parece-nos adequado dizer que Bakhtin se contrapõe de saída à concepção de que o discurso seja um espaço em cujo centro se fixa um sujeito. Ele – o sujeito – não fala senão através de outras falas – os enunciados típicos do gênero, disponíveis antes do seu enunciado e produzidos pela interação da totalidade dos falantes. Em outras palavras, o locutor fala através do discurso de outrem: “todo enunciado concreto é um elo na cadeia da comunicação discursiva de um determinado campo” (BAKHTIN, 2003: 296).

Podemos dizer que o lugar ocupado pelo sujeito no discurso é, então, o lugar vazio ocupado pelo outro no processo de enunciação. Isso ocorre simplesmente porque o meu discurso é o discurso de outro, ou mais precisamente de todos os outros que me precederam, que me sucederão e com quem dialogo. De alguma forma e num sentido bem concreto, eu não falo – sou falado... Essa interpretação é autorizada – ainda que com reservas, pois se trata de tradução de segunda mão – pelo confronto com a concepção de linguagem exposta em *Marxismo e filosofia da linguagem*: “não é a atividade mental que organiza a expressão, mas, ao contrário, é a expressão que organiza a atividade mental, que a modela e determina sua orientação” (BAKHTIN, 1979: 98). Ou mais precisamente: “a situação social mais imediata e o meio social mais amplo determinam (...) a estrutura da enunciação” (BAKHTIN, 1979: 99).

O próprio estilo – “o elemento expressivo, isto é, a relação subjetiva emocionalmente valorativa do falante com o conteúdo do objeto e do sentido do seu enunciado” (BAKHTIN, 2003: 289) – é determinado por esse(s) discurso(s) de outrem. Os estilos – para alguns o lugar privilegiado da individualidade – são eles mesmos determinados pelas formas típicas dos discursos, dados de antemão e produzidos no complexo processo de interação (discursiva) social: “todo estilo está indissolivelmente ligado ao enunciado e às formas típicas de enunciados, ou seja, aos gêneros do discurso” (BAKHTIN, 2003: 265). Essa proposição é inteiramente coerente com a obra assinada por Voloshinov, onde notamos que “a personalidade que se exprime, apreendida, por assim dizer, do interior, revela-se um produto total da inter-relação social” (BAKHTIN, 1979: 103).

No entanto, o ensaio que utilizamos como fonte não autoriza que se conceba o indivíduo como exercendo exclusivamente um papel passivo em face de forças incontroláveis. Se os

gêneros do discurso são dados como protótipos de enunciados disponíveis para os falantes, segue-se que o ato de enunciação principia pela seleção de determinada forma típica apropriada às tarefas postas ao “sujeito do discurso (ou autor) centradas no objeto e no sentido” (BAKHTIN, 2003: 289). E seleção implica, em alguma medida, vontade (discursiva). Há liberdade de escolha na enunciação: preso à totalidade dos discursos que o antecedem no seu domínio de atividade, o sujeito ainda assim escolhe, exerce seu arbítrio, ele dialoga em permanente tensão com os seus outros (que ainda assim são os outros do *seu* discurso).

Essa liberdade de escolha está ancorada diretamente no campo da comunicação discursiva ao qual se associam os gêneros disponíveis para uso nas situações concretas de comunicação. Bakhtin assinala que determinados gêneros afetam mais diretamente a natureza da enunciação que outros; em particular, os gêneros mais simples, associados a campos de atividade mais padronizados (digamos, a réplica do diálogo, toda sorte de correspondências oficiais, etc.), afetam mais a individualidade dos enunciados que outros (ver BAKHTIN, 2003: 265 e seguintes). Nos gêneros complexos, que “incorporam e reelaboram diversos gêneros primários (simples) que se formaram nas condições da comunicação discursiva imediata” (BAKHTIN, 2003: 263), a individualidade e especificidade – numa palavra, a criatividade – do enunciado em face das formas prontas é mais acentuada. Os casos limites estão, certamente, nos vários gêneros literários, tão importantes para o pensador russo.

No entanto, o ensaio consultado nos autoriza a dizer que, mesmo naqueles campos onde a liberdade e o arbítrio são mínimos – os campos padronizados da comunicação – eles, ainda assim, se exercem. Lembremos que as formas típicas, esses enunciados prototípicos são considerados relativamente estáveis, isto é, instáveis em alguma medida, ainda que a expressão de Bakhtin enfatize sua estabilidade. Ora, se o autor não nos esclarece, neste momento, a natureza dessa instabilidade – e, portanto, de sua inerente tendência à mudança – ela deve estar, inevitavelmente, associada ao caráter dialogal (melhor dizendo, dialógico) do discurso. A instabilidade e a mudança nas formas prontas do enunciado só podem estar associadas à interação dos sujeitos no discurso, onde estes exercem – em que grau pouco importa nesse momento – sua liberdade de escolha. Essa interpretação pode ser validada pela seguinte afirmação, extraída da obra que utilizamos para confrontação: “a língua constitui um processo de evolução ininterrupto, que se realiza através da interação verbal social dos locutores” (BAKHTIN, 1979: 113).

Desse modo, podemos dizer que o papel do sujeito do discurso (ou talvez do sujeito *no* discurso, para ser mais preciso) em Bakhtin é, no mínimo, ambíguo: ele tanto fala quanto é

falado; seu discurso é o discurso de *outrem*, mas ainda assim é *seu*. Ele não apenas seleciona os protótipos segundo os quais seus enunciados serão produzidos como também digere e reelabora ativamente tais formas prontas. Falamos apenas através de certos gêneros do discurso, é certo; mas esta fala é mais do que apenas discurso de *outrem* – é diálogo ativo e responsivo com todos os enunciados de um dado domínio (discursivo), no tempo e no espaço. O gênero impõe, mas o sujeito dispõe – e, ao dispor o imposto, torna-o instável, desestabilizando igualmente qualquer visão simplista que possamos ter do lugar – ou dos lugares – ocupado pela categoria de sujeito do discurso em Bakhtin. Em Bakhtin, o sujeito não é, ele está posicionado na cadeia de todos os enunciados – e sujeitos – que o precedem e todos os enunciados que a ele se sucedem. Ele é apenas um elo, um nó na cadeia de enunciações:

Cada enunciado é pleno de ecos e ressonâncias de outros enunciados com os quais está ligado pela identidade da esfera de comunicação discursiva. Cada enunciado deve ser visto antes de tudo como uma resposta aos enunciados precedentes de um determinado campo (BAKHTIN, 2003: 297).

Todo enunciado é resposta a outros enunciados – tal a lição bakhtiniana. E mais, todo enunciado se orienta para a sua própria resposta: “o traço essencial (constitutivo) do enunciado é o seu direcionamento a alguém, o seu endereçamento” (BAKHTIN, 2003: 301). Se é resposta, também requer uma resposta: a comunicação – mais que transferência, transporte de significações – é interação, onde o enunciado de um se limita pelo enunciado de outro, onde enunciar é passar a palavra a alguém, onde a “compreensão da fala viva, do enunciado vivo é de natureza ativamente responsiva (BAKHTIN, 2003: 271). Essa concepção dialógica de linguagem está na raiz do entendimento de que uma concepção típica do destinatário do enunciado é determinante de cada gênero do discurso.

A estabilidade – relativa, mas sempre estabilidade – das formas típicas do enunciado dá conta também de outro aspecto relevante do pensamento de Bakhtin: na fala – e não apenas na língua – encontramos igualmente regularidade e organização, muito mais do que desordem e caos individual. Mais precisamente, Bakhtin concebe os gêneros do discurso como um dos elementos preponderantes desse princípio de ordem: “os gêneros do discurso organizam o nosso discurso quase da mesma forma que o organizam as formas gramaticais (sintáticas)” (BAKHTIN, 2003: 283). No presente contexto, essa afirmação, conquanto fugaz, serve de indicação para o fato de podermos interpretar os gêneros do discurso em Bakhtin como mais que simplesmente um repertório de enunciados convencionais, mas como um conjunto de convenções de enunciação. Esta é uma interpretação um tanto arbitrária, é certo, já que há muito pouca evidência textual em seu favor; no entanto, é inteiramente possível. Poderíamos, então, ainda no contexto de Bakhtin, admitir que os gêneros do discurso sejam tanto formas

prontas de enunciados convencionais quanto convenções que dirigem (organizam) o uso dos enunciados concretos dos falantes.

Por que nos parece importante essa interpretação? Simplesmente porque queremos determinar se o pensamento de Bakhtin é uma boa base para raciocinarmos não apenas sobre os gêneros do discurso verbal como também sobre os gêneros semióticos em geral ou se, ao contrário, precisaremos abandonar suas importantes contribuições ao discutir situações de comunicação humana típicas do mundo digital. Como vimos, ainda que as línguas naturais (particularmente sua codificação em escrita) sejam centrais no ciberespaço, outras semioses ganham relevância cada vez maior na medida em que a tecnologia banaliza eventos de comunicação híbridos, baseados na hipermídia. Em se tratando de imagens e sons, formas prontas de enunciados podem não ser facilmente perceptíveis; o mesmo não ocorre com o que poderíamos chamar formas prontas da enunciação, incluindo num único olhar normas e convenções, atitudes comunicativas e estereótipos, tanto quanto enunciados (em sentido amplo) típicos.

Trata-se, portanto, de determinar se os novos domínios do ciberespaço se deixam interpretar pelas categorias já hoje convencionais do pensamento bakhtiniano ou se, ao contrário, precisaremos de conceitos e modos de pensar inteiramente novos. A nosso ver, se os argumentos apresentados são convincentes, isso não deve ser necessário. Basta-nos conceber os gêneros dos discursos (e não somente do discurso) como normas e convenções (ou protótipos de enunciações) que organizam a enunciação concreta para que isso seja possível. E isso está ainda dentro do horizonte conceitual de Bakhtin.

Mas retornemos ao problema principal. Bakhtin assinala que são os gêneros que tornam possível a comunicação: “se os gêneros do discurso não existissem e nós não os dominássemos, se tivéssemos de criá-los pela primeira vez no processo do discurso, (...) a comunicação discursiva seria quase impossível” (BAKHTIN, 2003: 283). Isso parece decorrer da distinção entre significado e sentido. Para Bakhtin, o significado (das palavras e sentenças da língua) é apenas a forma possível do sentido; este só se materializa na enunciação, isto é, nas condições concretas do discurso. O sentido é, pois, função do contexto do enunciado, que torna real o possível do significado no domínio da língua. Ora, os contextos possíveis dos enunciados devem ser infinitos (ou pelo menos incalculáveis); e se os significados da língua devem dar conta de todos os sentidos possíveis, segue-se que deve existir pelo menos uma quantidade indeterminada (ou mesmo infinita) de sentidos. Na comunicação real, o interlocutor deve se colocar, então, a questão: qual dos significados possíveis é o pretendido

pelo locutor, aquele segundo o qual o enunciado deve ser compreendido, numa palavra qual o sentido do enunciado?

Tudo indica que, ao organizar o enunciado, o gênero do discurso segundo o qual ele foi produzido fornece algumas das pistas necessárias a essa decisão e, portanto, à compreensão. Para Bakhtin, “imaginamos o que o falante quer dizer” (isto é, qual a sua vontade discursiva verbalizada) e utilizamos essas pistas para compreender o enunciado. Ora, “a vontade discursiva do falante se realiza antes de tudo na escolha de um certo gênero de discurso” (BAKHTIN, 2003: 282). Daí podermos afirmar que o gênero é também um dos princípios organizadores da compreensão.

Trata-se agora de retrazar o caminho percorrido até aqui, repondo a questão que nos move – determinar se a concepção de Bakhtin é uma boa base para raciocinarmos sobre os diversos eventos comunicativos ocorrentes no ciberespaço. Acompanhando o caminho do pensador russo, determinamos, em primeiro lugar, que os gêneros do discurso são produtos da interação discursiva nos diversos campos da atividade humana e se cristalizam em formas típicas relativamente estáveis. Ora, o ciberespaço é certamente um (e provavelmente mais de um) importante campo da atividade humana; segue-se que ele deve produzir novos gêneros do discurso – o que de fato ocorre: inúmeras pesquisas recentes dão conta dessa enorme ampliação dos horizontes dos gêneros do discurso. Quais são esses possíveis diferentes “campos” da comunicação e se ocorrem de fato no ciberespaço, quais suas características mais marcantes, sua função no contexto comunicativo, seu impacto nas relações sociais, etc. – tais questões podem ser postas por uma perspectiva coerente com a de Bakhtin.

O segundo saldo da discussão anterior refere-se ao fato de os gêneros do discurso estarem ancorados diretamente na comunicação concreta, tal como a concebe Bakhtin – como interação entre um locutor, um interlocutor e os outros sujeitos dos discursos dos quais o enunciado presente é apenas um elo. Ao enunciar, o locutor produz seus enunciados a partir da seleção de formas típicas presentes no repertório disponível no gênero mais adequado às suas necessidades comunicativas e ao seu objeto. Isso orienta o enunciado a enunciados antecedentes – ao discurso do outro, dialogando com eles diretamente (ver BAKHTIN, 1981: 157 e seguintes), seja indiretamente através do próprio gênero, as formas prontas do enunciado. O enunciado é igualmente orientado ao interlocutor (a audiência, na tradição retórica) ao qual passa a palavra e essa orientação é igualmente constitutiva, mesmo em gêneros simples e distantes dos domínios da retórica clássica. Concluimos assim que essa tríade – o locutor, o interlocutor e o(s) outro(s) do discurso – e sua interação nos domínios da comunicação concreta são essenciais à abordagem dos gêneros na tradição bakhtiniana.

Alguns problemas trazidos para a presente pesquisa por tal enfoque podem ser facilmente enumerados: a determinação da situação comunicativa implicada nos vários eventos discursivos ocorrentes no ciberespaço; a tipificação das *personae* discursivas – o locutor, o interlocutor e o outro – presentes (ou ausentes) nos vários eventos; o lugar do dialogismo discursivo nos diversos domínios do ciberespaço etc.

O terceiro aspecto relevante que cabe ressaltar a partir da discussão a propósito do pensamento bakhtiniano sobre os gêneros do discurso refere-se ao seu papel como ordenador da comunicação. Todo gênero se constitui como um conjunto de convenções discursivas e formas típicas do enunciado que devem ser utilizados – com grau variável de restrição sobre as escolhas do enunciador – na comunicação concreta. Por outro lado, se tais restrições afetam as escolhas disponíveis ao falante, elas afetam igualmente a audiência, aparecendo como um conjunto de expectativas a propósito do enunciado do outro tais que orientam e restringem a compreensão. Em outras palavras, tais convenções de enunciação incluem igualmente convenções de interpretação. Obviamente, a delimitação de tais convenções e a determinação do grau de restrição que impõem à enunciação e à compreensão são algumas das abordagens que devem estar presentes na pesquisa a propósito dos eventos comunicativos do ciberespaço.

2.1.2. Gêneros e eventos comunicativos: a perspectiva de Swales

Passemos agora à abordagem (retórica) de Swales. Preliminarmente, cabe assinalar (novamente) que nosso interesse nessa perspectiva é menos fundamentar a aplicação do seu clássico modelo de análise (CARS – *Create A Research Space*) no nosso próprio *corpus* que buscar idéias que possam fundamentar uma teoria de gêneros discursivos facilmente aplicável aos eventos semióticos incidentes no ciberespaço. Não que a noção de *movimento retórico*, idéia guia daquele modelo, seja inaplicável a uma ampla gama desses eventos; no entanto, essa noção aplica-se particularmente bem a textos cuja coesão é temporal, como os construídos estritamente com a linguagem verbal – e interessa-nos particularmente aquela classe de textos onde a coesão espacial desempenha papel significativo, onde seria mais apropriado falar de uma *distribuição retórica*. Assim, o que nos move aqui são os fundamentos teóricos daquela análise, mais especialmente uma possível teoria dos gêneros em Swales.

Swales introduz seu clássico estudo sobre os gêneros acadêmicos com uma definição – aliás bastante longa – do que sejam gêneros discursivos no contexto do seu interesse. Para ele, um gênero pode ser concebido como “uma classe de eventos comunicativos, onde cada um deles partilha um conjunto de propósitos comunicativos (...) reconhecidos pelos membros experientes da comunidade discursiva à qual pertencem (SWALES, 1990: 58). Três conceitos chave são utilizados na definição, a saber: evento comunicativo, propósito comunicativo e comunidade discursiva. Examinemos cada um deles rapidamente, buscando igualmente sua relevância para o propósito que nos move.

Por evento comunicativo, Swales entende qualquer evento onde “a linguagem (e/ou a *para-linguagem*) desempenha um papel significativo e indispensável” (SWALES, 1990: 45). Ainda que a definição admita eventos suportados por linguagens não verbais, Swales

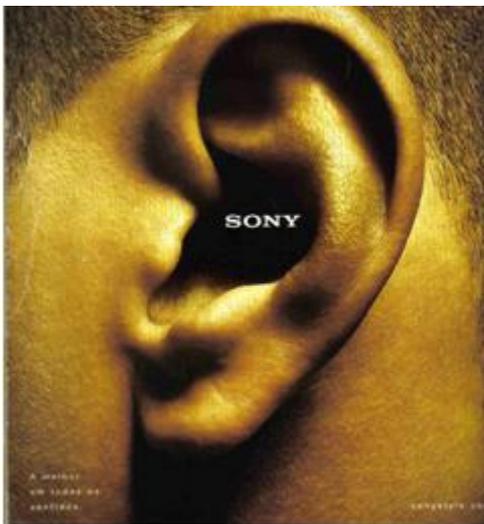


Ilustração 6: Exemplo de comunicação não verbal

explicitamente recusa *status* comunicativo a eventos semióticos que “envolvam os olhos e ouvidos em meios não verbais tais como admirar quadros ou ouvir música” (SWALES, 1990: 46). Diante dessa restrição, a inclusão entre parênteses de *para-linguagens* (ou simplesmente linguagens não verbais) na definição torna-se subitamente obscura e contraditória. Ela imediatamente exclui uma classe bastante ampla de eventos semióticos claramente comunicativos, ainda que apenas marginalmente suportados pela linguagem verbal. Um exemplo típico pode ser visto na Ilustração 6, uma peça publicitária escolhida arbitrariamente.

Nessa peça publicitária, a linguagem verbal é claramente acessória. Sony não é propriamente uma palavra, mas uma marca, e poderia ser facilmente substituído por um logotipo, se a Sony tivesse um logotipo universalmente. A frase em posição marginal (“A melhor em todos os sentidos.”) apenas agrega significação à composição – e só implica em polissemia se confrontada com a foto. Significação e sentido nesse tipo de evento não têm certamente o mesmo valor das comunicações estritamente verbais; no entanto, é possível dizer que estão presentes na composição, podendo ser “traduzidos” em linguagem verbal, ainda que as “traduções” possam não ser inequívocas ou consensuais. Aliás, que traduções entre diferentes idiomas o são? E mais: a organização textual, a composição fotográfica e verbal aqui desempenha uma função retórica inegável. No evento apresentado na Ilustração 6 – a interação entre o produtor da peça publicitária e o espectador – tanto a linguagem (verbal)

quanto algo que podemos chamar provisoriamente uma *para-linguagem* desempenham “um papel significativo e indispensável”, não havendo razões para não ser considerado um *evento comunicativo*.

Portanto, a restrição estabelecida por Swales à definição de *evento comunicativo* não parece ser justificada. Ela implica igualmente que categorias como “natureza morta” e “retrato”, aplicadas à pintura, ou “sinfonia” e “minueto”, aplicadas à música, não são, nos discursos pictórico e musical, do mesmo tipo das categorias de “romance” e “conto” no discurso literário. Aliás, aceitá-la implica na impossibilidade de falar de coisas tais como discursos pictórico e musical. Portanto, nossa abordagem ignorará, um tanto arbitrariamente, a restrição imposta por Swales, preferindo sua definição mínima de evento comunicativo. Feita esta ressalva, consideraremos como *comunicativo* qualquer evento semiótico e o conceberemos, em consonância com Swales, como composto não apenas pelo discurso em si e seus participantes, como também pelo papel desse discurso e pelo ambiente da sua produção e recepção, incluindo seu contexto histórico e cultural.

O segundo conceito-chave incluso na definição é o de *propósito comunicativo*, isto é, “os gêneros são veículos de comunicação para a realização de objetivos” (SWALES, 1990: 46). Ainda que tais propósitos (*compartilhados* pela comunidade sócio-retórica) possam ser de difícil identificação e, portanto, impliquem em dificuldades conceituais, Swales entende que é a existência de um ou mais propósitos comunicativos a principal característica que torna um conjunto qualquer de eventos comunicativos num *gênero*. A dificuldade conceitual está provavelmente ligada à dificuldade de reconhecimento desses propósitos; se o propósito de uma gênero não é facilmente reconhecível, como ele pode ser *compartilhado* pelos membros da comunidade sócio-retórica?

O terceiro conceito envolvido na definição é o de *comunidade discursiva*. Ao introduzir o problema dos gêneros, Swales assinala que nas comunidades sócio-lingüísticas de fala as necessidades comunicativas do grupo, a socialização e a solidariedade, “tendem a predominar no desenvolvimento e na manutenção das suas características discursivas” (SWALES, 1990: 24). Isso significa que os determinantes primários do comportamento lingüístico são sociais. Por outro lado, numa comunidade discursiva sócio-retórica, “os determinantes primários do comportamento lingüístico são funcionais, uma vez que uma comunidade discursiva consiste em grupos de pessoas que se associam para obter objetivos prioritários sobre aqueles ligados à socialização e à solidariedade” (SWALES, 1990: 24). Em outras palavras, nas comunidades discursivas, as necessidades comunicativas tendem a predominar sobre as necessidades sociais, ainda que estas acabem igualmente por ocorrer.

Swales elenca um conjunto de características que permitem reconhecer uma comunidade discursiva, melhor exploradas e estendidas em trabalho posterior (SWALES, 1992). Para o pesquisador, uma comunidade discursiva:

- i Dispõe de um conjunto de metas passíveis de descoberta; podem estar pública e explicitamente formulados e ser aceitos total ou parcialmente por seus membros;
- ii Dispõe de mecanismos de comunicação entre seus membros;
- iii Utiliza esses mecanismos de participação para inúmeras tarefas, entre elas fornecer informações de desempenho e *feedback*; inovações no canal; manutenção do sistema de crenças e valores da comunidade e aumentar o espaço profissional;
- iv Utiliza uma abrangente seleção de gêneros na perseguição de suas metas e como realização dos seus mecanismos participativos;
- v Desenvolve continuamente uma terminologia especializada;
- vi Dispõe de uma estrutura hierárquica explícita ou implícita para gerenciar o processo de progressão na comunidade discursiva.

Novamente, podemos fazer objeções às restrições impostas à noção. Pensamos principalmente em gêneros complexos como os literários. Se utilizarmos tais critérios restritivos, teremos dificuldade em associar tais gêneros discursivos a comunidades discursivas específicas. O fato é que tais características se aplicam facilmente às comunidades discursivas estudadas pela pesquisa de Swales, mas não é certo que se apliquem com a mesma simplicidade a *qualquer* contexto onde possamos observar a incidência e a utilização de gêneros discursivos – ou mesmo a todas as comunidades cujos objetivos compartilhados possam ser designados como primariamente *comunicativos*. Uma segunda objeção que se pode fazer à noção é a implícita homogeneidade assumida na caracterização das comunidades discursivas – e se há algo que possa ser dito sem maiores justificativas sobre a linguagem em uso é sua heterogeneidade constitutiva. Embora um corte desse tipo possa ser útil em algumas abordagens, assumi-las em qualquer contexto pode significar a perda de uma das mais relevantes dimensões do discurso. Essas dificuldades tornam, portanto, o conceito inútil para o propósito que nos move.

Esses elementos – evento comunicativo, propósito e comunidade discursiva – não esgotam o problema da identificação e descrição dos gêneros do discurso. Swales assinala que os propósitos comunicativos “constituem as razões do gênero” e “moldam a estrutura esquemática do discurso, influenciando e restringindo a escolha do conteúdo e do estilo”

(SWALES, 1990: 58). Em outras palavras, eles restringem as contribuições permitidas aos membros da comunidade discursiva para o conteúdo, o posicionamento e a forma do exemplar do gênero. Isso significa que todo gênero implica determinadas *convenções* (discursivas) que devem ser seguidas em alguma medida para que possamos identificar um determinado “exemplo” como *exemplar* do gênero. Tais convenções não são impositivas, fazendo com que os “exemplares ou instâncias dos gêneros variem em sua prototipicidade” (SWALES, 1990: 49). O conceito de *protótipo* é, pois, o quarto elemento chave presente na definição de um gênero.

Segundo Swales, esse conceito pode ser abordado de duas maneiras: pela abordagem da *definição* e pela abordagem da *semelhança de família*. A primeira abordagem remonta à tradição aristotélica e trata o processo de categorização como o reconhecimento de “um conjunto de propriedades simples que são individualmente necessárias e cumulativamente suficientes para identificar todos os membros e apenas os membros de uma categoria” (SWALES, 1990: 49). Essa abordagem, se bem sucedida em determinadas áreas do conhecimento (em especial o conhecimento de algumas disciplinas científicas), fracassa no tratamento de categorias emergentes na experiência cotidiana.

Uma abordagem distinta para o problema da categorização é a da *semelhança de família*, que remonta às discussões de Wittgenstein a propósito da noção de jogo. Por essa abordagem, os atributos de uma categoria são vistos como “uma rede complexa de similaridades que se sobrepõem e se cruzam, algumas delas semelhanças genéricas, outras semelhanças nos detalhes” (SWALES, 1990: 50). Isso significa que, para essa abordagem, a categoria devem ser apreendida com base em *exemplos* e não através da inferência de condições necessárias e suficientes. Swales assinala que foi com base na discussão de Wittgenstein (ver WITTGENSTEIN, 1979) do problema da semelhança de família na apreensão das categorias que se desenvolveu uma teoria dos “protótipos”, que ele se utiliza na compreensão dos gêneros do discurso. Como o autor não se estende nessa discussão, façamos uma rápida digressão.

Em pesquisa sobre o problema da categorização na Lingüística, Taylor assume que as entidades são categorizadas com base nos seus atributos. No entanto, “esses atributos não são os construtos binários da abordagem clássica” (TAYLOR, 1991: 41), cuja presença ou ausência é utilizada na determinação de condições necessárias e suficientes para a definição de uma categoria. Para ele, a “categorização de uma entidade não é uma questão de determinar se a entidade possui um atributo ou não”, mas de determinar quão perto de certos valores ótimos este atributo está. Citando análises de experimento clássico de Labov, Taylor

assinala que “nenhum atributo singular é essencial para distinguir uma categoria de outra”; a presença ou ausência de um determinado atributo simplesmente aumenta ou diminui a probabilidade de que uma determinada instância seja categorizada de tal ou qual maneira. Isto é, a presença de determinados atributos apenas aumenta o grau de *pertinência* (*membership*) de uma instância em relação a uma determinada categoria.

Por outro lado, a abordagem clássica considera que os atributos que intervêm na definição de uma categoria são “propriedades tangíveis” das coisas designadas. Taylor assinala que, em inúmeros casos, também atributos *funcionais* (concernentes ao uso dado aos objetos) e *interacionais* (concernentes ao modo como as pessoas lidam com os objetos) são utilizados no reconhecimento da categoria à qual pertence uma determinada instância. Mais precisamente ainda, tais atributos não têm a ver com as “propriedades inerentes aos objetos em si, mas com o papel desempenhado por esses objetos numa cultura em particular” (TAYLOR, 1991: 41).

Como, então, podemos distinguir “copos” de “xícaras” ou “vasos”, isto é, como construímos as categorias? Como vimos, utilizamos *atributos* heterogêneos (funcionais, interacionais, culturais, etc.) para categorizar. Tais atributos podem ser de dois tipos: binários, cuja presença ou ausência marca de alguma forma a categoria, ou contínuos, cuja variação também marca a categoria. Assim é que dizemos que uma “caneca” tem (ou costuma ter) uma alça e um copo, não; por outro lado, uma “caneca” tende a ser dita maior que uma “xícara”, etc.. Em qualquer caso, os diferentes atributos contribuem em conjunto para a categorização, isto é, não é o valor (contínuo ou binário) de *um* atributo que determina se uma instância pertence ou não a uma categoria, mas o valor do conjunto. Além disso, pertencer ou não pertencer a uma categoria são apenas os valores extremos do juízo: as instâncias são avaliadas como referência ao seu *grau de pertinência* à categoria. Isso, evidentemente, torna *nebulosos* (no sentido da *fuzzy logic*) os limites entre os conjuntos de instâncias, isto é, entre as categorias. Podemos imaginar que aquelas entidades cognitivas que chamamos *categorias* se componham de conjuntos de elementos formados por outras entidades cognitivas chamadas *instâncias*, onde não podemos afirmar com certeza se uma instância pertence inequivocamente ou não à categoria; tudo o que podemos afirmar refere-se ao seu *grau* de pertinência em relação à categoria. As instâncias, por sua vez, podem ser concebidas como um conjunto de elementos formados por outras entidades cognitivas chamadas *atributos*, onde, novamente, tudo o que podemos afirmar refere-se ao *grau* de relevância do atributo para definir a pertinência da instância em relação à categoria.

Nessa concepção, os “limites da categoria são nebulosos [*fuzzy*]” (TAYLOR, 1991: 38). Tudo se passa, portanto, como se pudéssemos conceber as categorias como os conjuntos nebulosos (*fuzzy sets*) como os exibidos na Ilustração 7, que procura esquematizar a construção cognitiva das categorias “xícara”, “copo” e “vaso”. Como se pode notar na ilustração, alguns elementos se encontram mais próximos às “bordas” do conjunto, enquanto outros estão mais próximos das suas regiões centrais, o que equivale a dizer que existem elementos com diferentes graus de pertinência em relação ao conjunto.

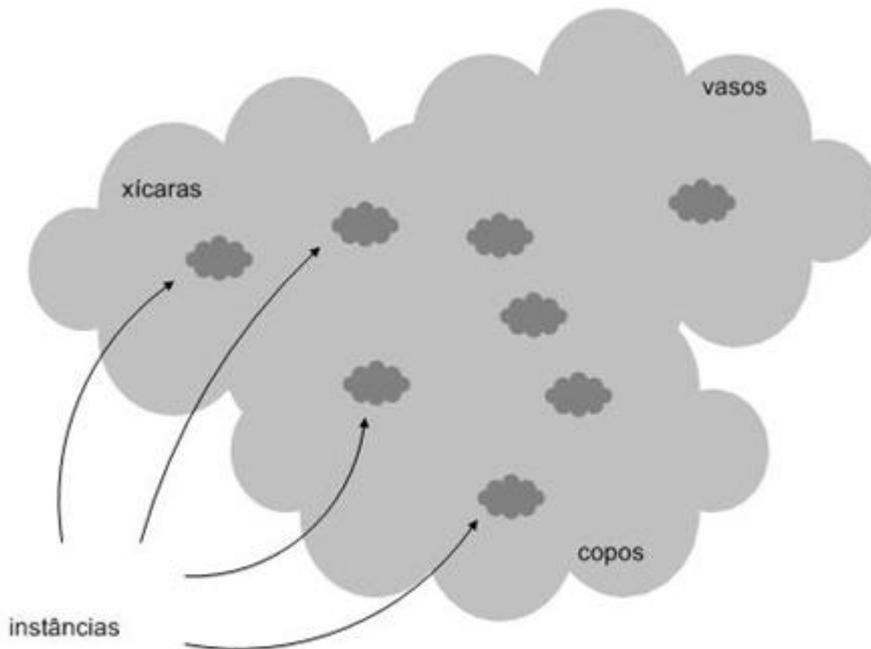


Ilustração 7: Representação visual de conjuntos nebulosos.

Do mesmo modo, para essa abordagem não é estranho que um mesmo elemento possa ser concebido como duas *instâncias* diferentes de duas categorias diversas. Isso decorre da difusão (*fuzziness*) dos limites do conjunto e resolve problemas tais como categorizar uma “baleia” (é um “mamífero” ou um “peixe?”). Desse modo, torna-se simples conceituar protótipos como aqueles elementos mais próximos do “centro” do conjunto nebuloso, isto é, aquelas instâncias cujo grau de pertinência à categoria é maior que todas as demais instâncias consideradas. É esta a noção de protótipo adotada por Taylor (e por Swales), que assinala que “esses protótipos servem como pontos de referência para a categorização de instâncias mal definidas” (TAYLOR, 1991: 42). É este o sentido da expressão “apreender a categoria através de exemplos”: trata-se de utilizar protótipos (isto é, aqueles elementos com maior grau de pertinência no conjunto) como marcos de referência de uma categoria.

Podemos agora retornar à caracterização de Swales, compreendendo mais precisamente sua afirmação de que “os exemplares ou instâncias do gênero variam em prototipicidade”. De modo mais preciso, Swales assinala que os “exemplares do gênero apresentam padrões variáveis de similaridade em termos de estrutura, estilo, conteúdo e audiência presumida”, de tal modo que, “se todas as expectativas mais prováveis forem satisfeitas, o exemplar será visto como prototípico pela comunidade discursiva à qual pertence” (SWALES, 1990: 58).

Outros elementos são considerados como relevantes por Swales na sua caracterização dos gêneros do discurso, em especial a terminologia elaborada pela comunidade discursiva. Os termos atribuídos ao gênero e a seus constituintes pelos membros mais experientes da comunidade discursiva são indicativos de como concebem a função retórica desses eventos comunicativos. No entanto, não nos alongaremos mais na discussão, já que o que vimos até o momento é suficiente para caracterizar os elementos centrais da sua abordagem.

A essa altura é conveniente uma pausa para avaliarmos comparativamente as contribuições de Bakhtin e Swales para o problema da caracterização dos gêneros do discurso. É certo que, pelo que vimos até o momento, persistem importantes diferenças nas duas abordagens; porém, notáveis concordâncias podem ser obtidas, pelo menos em aspectos importantes das definições com as quais trabalham.

É certo que a diferença mais relevante entre as duas abordagens é o fato de Bakhtin fundar explicitamente sua concepção de gêneros do discurso numa teoria da linguagem muito particular, enquanto que Swales apenas marginalmente aborda o problema e, sobretudo, não deriva sua concepção diretamente de uma teoria da linguagem. Swales se basta em assumir a premissa de que o uso da linguagem é uma forma de comportamento social, enquanto que Bakhtin deriva sua concepção de gênero de uma teoria dialógica da linguagem. Essa derivação implica na centralidade da noção de diálogo como constituinte do discurso (e, portanto, do gênero), noção estranha a Swales. Apesar disso, Swales concorda notavelmente com Bakhtin ao enfatizar a centralidade da comunicação na caracterização dos gêneros; ainda que por razões possivelmente distintas, ambos (e cada um com o seu jargão próprio) enfatizam o papel do propósito comunicativo (ou vontade discursiva), da presunção da audiência (ou orientação ao interlocutor) e da situação comunicativa concreta na determinação das convenções do gênero. Bakhtin, é certo, enfatiza o caráter ordenador do discurso conferido aos gêneros, afirmando mesmo que a comunicação seriam impossível sem eles; essa noção é igualmente estranha a Swales – talvez porque seu campo de pesquisa seja menos abrangente que o proposto pelo pensador russo.

É interessante notar que, apesar de Bakhtin fundar sua concepção de gênero numa teoria da linguagem, é Swales quem ancora a noção mais diretamente nas *atividades discursivas*, construindo-a em torno do conceito de *comunidade discursiva*. Bakhtin se basta em assinalar que os gêneros são produto dos diversos “campos da atividade humana”, mas não precisa que campos seriam esses. Swales, ao assumir a premissa (igualmente válida para a tradição bakhtiniana) do uso da linguagem como forma de comportamento social, deriva dela uma noção precisa desses “campos”. Ainda que possamos objetar à caracterização de Swales do que sejam (e de como podemos reconhecê-las) comunidades discursivas, é certo que essa noção agrega algum valor à teoria dos gêneros, pelo menos em determinados contextos.

É igualmente notável também a contribuição para a teoria dos gêneros que pode ser trazida pela noção de *protótipo*, tal como conceituamos há pouco. A Bakhtin basta caracterizar os gêneros como “formas típicas, relativamente estáveis”; a noção de protótipo nos permite determinar como essas formas se tornam “típicas” e sobretudo como são reconhecidas e utilizadas no contexto dos gêneros do discurso.

O que queremos deixar claro com essa breve avaliação é o fato de que as duas concepções, ainda que distintas, não são contraditórias: afirmar uma não precisa ser, necessariamente, negar a outra. Em outras palavras, os pontos de divergência que assinalamos não precisam ser *polêmicos*. Se estes argumentos forem pelo menos admissíveis, justifica-se nossa pretensão de procurar extrair das diversas tradições algumas idéias fundamentais que possam ser aplicadas produtivamente ao problema geral dos gêneros digitais e à melhor compreensão do nosso *corpus* em particular.

2.1.3. A multimodalidade e a produção do sentido em Kress e Van Leeuwen

Quando discutimos o conceito de evento comunicativo em Swales, objetamos à sua restrição de que semioses não verbais pudessem fazer parte dessa classe de eventos e procuramos assinalar que essa restrição não se justificava. Nossa preocupação, naturalmente, se devia ao fato de que, como já vimos, imagem, som, movimento, cor e linguagem verbal são elementos decisivos para uma compreensão abrangente dos eventos incidentes no ciberespaço. Trata-se agora de incorporar aos conceitos que vimos extraindo das diferentes abordagens do problema dos gêneros do discurso um corpo teórico que, presumivelmente, nos

auxiliará na caracterização dos textos de hipermídia. Trata-se das concepções de Kress e Van Leeuwen. Embora uma teoria dos gêneros seja marginal na concepção dos autores, sua relevância para o propósito que nos move nos obriga a rastrear o problema na tradição em que se inscrevem, buscando determinar sua coerência com a noção implícita que podemos inferir nos trabalhos mais recentes dos autores.

Esse rastreamento nos remete a antiga obra de Kress, essa em co-autoria com Hodge. Na obra, os autores conceituam os gêneros do discurso como uma categoria particular do que chamam *sistemas logonômicos*, definidos como “um conjunto de regras prescrevendo as condições de produção e recepção do sentido”. Tais sistemas determinariam “quem pode iniciar (produzir, comunicar) ou conhecer (receber, entender) sentidos a propósito de determinados tópicos” (HODGE; KRESS, 1988: 4). Dentro desse conjunto de conceitos, os gêneros seriam “regras que restringem a forma geral dos textos” e podem ser definidos mais precisamente como “formas típicas de textos que ligam tipos de produtores, consumidores, tópicos, meios, maneira e coesão” (HODGE; KRESS, 1988: 7), controlando o comportamento de produtores e as expectativas de consumidores e codificando as relações entre os participantes sociais.

Observamos aqui a recorrência da noção de formas típicas e prototipicidade associada ao problema dos gêneros. Além disso, os autores enfatizam o papel do gênero na determinação das condições sob as quais os textos operam, isto é, na situação comunicativa à qual estão associados. Por outro lado, os autores enfatizam o caráter *prescritivo* das condições de produção e recepção do sentido, dando maior ênfase ao que Bakhtin chamava o “papel ordenador” dos gêneros: os autores explicitamente se referem ao *controle* que exercem sobre o comportamento dos produtores e as expectativas dos consumidores dos textos.

Trata-se, portanto, de concepção que não contradiz a tradição bakhtiniana (até agora a mais adequada ao propósito que nos move), enfatizando alguns aspectos pouco aprofundados pelo pensador russo. No entanto, apenas essa concepção pouco agrega ao ferramental conceitual que já acumulamos. Assim, devemos prosseguir na abordagem dessa tradição teórica, saltando diretamente para um conceito sedutor – o de multimodalidade.

Para Kress e Van Leeuwen o conceito de *multimodalidade* é entendido “como o uso de vários modos semióticos na concepção de um produto ou evento semiótico” (KRESS & VAN LEEUWEN, 2001: 20). Para os autores, todos os modos (recursos semióticos – ou linguagens, em sentido amplo – como cores, imagens, linguagem verbal, sons, etc.) incluídos num artefato ou evento semiótico contribuem para a produção do sentido na interação humana. Do ponto de vista do objeto de estudos que nos move, uma teoria que pretenda dar conta da produção

do sentido em artefatos (e eventos) multimodais deve nos ajudar a compreender a produção do sentido nos eventos da Internet, onde o hipertexto e a hipermídia desempenham papel cada vez mais central.

Mas como se dá essa contribuição dos vários recursos semióticos para o sentido? Uma resposta preliminar pode ser buscada em trabalho anterior dos autores (KRESS & VAN LEEUWEN, 1996), dedicado ao estudo da gramática do *design* visual. Naquele trabalho, os autores se dissociam da tradição da “semiologia” baseada em algumas idéias de Saussure, particularmente na sua noção de signo tal como foi estendida a domínios trans-lingüísticos. Para os autores, o signo “não é a união pré-existente de um significante a um significado, destinado a ser reconhecido e utilizado em bloco, (...) mas um processo de sinalização (*sign-making*)” (KRESS & VAN LEEUWEN, 1996: 7). Trata-se de uma *atividade* de construção do sentido a partir de recursos disponíveis, onde a subjetividade desempenha papel relevante. Essa atividade seria conduzida por processos cognitivos de natureza metafórica. Os autores ilustram essa assertiva acompanhando ao longo do tempo o grafismo de uma criança de 2 a 3 anos, chamada a desenhar um carro. Os quatro momentos captados estão representados na Ilustração .

Os autores chamam a atenção especialmente para o processo analógico em curso ao longo de 14 meses na representação da criança. A propósito da última imagem da Ilustração , os autores assinalam que, para a criança, o carro é definido como algo que tem rodas e rodas podem ser representadas por círculos. Trata-se, portanto, de um duplo processo metafórico conduzindo dos círculos ao carro, a saber: carro → rodas → círculos.

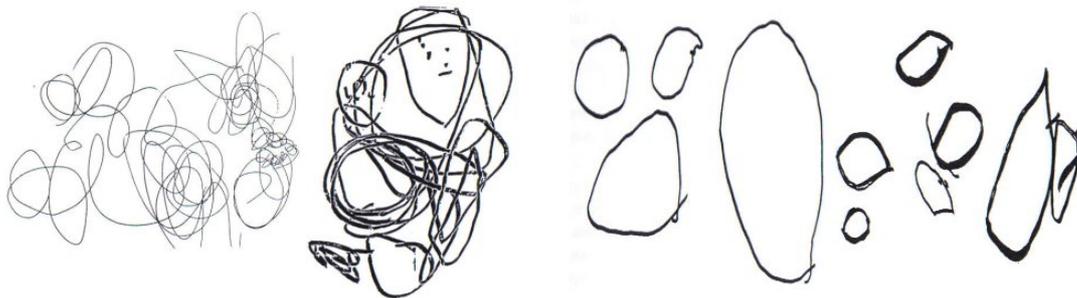


Ilustração 8: O sinal gráfico de um carro ao longo do tempo (fonte: Kress & Van Leeuwen).

Outro interessante aspecto é o próprio processo de conceitualização do carro, acompanhado pelas três imagens anteriores, produzidas com intervalo de alguns meses e representando o “mesmo” carro. Nota-se aí a progressiva conceitualização de uma imagem primitiva – a de movimento, – provavelmente fruto da experiência da criança (os autores utilizam a expressão *interesse*) com carros, num contínuo processo de abstração (na verdade,

uma classificação por analogia). Assim, os autores concluem que, na atividade de sinalização, os signos são “construtos *motivados* de significantes (formas) e significados (sentidos)” (KRESS & VAN LEEUWEN, 1996: 7). Note-se que a motivação aqui não é uma relação que se estabelece entre o sinal e seu (suposto) referente, mas entre o produtor do sinal e o próprio sinal, motivação nascida da experiência (do interesse) cognitiva em face dos recursos (semióticos) disponíveis.

Os processos cognitivos, porém, são apenas parte da atividade de significação. Igualmente relevante é o contexto social e cultural onde a comunicação ocorre. Os autores assinalam a existência de um “controle social sobre o sentido”, controle que se exerce seja na seleção e estruturação das imagens, seja na sua interpretação, mesmo naquelas situações de “leituras aparentemente abertas” (KRESS & VAN LEEUWEN, 1996: 26). Esse controle aparece sob a forma de “codificações” pré-existentes, “naturalizadas” no contexto da vida social. Para os autores, “a comunicação visual é sempre codificada”; essa codificação é transparente para os usuários da imagem “simplesmente porque nós já conhecemos seus códigos, pelo menos passivamente – mas sem saber que o sabemos, sem os meios de falar sobre o que fazemos quando lemos uma imagem” (KRESS & VAN LEEUWEN, 1996: 32).

Um exemplo bastante típico e ilustrativo desse “controle sobre o sentido” é discutido a propósito do problema da modalidade aplicado à semiótica da imagem. Os autores assinalam que o conceito de modalidade, tomado por empréstimo da Linguística, “se refere ao valor de verdade ou credibilidade de enunciados (...) sobre o mundo” (KRESS & VAN LEEUWEN, 1996: 160). No entanto, a modalidade “não expressa verdades ou falsidades absolutas; ela produz verdades compartilhadas, alinhando os leitores ou ouvintes com alguns enunciados e distanciando-os de outros” (KRESS & VAN LEEUWEN, 1996: 160). A modalidade cria um “nós” imaginário. Ora, as imagens podem representar o mundo como se fosse real ou, ao contrário, como se fosse imaginário. Essa representação é também modalizada, é também uma função de como as coisas são valoradas pelo seu produtor. No entanto, “também os julgamentos modais são sociais, dependentes do que é considerado real (ou verdadeiro, ou consagrado) no grupo social para o qual a representação foi originalmente destinada” (KRESS & VAN LEEUWEN, 1996: 160).

Isso pode ser facilmente notado submetendo-se à análise a escala de modalidade do padrão de cores na representação naturalista. Para essa representação, a realidade “é definida com base na correspondência entre a representação visual de um objeto e o que normalmente vemos deste objeto à vista desarmada” (KRESS & VAN LEEUWEN, 1996: 163). Esta é a definição “naturalizada” da realidade para esse tipo de representação; ela, contudo, é fundada

sobre a capacidade de resolução de cores e imagem da fotografia em 35 milímetros típica do cinema sob a tecnologia atual – e certamente não era a definição da pintura realista anterior à invenção da fotografia e do cinema.

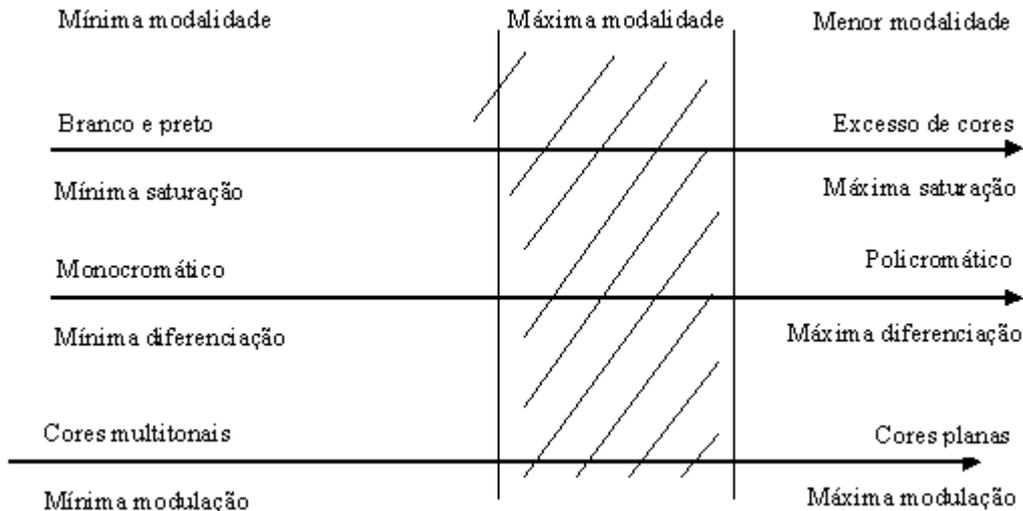


Ilustração 9: Escala de modalidade das cores na representação naturalista.

Essa escala de modalidade se exprime, por exemplo, na codificação das cores. Analisando nas cores as variáveis da saturação, diferenciação e modulação, os autores assinalam que o máximo da modalidade é obtido, na representação naturalista, em algum

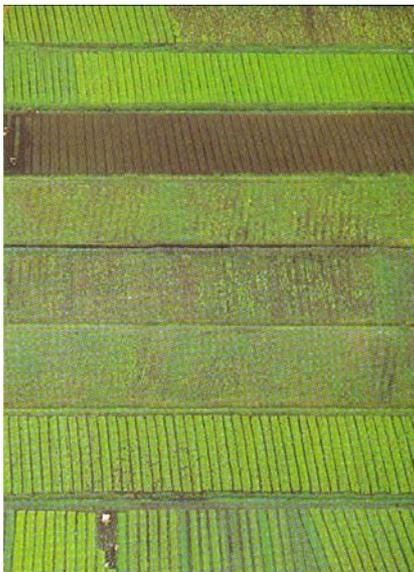


Ilustração 10: Fotografia de Hugo Faleiros

intervalo de pontos dentro da escala entre o branco/preto e o excesso de cores, os tons monocromáticos e os policromáticos e a multitonalidade e as cores planas, respectivamente, como visto na Ilustração . Para a representação naturalista, os extremos da escala “distorcem a realidade” ou, mais precisamente, “o que é entendido como real depende de como a realidade é definida por um grupo social em particular” (KRESS & VAN LEEUWEN, 1996: 163). Essas convenções, essas codificações são “naturalizadas”, aparecem como *transparentes* para os intérpretes das imagens; no entanto, elas não passam de convenções.

Veja-se, por exemplo, a fotografia exibida na Ilustração , colhida numa publicação de uso popular destinada a “ensinar” a fotografar. O autor do texto explicativo assinala que se trata da “visão diferente de uma horta” (COMO, 1988: 53). Que *diferença* é essa de que trata o autor do texto que acompanha a publicação

senão a “diferença da realidade” tal como concebida pelo padrão naturalista? Afinal, trata-se de uma horta, vista de cima! No entanto, ela não parece “representar a realidade” de uma horta... O autor do texto “sabe” que se trata de uma horta; no entanto, como a fotografia foi produzida deliberadamente com vistas à baixa modalidade, o que ele vê nela é sua *diferença da realidade*. O que torna a fotografia “diferente” é tanto a falta das convenções de perspectiva, falta obtida pela visão em ângulo reto com o objeto, quanto a pouca diferenciação de cores; o que a torna diferente é a *atitude* escolhida pelo fotógrafo, que produz um efeito de baixa modalidade (sempre na escala naturalista). Ou como diriam os teóricos que estamos interpretando: “a realidade está nos olhos do observador” (KRESS & VAN LEEUWEN, 1996: 163).

Obra conjunta mais recente de Kress e Van Leeuwen (2001) estende e refina a abordagem teórica tratada sumariamente até agora. Aqui o interesse é mais que discutir o papel das imagens na construção do sentido; é elaborar uma teoria abrangente a propósito da construção do sentido a partir de artefatos e eventos que incluem diversas semioses. Para isso, os autores introduzem quatro novas categorias. A primeira delas é a de discurso, definido como “conhecimentos socialmente constituídos sobre (algum aspecto de) a realidade” (KRESS & VAN LEEUWEN, 2001: 4). Embora os autores não se aprofundem na descrição de suas características, salvo a indicação de que ele é “relativamente independente do gênero, do modo e (às vezes menos) do *design*” (KRESS & VAN LEEUWEN, 2001: 5), o acompanhamento da análise que realizam de diferentes artefatos semióticos indica que, para eles, o discurso parece ser um conjunto (possivelmente estruturado) de enunciados (conhecimentos, julgamentos, opiniões, etc.) sobre um determinado objeto. Esse entendimento parece estar relacionado a uma das concepções de discurso adotadas por Foucault ao longo de sua trajetória, em particular a noção de “formação (ou dispersão) discursiva”, utilizada “no caso em que entre os objetos, os tipos de enunciação, os conceitos, as escolhas temáticas, se puder definir uma regularidade (uma ordem, correlações, posições e funcionamentos, transformações)” (FOUCAULT, 1986: 43).

Trata-se apenas de assinalar certo parentesco com as idéias de Foucault e não de buscar definir um espaço de conjunção dos referenciais teóricos de ambos os autores, até porque essa conjunção é periférica. No mínimo, Kress e Van Leeuwen conferem uma relevância à atividade social (às práticas sociais) que Foucault não concede na determinação do discurso. Embora, como dissemos, os autores não se aprofundem na sua concepção de discurso, sua discussão do papel das práticas sociais na formação do discurso os alinha às concepções de Fairclough, que igualmente incorpora em parte a noção foucaultiana de formações discursivas

e ordem do discurso. Esse entendimento, contudo, não implica que “a constituição discursiva da sociedade” emane para ele de “um livre jogo de idéias”. Ao contrário, Fairclough concebe a relação entre discurso e estrutura social como uma relação dialética: o discurso tanto determina a estrutura social quanto é determinado por ela; ele emana “de uma prática social que está firmemente enraizada em estruturas sociais materiais”; por outro lado, como prática social que é concebido, “os efeitos constitutivos do discurso atuam conjugados com os de outras práticas” (FAIRCLOUGH, 2001: 93).

Outra categoria introduzida por Kress e Van Leeuwen é a de *design*, utilizada para exprimir a “concepção” que guia a produção dos artefatos semióticos, situando-a entre o conteúdo, onde os discursos atuam, e a expressão, que materializa os discursos. Para os autores, o *design* “é o aspecto conceitual da expressão e o aspecto expressivo do conceito. Os *designs* são (usos de) recursos semióticos, em todos os modos e combinações de modos semióticos” (KRESS & VAN LEEUWEN, 2001: 5), sendo concebidos, portanto, como os meios de materializar discursos na comunicação concreta. Os *designs* ainda podem se materializar de diferentes modos: o arquiteto “projeta” uma casa, mas não a constrói – e pode fazê-la materializar-se no papel ou num programa de computador – e ela ainda assim é conhecida como a materialização de discursos arquitetônicos.

As duas últimas categorias introduzidas na descrição da comunicação multimodal são a produção, a “articulação material real do evento semiótico ou à produção material real do artefato semiótico” (KRESS & VAN LEEUWEN, 2001: 6) e a distribuição. Esta última normalmente não é vista como material semiótico, já que supostamente não agregaria significação ao artefato ou evento semióticos. No entanto, as condições de circulação podem afetar – às vezes significativamente – o sentido do artefato ou evento. Os autores citam especificamente o problema da perda da “aura” da obra de arte pela reprodutibilidade técnica, citada por Benjamin.

Por fim, cabe assinalar que os autores entendem a comunicação “como um processo no qual um produto ou evento semiótico é articulado ou produzido e interpretado ou utilizado” (KRESS & VAN LEEUWEN, 2001: 20). Segue-se dessa compreensão que a produção e uso de objetos projetados e ambientes são considerados como uma forma de comunicação. Nesse sentido, o projeto, a construção e finalmente a habitação de uma residência são considerados, no seu todo, um evento semiótico, uma atividade comunicativa.

Como vimos, os “modos são recursos semióticos que permitem a realização simultânea de discursos e tipos de (inter)ação” (KRESS & VAN LEEUWEN, 2001: 22). O que importa a essa altura é o fato de o *design* usar esses recursos, selecionando e combinando modos

semióticos de acordo com as características e necessidade de uma determinada situação comunicativa, todos contribuindo para a formação do sentido.

A questão que se põe a essa altura é determinar se podemos cruzar essas concepções sobre o discurso multimodal com a noção de gêneros originalmente formulada por Hodge e Kress, de modo a ampliarmos o nosso ferramental conceitual. Como assinalamos anteriormente, Kress e Van Leeuwen concebem os eventos e artefatos semióticos como articulados por três camadas: a camada do discurso, a camada do *design* e a camada da expressão. Vimos também que existem determinadas convenções discursivas que afetam o *design*; no caso, nos debruçamos mais detalhadamente sobre o problema da modalidade das cores na representação naturalista. É possível imaginarmos que tais representações convencionais sejam regras e codificações (numa palavra, convenções) de *design* que organizam os discursos (ou pelo menos alguns discursos). Igualmente, as convenções de cor para a representação do gênero (sexual) típicas da nossa sociedade são codificações de *design* que cristalizam discursos. Tais convenções, tal como caracterizadas nos trabalhos mais recentes, parecem se relacionar àqueles sistemas logonômicos genéricos discutidos sumariamente a propósito da obra de Hodge e Kress. Pelo menos as noções não são contraditórias. No entanto, batizemos essas convenções de codificações ideológicas de *design* ou simplesmente de representações convencionais.

Ora, se podemos falar de gêneros dos discursos no contexto da semiótica mais recente de Kress e Van Leeuwen, eles deverão ser conceituados, evidentemente, também como convenções de *design*. No entanto, as convenções de gênero são relativamente independentes dos discursos (no sentido preciso que lhe emprestam Kress e Van Leeuwen). Portanto, aquelas representações convencionais de que falávamos não se identificam com os gêneros semióticos. Estes devem codificar elementos da camada da expressão (produção). Aqui cabe recordar que, no contexto das idéias de Kress e Van Leeuwen, não podemos caracterizar a camada de produção simplesmente como a materialidade do texto, mas como sua articulação concreta num dado contexto comunicativo, que envolve igualmente suas condições de produção e de recepção. Feita a ressalva, uma representação visual do lugar ocupado pelos gêneros semióticos na concepção em camadas de Kress e Van Leeuwen pode ser vista na Ilustração . Para utilizarmos o vocabulário da obra de Hodge e Kress, as áreas hachuradas estão diretamente relacionadas àqueles sistemas logonômicos ali discutidos.



Ilustração 11: Um possível lugar para as convenções de gêneros na concepção de Kress e Van Leeuwen.

Vimos anteriormente que é possível utilizar as duas principais tradições de estudo do problema dos gêneros do discurso no escopo de um único conjunto complementar de idéias, desde que algumas noções insatisfatórias de Swales sejam convenientemente ignoradas. O que esperamos ter assinalado agora é que essas idéias podem se estendidas pelas de Kress e Van Leeuwen de modo a abarcar um conjunto mais amplo de eventos comunicativos, conjunto mais amplo que aqueles determinados pela linguagem verbal, particularmente os eventos multimodais. Trata-se então de sintetizarmos essas idéias. Em primeiro lugar, julgamos mais abrangente conceber os eventos comunicativos como aquelas ocorrências que envolvem a produção e a utilização de artefatos semióticos. Neste caso, a interpretação pura (típica, por exemplo, da fruição estética) será apenas um caso particular da utilização dos artefatos semióticos, invariavelmente utilizados em contexto interpretativo. Portanto, qualquer interação onde possamos reconhecer a significação (isto é, aquelas interações onde as diferentes linguagens desempenham papel relevante) será interpretada como um evento comunicativo.

No contexto dessa definição, aqueles eventos cujos artefatos são implementados tendo em vista a satisfação de um ou mais propósitos comunicativos e que obedeçam a algumas convenções de *design* multimodal bem definidas podem ser considerados como instâncias de um (ou mais) gênero(s). Tais convenções afetam o modo como os artefatos são produzidos e interpretados, o que equivale dizer, utilizando o vocabulário bakhtiniano, que elas determinam os elementos composicionais e temáticos utilizados na sua concepção. Tais convenções se

cristalizam em formas prototípicas com referência às quais os exemplares do gênero são avaliados pelos membros da comunidade de seus usuários. Por outro lado, é conveniente ressaltar que tais convenções afetam igualmente as condições de produção e de recepção dos artefatos semióticos.

A essa altura, é conveniente assinalarmos que é preferível batizarmos essas convenções e formas prontas de *gêneros semióticos*. Isso é necessário em vista da noção precisa de discurso em Kress e Van Leeuwen: como vimos, tais convenções de *design* não estão orientados ao discurso tal como o concebem os autores, mas à produção e à materialização dos artefatos semióticos. De modo a não permitir confusão com aquelas representações convencionais que abordamos mais atrás (estas, sim, podendo mais apropriadamente ser batizadas de gêneros de discursos), julgamos aconselhável passar a utilizar essa designação.

Do mesmo modo, passaremos a utilizar a designação de gêneros digitais para aquelas categorias de artefatos utilizados no contexto dos eventos comunicativos incidentes no ciberespaço. Este último termo será utilizado para designar não a Internet como uma rede mundial de computadores, mas como um campo específico da atividade discursiva. A caracterização dessa categoria particular de gêneros semióticos será discutida mais adiante.

2.1.4. As mensagens do meio: gêneros do discurso em Maingueneau

Em conformidade com a tradição bakhtiniana, Maingueneau começa a abordar o problema assegurando que “todo texto pertence a uma categoria de discurso, a um *gênero de discurso*” (MAINGUENEAU, 2002: 59). Essas categorias podem aparecer para os interlocutores tanto como textos-modelo (protótipos) que estabelecem o padrão do gênero, quanto como “*rotinas*, de comportamentos estereotipados e anônimos que se estabilizaram pouco a pouco” (MAINGUENEAU, 2002: 65) e funcionam no sentido de proporcionar considerável *economia cognitiva* e assegurar a comunicação. Portanto, Maingueneau entende os gêneros como um conjunto de *regras* (mais ou menos estáveis) que afetam (ou talvez *determinem*, embora isso não fique claro no texto analisado) o comportamento discursivo dos interlocutores, isto é, como *atividades discursivas sociais*. Mais especificamente, Maingueneau traz em seu auxílio a teoria dos *speech acts* de Austin, e entende que, como um

“ato de linguagem de um nível de complexidade superior, um gênero do discurso encontra-se também submetido a um conjunto de condições de êxito”.

A primeira dessas condições de êxito é equivalente ao “propósito comunicativo” de Swales: um texto – e por extensão um gênero – deve ter uma finalidade reconhecida, tal que sua determinação “é indispensável para que o destinatário possa ter um comportamento adequado ao gênero do discurso utilizado” (MAINGUENEAU, 2002: 66). O autor não problematiza a condição, mas é possível levantar o problema de textos que ocultam seus propósitos por trás de finalidades aparentes distintas. Um modo de tratar esse problema, coerente com outros exemplos de Maingueneau, é assinalar aqui a *transgressão* (significativa) da norma (não coercitiva) do gênero. Uma atitude dessas pressupõe que a criatividade (em face do gênero) é transgressiva (ou, ao contrário, a transgressão é criativa). Isso poderia implicar que não existem gêneros com propósitos ambíguos, já que os *exemplos* é que seriam transgressivos e não os *exemplares*.

No entanto, como veremos mais adiante, pelo menos na correspondência eletrônica é possível identificar categorias de textos cujos exemplares ocultam seu propósito como artifício retórico para melhor atingi-lo. Isso significa que deveríamos considerar a hipótese de que algumas categorias de textos podem implicar numa situação comunicativa tal que a fórmula de Charaudeau, citada por Maingueneau (“Estamos aqui para dizer ou fazer o quê?”), seja respondida de modo diverso pelo locutor e pelo interlocutor, cada um deles reconhecendo uma finalidade distinta para a situação. Essa poderia ser uma hipótese interpretativa para o contexto comunicacional do *press-release* eletrônico, discutido sumariamente mais adiante, onde uma das partes “está ali” para “ser informado” a respeito de um “fato” e a outra “está ali” para “vender” a publicação de uma “notícia”. E essa talvez não seja uma situação comunicativa incomum!

Uma segunda condição de êxito dos gêneros seria o “estatuto de parceiros legítimos”, dando conta do que em outro contexto chamávamos o “papel” desempenhado pelos envolvidos numa situação comunicativa convencional. O autor assinala que “nos diferentes gêneros do discurso, já se determina de quem parte e a quem se dirige a fala” (MAINGUENEAU, 2002: 66), sendo que tais “papéis” só podem ser “legitimamente” representados por determinados atores sociais e não quaisquer outros. Essa concepção sugere relações implícitas de *poder* para os papéis representados pelos envolvidos; sugere também que tais relações são *assimétricas* também no interior da situação comunicativa. No texto citado, Maingueneau interpreta essas relações como pré-condição para o “êxito” da instância do gênero; por exemplo: “um curso universitário deve ser ministrado por um professor, que se

supõe deter um saber” (MAINGUENEAU, 2002: 66). No entanto, podemos (também “legitimamente”) interpretar de modo inverso o estatuto dos parceiros, afirmando que é o próprio estatuto do gênero que legitima aquele que assume um determinado papel, o que seria o mesmo que dizer que, se ministro um curso universitário, *então* tenho o saber para tanto... Por esse olhar, os gêneros intervêm como fatores de *naturalização* das relações (assimétricas) de poder envolvidas nas atividades sociais. Poderíamos, portanto e sem fugir à coerência com as idéias de Maingueneau, dizer que, na instanciação dos gêneros do discurso, o exercício do poder inclui suas próprias condições de legitimação.

Duas outras condições de êxito para os gêneros são ainda assinaladas: o lugar e o momento legítimos e a organização textual. A primeira inclui-se nas características da situação comunicativa em que se insere o evento e é coerente com os conceitos discutidos anteriormente; a segunda assinala o fato igualmente conhecido de que as convenções de gênero afetam o “conteúdo composicional” (para utilizarmos o jargão bakhtiniano) dos enunciados. No entanto, Maingueneau inclui um novo elemento, não discutido no nosso capítulo a propósito das várias tradições de pesquisa em gêneros: o suporte material onde as instâncias do gênero se manifestam. Para o pesquisador, todo “texto é inseparável de seu modo de existência material”, de tal modo que “uma modificação do suporte material de um texto modifica radicalmente um gênero do discurso” (MAINGUENEAU, 2002: 68).

Para dar conta dessa característica, Maingueneau introduz o conceito de *mídiun*, que se refere tanto ao “modo de manifestação material dos discursos” quanto ao seu “modo de difusão”. O que importa para Maingueneau é o fato de que “o mídiun não é um simples ‘meio’ de transmissão do discurso, mas que ele imprime um certo aspecto a seus conteúdos e comanda os usos que dele podemos fazer” (MAINGUENEAU, 2002: 71). A idéia não é nova e remonta a McLuhan; no entanto, parece ser nova no contexto da teoria dos gêneros do discurso. Um exemplo típico, citado pelo autor, seria a transformação sofrida pelo discurso político a par daquelas ocorridas no seu meio material de enunciação: da reunião eleitoral do século XIX ao debate televisivo contemporâneo, passando pelas “conversas ao pé do rádio” da primeira metade do século passado. Tanto as atitudes do enunciador quanto as expectativas da audiência foram profundamente transformadas em cada um dos cenários, afetando igualmente o modo como o discurso se constrói, de tal modo que torna possível a Maingueneau dizer que “as modificações das condições ‘materiais’ da comunicação política transformam radicalmente os ‘conteúdos’ e as maneiras de dizer, a própria natureza do que se chama ‘discurso político’” (MAINGUENEAU, 2002: 73).

Embora não tenha sido discutida detalhadamente no texto citado, sua aplicação na análise parece-nos indicar que a categoria de *mídiu*m talvez seja excessivamente abrangente, referindo-se um conjunto de elementos da comunicação que poderiam (ou talvez deveriam) ser tratados separadamente. Tomemos um cenário hipotético: A deseja convidar B para um evento qualquer. Para isso, ele pode utilizar inúmeros meios distintos, como por exemplo, uma conversa telefônica (a), uma conversa pessoal (b), um e-mail (c), uma carta (d), etc. Nossa experiência indica que é fácil determinar que o “conteúdo composicional” (para utilizarmos o jargão bakhtiniano) dos enunciados será distinto em cada um dos casos, apesar do propósito comunicativo dos eventos e os interlocutores se manterem constantes. O problema é que não foi apenas o “suporte material” da comunicação que mudou em cada caso: todo o circuito de comunicação foi modificado em cada um dos cenários. Nos cenários (a) e (b), os interlocutores interagem diretamente, ao contrário dos cenários (d) e (c). No cenário (a) apenas a voz intervém, enquanto que em (b) elementos visuais são incorporados, afetando o conteúdo da comunicação. Do mesmo modo, a presença pessoal em (b) e sua ausência em (a) devem afetar igualmente o modo como o diálogo é construído. Por outro lado, A e B podem ser as mesmas “pessoas” em cada cenário; porém, é certo que desempenham “papéis” distintos, com expectativas distintas a propósito da conversação, submetendo-se a regras de interação social distintas em cada caso, etc. O elenco de elementos diferentes envolvidos é bastante grande. Nesse contexto, a utilização de uma noção simplificadora como *mídiu*m parece equivalente à definição “tudo o que participa de uma interação comunicativa e que não é o enunciado verbal”, não agregando conhecimento algum ao problema, já que nada se segue da definição...

Note-se que não estamos negando que as condições materiais em que a interação comunicativa ocorre afetem seu conteúdo; estamos apenas questionando a necessidade da introdução de uma categoria analítica nova para o entendimento do problema de *como* essas condições materiais afetam a comunicação – ou pelo menos a introdução de uma categoria tão abrangente. Nosso argumento consiste em entender que uma categoria analítica qualquer, para ser útil, deve agregar valor de conhecimento à análise – o que não conseguimos encontrar na categoria de *mídiu*m. Em todos os cenários citados no exemplo, mas também em qualquer cenário imaginável, não é apenas a alteração do “modo de manifestação material dos discursos” e do seu “modo de difusão” que tornam diferentes o conteúdo dos enunciados. Isso é igualmente verdadeiro para os exemplos de “discurso político” utilizados para validar a categoria de *mídiu*m. O que distingue um discurso político pela TV, pelo rádio, num comício ou numa reunião não é apenas o “suporte material” dos enunciados, mas *todas* as condições

em que o discurso ocorre. Nesse último caso, por exemplo, sequer a *audiência* de cada um dos discursos é a mesma, ainda que imaginemos que eles sejam ouvidos pelas mesmas “pessoas”. Não apenas o caráter voluntário do comício ou da reunião, como também o fato de serem atividades *coletivas* altera as condições em que o discurso ocorre, afetando seu conteúdo. O próprio Maingueneau constata isso – e essa constatação não é nem precisa ser suportada por uma nova categoria analítica.

Um outro aspecto que é preciso considerar refere-se ao fato de que “suporte material”, quando aplicado ao texto, não ser uma categoria homogênea e facilmente discernível. Em que medida podemos dizer, como Maingueneau, que “todo texto é inseparável de seu modo de existência material: modo de *suporte/transporte* e de *estocagem*, logo, de *memorização*” (MAINGUENEAU, 2002: 68)? Pensemos num texto concreto, digamos num exemplar de uma crônica. À primeira vista, esse texto só é inseparável do seu “modo de existência” escrito, já que ele não parece ser afetado pelo fato de ser publicado num jornal, numa revista, num livro ou mesmo num *site* na Internet. Para que a proposição de Maingueneau seja válida nestes casos, apenas a codificação escrita da linguagem deveria ser considerada como o modo de suporte/transporte/estocagem/memorização do texto, caso contrário não poderemos afirmar que ele é “inseparável” do seu suporte físico. No entanto, parece plausível considerarmos livros e jornais, por exemplo, pelo menos como parte do *suporte* material para o transporte/estocagem/memorização de textos de tipos diferentes; pelo menos eles desempenham essa função.

Por outro lado, se o “conteúdo composicional” (a organização textual) da crônica do nosso exemplo não é afetado pelo fato de ser publicado num livro, numa revista, num jornal ou num *website*, as “condições materiais” da sua circulação e do seu consumo foram afetadas – pelo menos em alguma medida! As condições de leitura e possivelmente as atitudes e expectativas do leitor parecem ser diferentes em cada caso. Neste caso, se o “texto” é o mesmo, o *evento comunicativo* em que ocorre não é o mesmo em cada caso. Para que a proposição de Maingueneau seja aceitável aqui, é necessário que o conceito de *texto* se expanda para além das suas fronteiras estritamente “verbais”. Ora, essa operação interpretativa não nos parece necessária, já que dispomos da categoria de evento comunicativo para dar conta dessa dimensão da *circulação* do texto. Novamente, parece-nos que a categoria de *mídiun* (ou a equivalente *suporte*) é excessivamente abrangente para agregar valor aos nossos conhecimentos a propósito do problema dos gêneros.

A pergunta que nos move – o que é a dimensão do *suporte material* do texto e *como* ela afeta sua construção e consumo – ainda não foi, porém, respondida. Ainda que concordemos

com a afirmação de Maingueneau de que as condições materiais em que a interação comunicativa ocorre afetam seu conteúdo e condições de produção, circulação e consumo, ainda não fomos capazes sequer de conceituar adequadamente o que poderia ser esse *suporte*, pelo menos como uma categoria distinta daquelas que vimos discutindo ao longo das últimas seções. Devemos, pois, prosseguir em outra direção.

2.2. Canais (marcianos) e veículos de comunicação

Ainda que a tradução consultada da obra de Maingueneau se utilize de um novo termo para a noção de suporte, aparentemente no intuito de lhe conferir sentido técnico, o esforço não parece ser bem sucedido em precisar a noção. Pelo que pudemos notar, de uma forma geral a utilização da expressão “suporte” (ou alternativas, como *mídium*) tem como finalidade estabelecer uma fronteira mais ou menos nítida com a noção de *canal de comunicação*, cara às teorias convencionais da comunicação. Talvez uma reflexão sobre o problema nos traga auxílio aos objetivos deste capítulo.

Num interessante artigo, Bonini (2003) questiona a noção de *veículo* como componente da noção de canal de comunicação, introduzida por Berlo (1979) na sua tentativa de adaptar a teoria de Shannon à comunicação humana, assinalando a incoerência dessa noção com as observações. Dando como exemplo o jornal, Bonini demonstra que é muito difícil – senão impossível – separar a noção de veículo do conceito de gênero. Como na teoria de Berlo o gênero é entendido como parte do código e o veículo como parte do canal, as duas noções seriam incompatíveis. Para além da análise efetuada sobre as características sógnicas de um jornal, o núcleo conceitual da argumentação de Bonini pode ser assim resumido:

Sendo o gênero um material sógnico relativamente compartilhado entre os membros de determinada comunidade, e (...) tendo o jornal a mesma propriedade cultural e sógnica do gênero, ele seria também um tipo de gênero e não um veículo propriamente. Se assumirmos o ponto de vista de que a teoria da informação (Shannon e Weaver) não é uma moldura adequada para se pensar a noção de gênero, o jornal também pode ser entendido como um gênero construído a partir do encaixe de outros, um hiper-gênero (BONINI, 2003, : 77-78).

O primeiro dos argumentos de Bonini em favor da proposição consiste em afirmar a incapacidade das teorias de base comunicacional sustentarem a distinção entre veículo e gênero a partir da distinção entre canal e código. Em outras palavras, tais teorias (na verdade, a teoria *de* Berlo) situam o que chamamos “veículos de comunicação” no domínio conceitual

da noção de *canal*, enquanto que os exemplares de mensagens que situamos no interior da noção de gêneros estariam vinculados à noção de *código*. No intuito de refutar a teoria, Bonini procura demonstrar – com sucesso – a constituição sónica e convencional do que se costuma chamar “veículo” (de comunicação).

De saída, é conveniente assinalarmos que concordamos que a teoria da informação de Shannon (mas não de Weaver, apenas um comentarista e divulgador) “não é uma moldura adequada para se pensar a noção de gênero”, mas não pelas razões apontadas por Bonini. Ela não é adequada à Teoria dos Gêneros por que se destina a um *outro* propósito. Ela se dedica, utilizando os comentários de Weaver, a determinar “com que exatidão [fidelidade] podem ser transmitidos os símbolos de comunicação” (WEAVER, 1975: 4), isto é, ela pretende resolver problemas tais como determinar como é possível quantificar a informação, calcular a capacidade de um canal de comunicação, minimizar as perturbações causadas por fontes de ruídos, etc. Daí Shannon assinalar que “o problema fundamental da comunicação é reproduzir num dado ponto uma mensagem selecionada em outro ponto de modo tão exato quanto possível” e concluir, coerentemente, que “os aspectos semânticos da comunicação são irrelevantes para o problema da engenharia” (SHANNON, 1948, 1). Ora, por definição, para a Teoria dos Gêneros, os “aspectos semânticos da comunicação” não podem ser irrelevantes; daí a teoria de Shannon não ser aplicável ao problema, senão de modo extremamente limitado e cauteloso.

A teoria de Shannon não é uma Teoria de Tudo da comunicação. No entanto, Weaver sugere que ela pode se tornar uma, se generalizada (ver WEAVER, 1975: 26 e seguintes). As generalizações propostas por Weaver, contudo, parecem-nos ingênuas; ele propõe, por exemplo, adicionar novas caixas ao clássico diagrama de Shannon, como a “Semântica do Receptor” ou “Ruídos da Semântica”. Isso, parece-nos evidente, não é o bastante. Não é o que pensa, porém, Berlo (1979). Ao tentar tornar a teoria de Shannon uma Teoria de Tudo, Berlo torna, por exemplo, vaga uma noção precisa – a de canal. Para Shannon, “canal é simplesmente o meio utilizado para transmitir o sinal do transmissor ao receptor. Pode ser um par de fios, um cabo coaxial, uma banda de radiofrequências, um raio de luz, etc.” (SHANNON, 1948: 2). Trata-se, portanto, de uma noção física. Para Berlo (e isso é explorado por Bonini), um canal deixa de ser simples e torna-se um complexo amálgama de elementos físicos, culturais e lógicos, abrangendo desde os fótons aos jornais e às aulas. Ora, pela definição de Shannon, o papel – um meio físico – pode ser utilizado como canal de comunicação por onde trafegam sinais gráficos; o jornal, por outro lado, é um artefato cultural (um conjunto de mensagens) codificado a partir de sinais gráficos e transportado pelo papel.

A “generalização” de Berlo, portanto, não passa de uma extrapolação vaga e imprecisa, alterando os domínios de conceitos bem delimitados.

A nosso ver, Berlo simplesmente “embarca na canoa (furada)” dos próprios “veículos de comunicação”, que procuram se caracterizar enquanto “meros” portadores (ou transportadores) de informação. Trata-se de uma operação ideológica de naturalização, que quer ter como resultante esconder o fato de que se há algo que um jornal “transporta” são as opiniões de seus proprietários, editores e jornalistas a propósito dos assuntos abordados; essa operação tenta nos convencer de que o “veículo de comunicação” é passivo (como um canal de comunicação), limitando-se a servir de meio através do qual uma outra coisa trafega; ela esconde a natureza *ativa* do “veículo”, que *escolhe* o conteúdo que publica (ou “veicula”) dentre um sem-número de conteúdos possíveis. Portanto, a noção de *veículo* esconde a natureza daquilo que designa, não devendo ser incorporada ao acervo de uma teoria que procura compreender aquilo que comumente chamamos “veículos de comunicação”.

Nosso argumento, portanto, consiste em afirmar que é a teoria de base comunicacional *de Berlo* (e não as teorias de base comunicacional em geral) que é incapaz de sustentar “a distinção entre veículo e gênero a partir da distinção entre canal e código”. Ao contrário, podemos imaginar teorias de base comunicacional que afirmem que veículo e gênero *não* são distintos, tomando por base exatamente a distinção entre canal e código. Afinal, por que as noções de “veículo” e “código” *deveriam* ser distintas? Por que “veículos” como jornais e

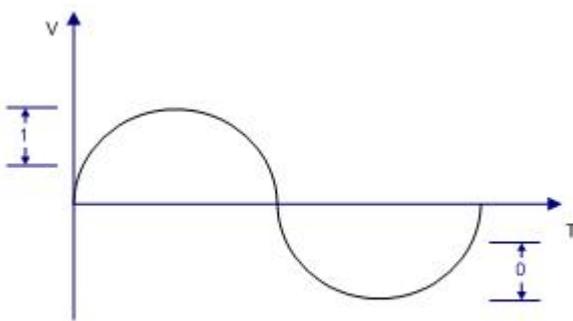


Ilustração 12: Uma codificação simples.

revistas não poderiam ser uma subcategoria de gêneros do discurso, aliás como sustenta Bonini? E acreditamos que para afirmar isso *não* é necessário negarmos categorias como “canal” e “código”. O fato daqueles artefatos que chamamos “veículos” serem convenções (ou mais precisamente *codificações*) que também “transportam” outras convenções (ou gêneros que suportam outros gêneros) não afeta o argumento. O que há de tão estranho em codificações que codificam outras codificações? Eco chama esse fenômeno de *hipercodificação* (ECO, 1980: 117 e seguintes); em outra vertente da teoria (tecnicista, no linguajar de Bonini), Ashby chama a atenção para a *ubiquidade da codificação* (ASHBY, 1957: 140 e seguintes). Para abordar corretamente o problema devemos, portanto, discutir a natureza das codificações.

Ora, uma das características das codificações é que elas podem ser arbitrárias e, portanto, podemos codificar diferentemente os mesmos significados. Suponhamos que seja

possível controlar de modo regular os valores V alcançados por um sinal físico qualquer e que esses V valores ao longo de um intervalo de tempo arbitrário T obedeam a uma função contínua e definida como a exibida na Ilustração 12. Podemos arbitrar que a faixa de valores no topo do gráfico representem o sinal lógico (binário) 1, enquanto que a faixa no vale representa o sinal binário 0. Com base nessa codificação muito simples, podemos construir todo um sistema de representação baseado na matemática binária. O importante aqui é que o sistema é indiferente ao sinal físico considerado. Podemos utilizar, por exemplo, a variação de tensão numa corrente elétrica e arbitrar como sendo o valor binário 0 qualquer tensão entre 0 e 1 volts, enquanto que tensões entre 2 e 5 volts são interpretadas como o valor binário 1: este é o caso dos circuitos digitais dos computadores. Mas podemos, por outro lado, utilizar quaisquer outros sinais físicos, como, por exemplo, a amplitude de ondas eletromagnéticas, a vazão de uma corrente de água, variações de tom, altura e sonoridade das ondas sonoras, etc.

Podemos agora empreender uma segunda codificação e arbitrar que os valores binários sejam interpretados em blocos de 8 dígitos; isso significa que teríamos 256 valores distintos à disposição para a nossa codificação. Por convenção, definimos então que o intervalo discreto entre 00010110 e 10100111 (a faixa de 97 a 122 na representação decimal) corresponde ao nosso alfabeto. Desse modo, se recebemos a seqüência binária correspondente ao número 164.655.220.502 (na escala decimal), decodificamos a mensagem enviada como sendo a seqüência de símbolos alfabéticos b-e-s-t-a. Por outro lado, podemos igualmente arbitrar blocos maiores para a escala digital, digamos 32, 64 ou 128 dígitos binários (bits) – valores limitados somente pela tecnologia disponível.

Com base nessa codificação muito simples podemos elaborar codificações (e hipercodificações) cada vez mais complexas, correspondendo a níveis de generalidade e abstração cada vez *menos* elevados, isto é, com finalidades cada vez mais especializadas. Por exemplo, alguns engenheiros da Internet elaboraram uma codificação destinada a estabelecer as condições de troca de sinais entre dois computadores conectados em rede – o protocolo TCP-IP (*Transfer Control Protocol – Internet Protocol*), que governa todas as transmissões nas redes conectadas à Internet. Utilizando sinais transmitidos segundo o protocolo TCP-IP, isto é, tomando como portadora dos seus sinais aquele protocolo, outros engenheiros elaboraram uma outra codificação, destinada a sustentar a transferência de *hiperdokumentos* entre dois computadores – o protocolo HTTP (*Hipertext Transfer Protocol*). Para codificar os dados transmitidos pelo protocolo HTTP, mais precisamente aqueles sinais de retorno das requisições do protocolo, outros engenheiros elaboram uma outra codificação – as convenções HTML (*Hipertext Markup Language*), um conjunto de convenções destinadas a interpretar os

hiperdocumentos. Nesse cenário, as codificações vão se superpondo em camadas, as inferiores utilizadas como portadora dos sinais codificados pelas superiores. Codificações dentro de codificações; camadas de transporte construídas sobre outras camadas de transporte.

Até esse ponto, as fronteiras entre as diferentes codificações são bastante claras, simplesmente porque estamos no domínio da técnica e da engenharia: as codificações se destinam ao uso exclusivo de dispositivos computacionais. A partir daí, no caso da Internet, contudo, as novas codificações construídas a partir das anteriores tornam-se difusas e vagas, não impositivas e ambíguas. Estamos no domínio das práticas sociais, da interação humana, do discurso, do *uso* humano das coisas humanas – e não no domínio da engenharia dos sistemas computacionais. No entanto, se as codificações se tornam imprecisas, nem por isso deixam de ser *codificações*: um conjunto (vago e impreciso) de regras (mais ou menos impositivas) passa a governar, por exemplo, a circulação de diferentes tipos de mensagens através das codificações HTTP e HTML. Talvez fosse o caso passarmos a utilizar a expressão “convenções” em lugar de “codificações”, procurando assinalar lexicalmente essa diferença no *grau* de precisão das regras segundo as quais os novos “códigos” são construídos. No entanto, o fato de se tratarem de produtos de atividades sociais mais ou menos ao acaso e não do consenso entre alguns poucos engenheiros não muda a *natureza* de código desses artefatos sociais.

Por que a imensa maioria dos *websites* é dividida em seções temáticas e, normalmente, cada uma delas é acessível diretamente da primeira página por um conjunto de *links* (conhecidos como *menus*)? Por convenção. Por que se espera que alguns *sites* tematizem principalmente as experiências quotidianas, as impressões, as idéias, os sentimentos dos seus autores? Por que simplificamos e/ou alteramos inúmeras convenções da escrita em eventos de interação direta na Internet? Por que os *links* são frequentemente indicados por texto sublinhado na cor azul? Por que muitas mensagens de correio pessoal na Internet são concluídas com a seqüência de símbolos “[s]”? Por que alguns *websites* incorporam regularmente textos de diferentes tipos sobre diferentes assuntos, de crônicas a notícias, de confissões a material publicitário? Em todos os casos, a resposta é (enganosamente) simples: por convenção – por convenções produzidas e assumidas ao longo das práticas sociais em torno desses eventos e artefatos comunicativos, mas ainda assim *convenções*. Como se pode notar, a partir das poucas premissas que viemos discutindo, chegamos naturalmente a algum tipo de noção de “gênero” ou categoria de artefatos de linguagem.

É interessante observar que, no domínio da engenharia, o código é inteiramente indiferente à natureza do canal. Aquela codificação arbitrária vista na Ilustração 12 e utilizada

como ponto de partida para todas as demais pode ser elaborada a partir de inúmeros tipos de suportes físicos (alguns ali citados) e transmitida por inúmeros tipos diferentes de canais – e ainda assim preservar as mesmas características de codificação e transmitir as mesmas mensagens. No caso dos computadores digitais, assinalamos que se convencionou utilizar as variações na tensão elétrica de um circuito para codificar os dois símbolos elementares daquele código. No entanto, podemos perfeitamente imaginar a mesma codificação utilizando, por exemplo, o volume de água circulante numa rede hidráulica e construir uma rede mundial de computadores em torno dela. Pode não ser prático, mas é teoricamente possível. Podemos imaginar codificações baseadas em reações químicas e, a partir delas, elaborar todo um sistema de comunicação baseado no cheiro, onde um determinado padrão olfativo corresponde ao 0 binário e outro ao 1 binário. Não há razão plausível para que todas as mensagens passíveis de serem elaboradas e transmitidas num sistema não o possam igualmente ser no outro. Nesse nível básico, a codificação é inteiramente independente do canal de comunicação e da natureza física dos sinais.

Um exemplo simples será o bastante para ilustrar o que pretendemos. Imaginemos um outro sistema de comunicação, este baseado no contraste entre formas regulares pintadas contra um fundo de cor nitidamente distinto. Um sistema de codificação simples pode consistir em utilizar pequenas linhas retas verticais de espessura variável, separadas por espaços da cor do fundo. As linhas mais finas (digamos até 1 milímetro de espessura) podem representar o valor binário 0; as mais grossas (entre 2 e 3 milímetros), o valor binário 1. Essa codificação é rigorosamente análoga à descrita nos parágrafos anteriores, podendo, portanto, transmitir rigorosamente as mesmas mensagens que aquela, com a vantagem de ser facilmente perceptível ao olho humano, não requerendo dispositivos de decodificação e recodificação dos sinais para ser utilizável por seres humanos. Do mesmo modo, as mensagens que podemos transmitir não são afetadas pelo canal de comunicação escolhido. Podemos, por exemplo, utilizar o papel como meio de armazenagem desses sinais, transmitindo-os amarrados a patas de pombos⁵, novamente construindo uma rede de computadores análoga à

⁵ Parte dessa fantasia divertida foi publicada em 1999 na mais séria das instituições reguladoras da Internet, o IETF (*Internet Engineering Task Force*), responsável pela criação e manutenção dos principais padrões utilizados na rede, através da RFC (*Request for Comments* – o mesmo gênero discursivo utilizado para discutir e definir os padrões técnicos da rede) 2549,

Internet utilizando esse sistema. Há uma implicação evidente aqui: se as codificações são indiferentes à natureza física dos sinais e do conteúdo das mensagens, isto é, se podemos mapear qualquer significado a qualquer sinal, segue-se que a língua natural não precisa ser o único mecanismo de codificação utilizado na comunicação humana.

Isso é o mesmo que dizer que tudo pode significar para esse *Homo semioticus*. De fato, as interações humanas são conduzidas não apenas pela linguagem natural, mas também por sistemas de codificação mais fracos, para-lingüísticos ou simplesmente não-lingüísticos. Se tais componentes da comunicação humana fazem ou não parte do fenótipo da espécie não é relevante para o que pretendemos: basta reconhecer que, à medida que as técnicas se modificam, introduzindo novas tecnologias de transmissão da informação, tais componentes se tornam cada vez mais decisivas para as práticas sociais comunicativas – em especial para aquelas incidentes no ciberespaço.

Argumentamos que as codificações são independentes do canal e da natureza física dos sinais, quando tratamos do nível da engenharia das codificações. No entanto, no domínio das práticas sociais construídas em torno dele as coisas não são tão simples assim. Aqui as convenções *podem* não ser inteiramente indiferentes à natureza do suporte físico utilizado pelas codificações de domínio mais baixo. Isso é bastante diferente da afirmação de Maingueneau de que “*todo texto é inseparável de seu modo de existência material*”. Seria o caso de afirmarmos, quando muito, que *alguns* textos o são... Por exemplo, Xavier, em sua tese de doutoramento, ao postular um novo *modo de enunciação* – o digital –, chama a atenção para o fato de que ele “só se realiza na tela. É o computador eletro-digital que promove o encontro entre todos os modos de enunciação no Hipertexto e os faz ganhar visibilidade na tela” (XAVIER, 2002: 99). Na verdade, Xavier não percebe que o *mouse* e outros dispositivos de *entrada* (o monitor é um dispositivo de *saída*) são igualmente essenciais ao hipertexto, como discutimos na seção 1.5. *O hipertexto e as teias do ciberespaço*, onde assinalamos o caráter *inter-ativo* do hipertexto. Isso, porém, não afeta a força do argumento. Isso ocorre porque essas convenções se produzem no interior de práticas sociais ancoradas em artefatos técnicos já existentes: o “demiurgo” se utiliza dos materiais

que encontra à mão... Podemos, é certo, imaginar uma Internet baseada no cheiro; porém, não somos capazes de imaginar as convenções de domínio discursivo construídas a partir dela simplesmente porque não temos experiência alguma nesse tipo de comunicação e, portanto, sequer podemos imaginar as práticas sociais construídas em torno desse sistema de comunicação... Nesse sentido, as convenções, os códigos no domínio das práticas sociais nunca são inteiramente arbitrários, como as codificações no domínio da engenharia. Podemos dizer que tais códigos são *relativamente* motivados (pelos meios à disposição dos agentes sociais), modalizando (e restaurando) a afirmação de Maingueneau.

Nos parágrafos anteriores imaginamos um sistema de comunicação baseado em códigos de barras e transmitidos por pedaços de papel atados a patas de pombos. Mesmo que alteremos o canal de comunicação utilizado, substituindo-o agora por folhas retangulares de papel branco encadernadas em conjunto e transmitidas por portadores humanos, ainda assim é possível transmitir rigorosamente as mesmas mensagens do sistema anterior, incluindo imagens (embora estas sejam codificadas em binário e não nas formas convencionais da fotografia). Por que então aqueles artefatos chamados “livros” *não* transmitem quaisquer mensagens que podemos transmitir fazendo uso das convenções da escrita em língua natural? Por que, por exemplo, não escrevemos bilhetes de amor em livros e os transmitimos para a(o)s namorada(o)s através de um amigo comum, enquanto que romances epistolares são admissíveis nesse “suporte”? Porque “livros” não são folhas retangulares de papel branco encadernadas em conjunto e transmitidas por portadores humanos (um *canal de comunicação*), mas *artefatos culturais* do domínio das codificações, artefatos construídos por práticas sociais envolvendo os meios técnicos disponíveis para os agentes sociais (por exemplo: pergaminho e escrita cursiva vazada a bico de pena por monges beneditinos, papel impresso por dispositivos mecânicos, etc.). Utilizando uma metáfora, podemos dizer que “livros” são signos – e sua natureza de signo fica bem clara quando alteramos o “canal de comunicação” de um romance e o publicamos no formato de arquivo do Adobe Acrobat, cujo aplicativo de exibição chega mesmo a “imitar” a folha branca de papel impressa como na prensa de Gutenberg, muito embora esta metáfora visual seja desconfortável para leitura na tela de um dispositivo computacional. Do mesmo modo, uma placa de bronze não é simplesmente um “suporte” para alguns discursos, mas *parte* das convenções de um gênero discursivo (digamos uma placa comemorativa, o “texto” de um evento comunicativo chamado “inauguração”, cujo propósito parece ser comunicar às infortunadas gerações futuras como eram empreendedoras certas autoridades públicas do passado).

Ora, o saldo que julgamos mais importante nessa discussão consiste no fato de não precisarmos da categoria de *suporte* (ou veículo) ou qualquer categoria intermediária entre *canal* e *código*, entre o meio de transmissão físico e os artefatos sígnicos para apreender e interpretar os artefatos da comunicação humana. Argumentamos que o fato de alguns códigos permitirem a codificação de outros a partir deles é meramente acidental, já que, em princípio, todo código pode ser hipercodificado e reutilizado para uma nova finalidade. Por esse argumento, gêneros discursivos são códigos fracos, elaborados a partir das práticas sociais coletivas, relativamente motivados pelos meios materiais disponíveis, códigos que não apenas dão conta dos artefatos produzidos no seu interior como também dos papéis desempenhados pelos agentes sociais envolvidos e das condições de produção, circulação e consumo desses artefatos. Utilizando essa linha de raciocínio, falar em “veículos de comunicação” como parte do domínio conceitual do canal de comunicação (mesmo que para desqualificar essa última noção) é pura ilusão de ótica: como os legendários canais marcianos, vemos o que queremos ver e não o que está lá...

Com base nesse raciocínio e numa teoria comunicacional que mantém a distinção entre canal e código, é possível concluir que não podemos realmente sustentar uma distinção entre veículo e gênero... Na verdade, *não queremos* sustentar essa distinção. É possível, ao contrário, demonstrar a indistinção entre as duas categorias, partes de um mesmo domínio de problemas: os chamados “veículos de comunicação” devem ser vistos, portanto, como gêneros do discurso. Por outro lado, esse mesmo raciocínio nos conduziu à conclusão que, do ponto de vista das codificações, é possível traçar uma linha contínua e fracamente distinta entre os canais da comunicação humana e os artefatos sígnicos, linha contínua constituída pelas codificações produzidas no interior das práticas sociais discursivas. Do nosso ponto de vista, o uso de uma categoria analítica como “suporte” para caracterizar essa linha contínua não é necessário nem aconselhável, já que agrega pouco ou nenhum valor de conhecimento ao problema.

2.3. O lugar dos gêneros digitais no ciberespaço

Com a discussão travada anteriormente, supomos ter recolhido ferramental conceitual suficiente para a caracterização de gêneros semióticos (emergentes) incidentes no

ciberespaço. Com base naquelas idéias, essa operação deve envolver os seguintes passos bem definidos:

- i Identificar os eventos comunicativos candidatos a gêneros num dado campo de atividade humana;
- ii Caracterizar a situação comunicativa desses eventos candidatos, o que implica tanto em determinar os agentes comunicativos envolvidos quanto as *personae* discursivas que assumem durante o evento;
- iii Determinar que convenções e regras controlam o comportamento dos produtores e as expectativas dos consumidores dos eventos, definindo com isso tanto as condições sob as quais o evento é produzido quanto aquelas sob as quais é recebido e interpretado.

Isso significa que devemos, preliminarmente, elaborar, com base nas idéias discutidas nas últimas seções, um modelo de identificação dos eventos comunicativos incidentes no ciberespaço. No nosso caso, queremos adicionalmente determinar o lugar do acesso aos portais de interesse geral dentro dessa gama de eventos comunicativos. É o que trataremos a seguir.

2.3.1. Um modelo multidimensional para o ciberespaço

Não há dúvida de que a Internet constitui um campo de atividade comunicativa complexo. No entanto, para simplificar o problema, podemos imaginar que um dado evento comunicativo na Internet ocorra num ciberespaço multidimensional heterogêneo e indagar: qual a situação comunicativa em que ocorre este evento? Que agentes comunicativos estão envolvidos (locutores, interlocutores e *outros* discursivos)? Que *personae* discursivas são construídas ao longo da sua ocorrência? Se as mesmas respostas – ou pelo menos respostas suficientemente próximas entre si – puderem ser obtidas para uma quantidade qualquer de diferentes eventos, isto é, se o evento comunicativo for *repetível*, então haverá boas possibilidades de que tais eventos ocorram no interior de (ou como produtos de) gêneros semióticos.

Trata-se de determinar agora se tais respostas devam ser inerentemente qualitativas ou se, ao contrário, podem ser encontradas variáveis quantificáveis para representar tais

respostas. Em outras palavras, se nossas perguntas puderem ser feitas de tal modo que suas respostas possam ser representadas, mesmo que aproximada e arbitrariamente, por funções que variem de algum modo uniforme, então poderemos quantificar as diversas *dimensões* do ciberespaço de tal modo que isso represente mais que uma metáfora. Desse modo, um evento comunicativo no ciberespaço poderia ser definido pelo valor dessas variáveis, isto é, poderíamos concebê-lo como um *ponto* e localizá-lo no ciberespaço multidimensional por suas *coordenadas* específicas. Esse artifício tornaria mais preciso o raciocínio do parágrafo anterior. Suponhamos que possam ser necessários apenas dois valores para plotar qualquer evento comunicativo no ciberespaço; isto é, suponhamos um ciberespaço plano. Uma configuração tal como a exibida na Ilustração 13 seria a esperada para uma topologia

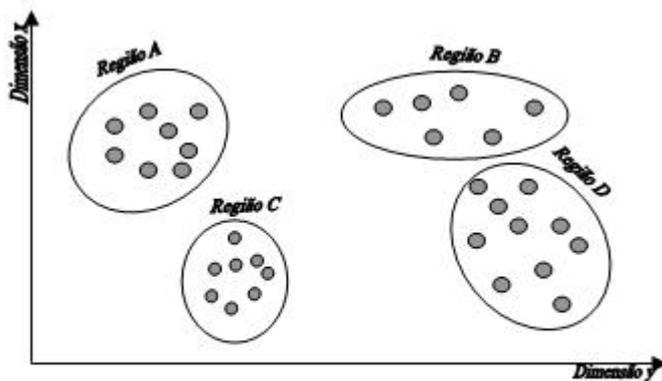


Ilustração 13: Ciberespaço plano com topologia heterogênea

heterogênea onde as irregularidades locais correspondem a famílias de eventos comunicativos tais como os gêneros semióticos. Nesta configuração, a utilização de diferentes gêneros semióticos poderia produzir eventos comunicativos que tenderiam a se agrupar em determinadas regiões do ciberespaço.

Ora, essa representação deve corresponder de algum modo ao ciberespaço, já que tais regiões podem ser intuitivamente inferidas. A troca de e-mails entre dois usuários parece corresponder a uma situação de comunicação distinta da envolvida quando esses mesmos dois usuários conversam diretamente através de um software de *messaging* como o Microsoft Messenger, e essa distinção parece transcender o simples retardo na interação introduzido pelo correio eletrônico: convenções distintas são seguidas; construções frasais diferentes são utilizadas; fórmulas de etiqueta diversas são empregadas, etc. Assim, a imagem de um ciberespaço multidimensional com topologia heterogênea parece ser a mais adequada à descrição desse campo da atividade humana.

Um saldo evidente da discussão anterior sobre gêneros semióticos é que as regiões demarcadas na Ilustração 13 não podem ser bem delimitadas como indicado ali. Se concordarmos com a argumentação a propósito do problema da prototipicidade, os limites entre as regiões demarcadas deve ser difuso. Isso significa que haveria incerteza sobre como

categorizar eventos fronteiriços. Isso não afeta a essência do raciocínio, apenas altera o modo como devemos conceber a interpretação das irregularidades locais observadas.

Até o momento, porém, estamos apenas “postulando” o que já sabemos existir – afinal, todos os pesquisadores que abordaram o problema dos gêneros semióticos na Internet concordam com a existência de algo que podemos chamar *gêneros digitais*. Para que a abordagem proposta seja algo mais que trivial, ela deve acrescentar algum elemento novo às descrições atuais.

Em sua apresentação do problema, Marcuschi inventaria doze gêneros digitais emergentes, a saber: *e-mail*, *chat* em aberto, *chat* reservado, *chat* ICQ, *chat* em salas privadas, entrevista com convidado, *e-mail* educacional, aula *chat*, vídeo-conferência interativa, lista de discussão, endereço eletrônico e *blog*. Do inventário são explicitamente removidos os *sites*, jogos interativos e MUDs (*Multi-User-Dungeon*). Para Marcuschi, a *home page* (isto é, a primeira página de um *site*) “não passa de um ambiente específico para localizar uma série de informações, operando como um suporte e caracterizando-se cada vez mais como serviço eletrônico” (MARCUSCHI, 2004: 20). Os jogos interativos são considerados como “suportes para ações complexas envolvendo vários gêneros na sua configuração” (MARCUSCHI, 2004: 26) e igualmente excluídos da descrição. Por fim, “no caso dos MUDs temos um tipo de relação irreal, relação com a fantasia e não com seres reais e trata-se de um jogo” (MARCUSCHI, 2004: 31-32), o que faria com que também não tivesse direito ao *status* de gênero digital.

O inventário parece no mínimo discutível: por um lado, a introdução do endereço eletrônico (endereço de *e-mail* e os URLs – *Uniform Resource Locators* – utilizados para identificar *sites*) parece dever-se exclusivamente a seu caráter convencional, não havendo indicações de corresponder a uma situação comunicativa efetivamente encontrável no ciberespaço (para não falar num *evento* de comunicação); por outro lado, a especialização dos *chats* parece ser excessiva, já que os diversos tipos inventariados parecem corresponder a conjuntos de eventos comunicativos não inteiramente disjuntos. Por outro lado, pode-se legitimamente perguntar a razão de se incluir no inventário o endereço de um recurso e não o próprio recurso, no caso, os *sites* Internet.

Poderíamos argumentar indefinidamente em torno do inventário proposto. No entanto, parece-nos mais eficaz procurar determinar se existe outra descrição satisfatória para o lugar dos vários eventos comunicativos do ciberespaço tal que seja possível prever a existência e as características dos vários gêneros digitais. Essa descrição deve basear-se nas indagações implícitas no quadro teórico levantado anteriormente e satisfazer aos requisitos anteriormente

estabelecidos, a saber: utilizar-se tão somente de grandezas homogêneas tais que permitam *quantificar* a descrição, ainda que aproximada e arbitrariamente, de modo a assegurar-lhe alguma precisão. Devemos buscar, portanto, definir o que e quais seriam as *dimensões* do ciberespaço.

A partir do levantamento de Marcuschi, selecionamos um pequeno conjunto de eventos comunicativos para análise. A seguir, formulamos hipótese preliminar a propósito dos – possíveis – gêneros digitais aos quais tais eventos estariam relacionados. Os resultados estão ilustrados na Tabela 3. Como se pode notar, o principal critério que nos levou a formular a hipótese do gênero associado foi o *software* que fornece suporte ao evento. Como hipótese preliminar, esse critério não deve ser inteiramente descabido, já que na Internet todos os eventos comunicativos se realizam com base em suporte de *software*.

Evento comunicativo	Gênero relacionado
Conversa bilateral	Messaging
Conferência bilateral	Vídeo-conferência (individual)
Conferência multilateral	Vídeo-conferência
Conversa reservada	Chat (reservado)
Troca de e-mail	E-mail
Busca de informação	Search engine
Interação lúdica	Jogo interativo
Entrevista	Entrevista por chat
Aula virtual	Chat (educacional), E-mail (educacional)
Conversa em grupo aberto	Chat
Discussão coletiva	E-mail (fórum), newsgroups
Multi-interação lúdica	MUD's, adventures
Leitura de diário	Blogs
Acesso a um site	Sites, home pages
Acesso a um portal	Portal

Tabela 3: Eventos comunicativos de interesse no ciberespaço.

Note-se também que o mesmo *software* pode suportar diferentes classes de eventos comunicativos; reciprocamente, estamos considerando que diferentes eventos podem ser suportados pelo mesmo *software*, sem alterações substanciais nas suas características. Tais suposições não são absolutamente evidentes; portanto, devemos assinalar que tais hipóteses deveriam ser submetidas a teste empírico.

Alguns critérios satisfazem as condições estabelecidas na seção anterior e podem ser utilizados para definir um ciberespaço multidimensional, em conformidade com as noções de gênero discutidas. O primeiro deles pode ser chamado dimensão da interação. Podemos

tipificar as interações como variando continuamente das interações diretas às indiretas, identificando uma região intermediária – a das interações mediadas. Interações diretas envolvem troca imediata de turnos e locução orientada diretamente a interlocutores bem definidos. Os diversos tipos de conversas e conferências, além da troca de *e-mail*, são interações desse tipo. As interações mediadas implicam trocas de turnos e interação contínua, mas são intermediadas por uma terceira parte (um mediador, um professor, um entrevistador, um dispositivo computacional, uma personagem, etc.), como, por exemplo, as entrevistas e diversos tipos de interações lúdicas. Interações indiretas não implicam em troca de turnos nem interlocutores bem definidos, como, por exemplo, a leitura de um diário e o acesso aos *sites*.

Podemos tipificar uma segunda dimensão – a locução – e interpretá-la como variando continuamente de locuções 1–1 até locuções n–n, intermediadas por locuções 1–n. As locuções 1–1 são bilaterais, envolvendo apenas dois participantes. As locuções 1–n envolvem a interação entre um locutor e uma platéia, onde os membros da platéia podem também interagir com o locutor, mas não entre si. É o caso, por exemplo, de uma entrevista. As locuções n–n são multilaterais, envolvendo um número indefinido de participantes que interagem mutuamente, como nos *chats* e conferências.

Num ciberespaço bidimensional poderíamos dividir cada dimensão em três regiões vizinhas de modo a construir uma representação de nove quadrantes e tentar plotar em cada um deles, numa primeira aproximação, os eventos comunicativos listados na Tabela 3. Esse esforço poderia resultar numa configuração como a exibida na Tabela 4, bastante intuitiva e simples. Note-se que plotamos a dimensão da interação no eixo horizontal Y e a dimensão da locução no eixo vertical Z e numeramos os quadrantes resultantes da esquerda para a direita. Como se pode notar, uma representação bidimensional do ciberespaço é insuficiente para distinguir todas as famílias de eventos comunicativos selecionadas e, portanto, uma terceira dimensão é necessária.

00 conversa bilateral, conferência bilateral, conversa reservada	01 busca de informação, interação lúdica	02
03 troca de e-mail	04 entrevista, troca de e-mail	05 leitura de um diário, acesso a um site
06 conferência multilateral, conversa em grupo aberto	07 aula virtual, discussão coletiva, multi-interação lúdica	08 acesso a um portal

Tabela 4: Distribuição de eventos comunicativos num ciberespaço bidimensional.

Suponhamos que o modo como a audiência é concebida possa ser caracterizado como uma dimensão contínua variando de uma audiência conhecida a uma audiência construída (convencionada) e terminando numa audiência totalmente imaginada. Na correspondência pessoal através do correio eletrônico, por exemplo, a imagem que temos do nosso interlocutor é, geralmente, construída antes e independentemente dessa interação.

Trata-se, portanto, do ponto de vista da interação corrente, de uma audiência conhecida. Isso, porém, não é verdadeiro para o *spam*. Situação similar ao *spam* ocorre quando produzimos um *blog* (ou quando interagimos numa sala de *chat*). Nestes casos, nossa locução é determinada em alguma medida por um conjunto de suposições a propósito dos nossos possíveis interlocutores. Essa situação configura o que chamamos interação com audiência construída. Por fim, numa interação lúdica, como nos *adventures*, por exemplo, interagimos com personagens inteiramente imaginários. As suposições feitas a propósito dessa audiência são, portanto, de ordem inteiramente diversa das anteriores. Uma representação visual desse modelo de ciberespaço pode ser vista na Ilustração . Esse artifício permitiria adicionar várias distinções ao modelo anterior, fazendo com que várias famílias de eventos se deslocassem ao longo do eixo da audiência. Os eventos relacionados ao universo dos *chats* poderiam se distinguir dos outros tipos de conversa: de uma forma geral, os participantes dos *chats* não se conhecem fora do ciberespaço, tendo à disposição apenas as próprias interações para construir

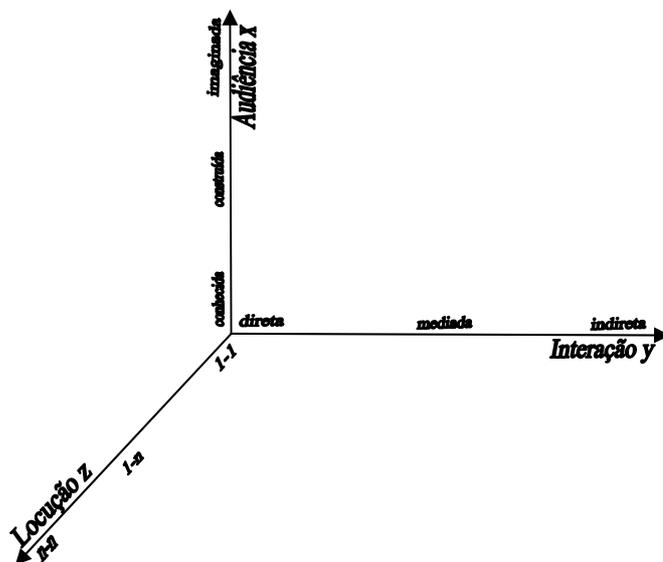


Ilustração 14: O ciberespaço tridimensional.

uma imagem dos seus interlocutores. Além disso, os jogos seriam igualmente diferenciados das famílias próximas, já que pressupõem uma audiência inteiramente imaginária. A Tabela 5, cujos quadrantes são numerados em representação hexadecimal, exprime essas diferenciações introduzidas pelo acréscimo de uma terceira dimensão. Note-se a permanência de algumas indistinções importantes. Note-se também que a ambigüidade referida a respeito da correspondência eletrônica

está ainda expressa na tabela.

00 conversa direta, conferência bilateral	09 conversa reservada	12
01	0A busca de informações	13 interação lúdica
02	0B	14
03 troca de e-mail	0C troca de e-mail	15
04	0D Entrevista	16
05	0E leitura de um diário, acesso a um site	17
06 conferência multilateral	0F conversa em grupo aberto	18
07 aula virtual	10 discussão coletiva	19 multi-interação lúdica
08	11 acesso a um portal	1A

Tabela 5: Distribuição dos eventos comunicativos num ciberespaço tridimensional.

Até o momento, os critérios propostos não foram capazes de distinguir entre a conversa direta e a conferência bilateral, por um lado, e a leitura de um diário e o acesso a um site, por outro, mesmo por hipótese. Quanto aos dois primeiros eventos a distinção é intuitiva, já que tais interações são suportadas por semioses bastante diversas. Poderíamos, talvez, afirmar que, em ambos os casos, os propósitos comunicativos são os mesmos, o que possivelmente explicaria a indistinção – tratar-se-ia da mesma família de eventos comunicativos. No entanto, a *situação comunicativa* onde ocorrem é bastante diversa, para não falar nos recursos utilizados na interação: na conversa direta, sustentada pelos dispositivos de *messaging*, a linguagem natural escrita é central na interação, determinando suas características; na conferência bilateral, sustentada por dispositivos de vídeo-conferência, a linguagem natural oral desempenha esse papel em colaboração com outras semioses como gestos, imagem, voz, etc. Logo, trata-se de eventos comunicativos distintos – e, portanto, é necessário introduzirmos uma nova dimensão na análise.

No escopo do presente trabalho, para dar conta do dialogismo implícito e explícito em todo enunciado, consideramos mais produtivo utilizar, em lugar da categoria bakhtiniana de polifonia, a expressão intertextualidade (FAIRCLOUGH, 1996). Estaremos particularmente

interessados em determinar o lugar ocupado pelas marcas visíveis de outros textos em sua incidência típica do ciberespaço – o *link* – na caracterização dos eventos comunicativos.

Nosso segundo foco de interesse no problema das marcas da intertextualidade tipicamente incidentes no ciberespaço recai sobre a multimodalidade, como discutido anteriormente. Os eventos multimodais articulam e integram múltiplos modos discursivos, e essa articulação pode ser considerada como marca da heterogeneidade textual. Interessa-nos aqui, ainda uma vez, sua incidência típica do ciberespaço, o que Berners-Lee chamou *hypermedia*, que integra imagem, som e vídeo ao hipertexto (BERNERS-LEE, 1989). Essa articulação multimodal integra o texto numa nova dimensão típica do ciberespaço – a hipermodalidade. Nesse contexto, a hipermodalidade é mais que uma dimensão constitutiva da heterogeneidade discursiva: ela é também marca visível dessa heterogeneidade, que aflora à superfície do texto e aponta – por meio do *link* – para outros textos. A hipermodalidade pode, portanto, ser concebida como a confluência da intertextualidade manifesta e da interdiscursividade numa única articulação – o texto, concebido agora e de uma vez por todas como hiperdocumento.

Posto o problema desse modo, podemos aventar um outro critério capaz de satisfazer às condições estabelecidas anteriormente e que pode ser produtivo: a hipermodalidade. Podemos concebê-la como uma dimensão contínua e assinalar pelo menos as seguintes regiões distintas: a região das enunciações monotonais, caracterizadas pela incidência nula de *links* e

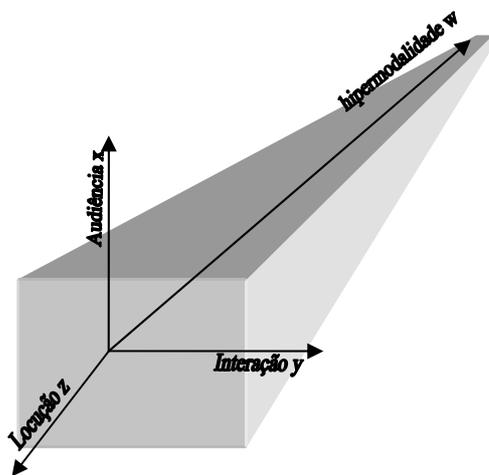


Ilustração 15: Uma representação quadridimensional do ciberespaço

por um único modo de apresentação, e a região das enunciações politonais, caracterizada por alta incidência de *links* e utilização de múltiplos modos de apresentação. Assim, uma representação possível de um ciberespaço quadridimensional pode ser vista na Ilustração . Com essa nova dimensão, os vários eventos suportados pelos recursos de vídeo-conferência e aqueles mais tipicamente associados à hipermídia (*sites*, portais, etc.) se deslocariam acentuadamente no eixo W em direção à politonalidade, distinguindo assim os eventos considerados. Se utilizarmos esses critérios para

localizar no ciberespaço os eventos comunicativos considerados, poderemos obter a caracterização da Tabela 6 como uma aproximação preliminar.

Eventos	Y ₁	y ₂	y ₃	z ₁	z ₂	z ₃	x ₁	x ₂	x ₃	w ₁	W ₂	v ₁	v ₂	v ₃
Leitura de diário			■		■			■			■			
Acesso a site			■		■			■			■			
Conversa bilateral	■			■			■			■				
Conferência bilateral	■			■			■				■			
Conferência multilateral	■					■	■				■			
Conversa multilateral	■					■		■		■				
Busca de informação		■		■				■			■			
Interação lúdica		■		■					■		■			
Aula virtual		■				■	■			■				
Discussão coletiva		■				■		■		■				
Entrevista virtual		■			■			■		■				
Multi-interação lúdica		■				■			■		■			
Acesso a portal			■			■		■			■			
Troca de e-mail	■			?	?		?	?		?	?			

Tabela 6: Distribuição dos eventos comunicativos no ciberespaço multidimensional

Legendas:

Y ₁ Interação direta	Z ₁ Locução 1-1	X ₁ Audiência conhecida
Y ₂ Interação mediada	Z ₂ Locução 1-n	X ₂ Audiência construída
Y ₃ Interação indireta	Z ₃ Locução n-n	X ₃ Audiência imaginada
W ₁ Hipermodo monotonal	V ₁ ?	
W ₂ Hipermodo politonal	V ₂ ?	
	V ₃ ?	

No entanto, o esforço não foi suficiente para distinguir dois eventos: o acesso (genérico) a um *site* e o acesso a um *blog*, este último inequivocamente considerado um gênero digital emergente em todas as pesquisas consultadas. Essa indistinção concede força ao argumento de Marcuschi, apesar de toda a discussão travada nas seções anteriores. Por outro lado, ela pode simplesmente indicar a insipiência do modelo proposto – o que parece uma solução mais simples.

Não cremos, porém, que o modelo deva ser inteiramente descartado. A razão disso é sua produtividade aparente. A dificuldade que tivemos em localizar inequivocamente os eventos de correio eletrônico nas dimensões da locução, da audiência e da hipermodalidade sugere a possibilidade de se tratar aqui de diferentes eventos comunicativos. Nossa experiência com o correio eletrônico encontra *e-mails* enviados para um ou mais interlocutores, endereçados a

destinatários conhecidos ou simplesmente supostos, em texto plano ou incorporando *links* e elementos de multimídia. Pensamos aqui nos diversos tipos de *spam* que circulam na rede. Isso pode indicar que o modelo é capaz de prever a emergência de eventos ainda não reconhecidos, que poderão vir a ter a estabilidade convencional de gêneros digitais.

Se pudermos demonstrar que essa sugestão não é acidental, teremos uma boa indicação de que o modelo é suficientemente produtivo para não ser descartado, ainda que o consideremos insatisfatório. Se isso puder ser feito, teremos assegurado um lugar para o acesso aos portais entre os eventos comunicativos do ciberespaço, já que pudemos facilmente plotá-los no modelo. Nosso esforço subsequente deverá se concentrar, pois, em demonstrar a validade do modelo, através da determinação desses eventos comunicativos possivelmente emergentes na correspondência eletrônica. Isso deverá ser feito como condição prévia ao refinamento do modelo. Como dissemos anteriormente, o modelo multidimensional do ciberespaço caracteriza apenas a *situação comunicativa* implícita nos eventos. Ele nada nos diz sobre os outros aspectos necessários para considerar tais eventos gêneros (digitais).

2.3.2. Um experimento identificador de novos gêneros digitais

Se o modelo proposto para a caracterização dos eventos comunicativos é relevante para a análise, devemos ser capazes de estabelecer uma correlação inequívoca entre algumas famílias de eventos comunicativos e a incidência de marcas das dimensões sugeridas, correlação que não possa ser estabelecida com outras famílias rotuladas inicialmente como “troca de *e-mail*”. Se isso for possível, teremos demonstrado igualmente a produtividade do modelo proposto, que nos tornou capazes de prever a incidência de novos eventos, mais do que simplesmente situar os já conhecidos. Precisamos, portanto, formular o problema de modo rigoroso.

Na caracterização esboçada até aqui não fazemos qualquer referência ao propósito comunicativo dos vários eventos, simplesmente dando como aceito que eles são identificáveis e diferentes. Isso foi feito porque se trata de critério qualitativo, que não atende à condição de ser uma grandeza homogênea que varie de algum modo contínuo e definido, estabelecido logo de saída. Neste momento, porém, o critério pode ser útil para tornar rigorosa a formulação do problema. Uma outra consideração deve incluída aqui. As dimensões que propusemos para o

modelo se referem especialmente à situação comunicativa em que o evento ocorre. Ora, o ferramental conceitual que recolhemos no capítulo anterior indica que a situação comunicativa *afeta* o modo como o texto é produzido. Isso significa que devemos ser capazes de rastrear no texto marcas visíveis dessas dimensões. Essas considerações nos permitirão formular o problema de modo rigoroso, utilizando variáveis facilmente observáveis num *corpus* textual qualquer. Diremos então que toda ocorrência de um evento com propósito comunicativo x deve corresponder à incidência de uma quantidade p definida de marcas textuais, quantidade esta que é necessariamente distinta da incidente num evento com propósito comunicativo y . Isto significa que se dois eventos tiverem o mesmo propósito comunicativo porém quantidades distintas de marcas, a proposição terá sido falseada e suas premissas refutadas.

Infelizmente, não podemos formular o problema dessa maneira porque o fenômeno abordado não se presta a medidas com esse grau de precisão: se as premissas conceituais que adotamos estão corretas, as convenções dos gêneros semióticos não são rigorosamente impositivas de tal modo que sempre ocorram em todos os casos concebíveis. Isto significa que a incidência de qualquer marca não é uniforme para qualquer ocorrência do evento.

Um modo de levar essa variabilidade em consideração e, ainda assim, manter o rigor da proposição é considerar a incidência média das marcas das dimensões escolhidas, isto é, adotarmos um enfoque *probabilístico*. Isso é inteiramente consistente com o enfoque que demos anteriormente ao problema da prototipicidade. Pela nossa abordagem, o melhor que podemos esperar de um processo de categorização é afirmar o grau de pertinência p de uma instância i qualquer à categoria G – o que é o mesmo que dizer que a instância i tem probabilidade p de ser categorizada como um gênero semiótico G .

Assim, podemos reformular a proposição, afirmando que a média p_m da incidência das marcas textuais é uma função definida do propósito comunicativo do evento, de modo que diferentes propósitos devem corresponder a médias diferentes em pelo menos uma ordem de grandeza. Nesse enfoque, assumimos que se formos capazes de observar um *corpus* suficientemente representativo do universo de mensagens possíveis, essa média então deverá corresponder à *probabilidade* de ocorrência de uma determinada marca em qualquer mensagem observável. Essa proposição pode ser facilmente falseada por inúmeros experimentos e, portanto, pode ser usada para corroborar nossas idéias.

Nosso inventário inicial de eventos comunicativos incluía três famílias distintas que podem ser suportadas pelos aplicativos de correio eletrônico, a saber: a troca de e-mail, a aula virtual e a discussão coletiva. Há consenso quanto a classificar como gêneros digitais as duas

últimas; no entanto, os critérios propostos anteriormente sugerem que esse consenso não deve ser tomado como tácito no que concerne à correspondência eletrônica em geral. Assim, o objetivo desse nosso primeiro experimento é verificar a incidência de diferentes famílias de eventos comunicativos na correspondência eletrônica que não possam ser classificadas como aula virtual ou discussão coletiva. Em particular, queremos verificar essa incidência naquela correspondência comumente classificada como *spam*. Aliás, é mais do que tempo de definir claramente o que se entende por *spam*. Uma definição bastante extensa e que dispensa maiores comentários pode ser obtida na Internet, no *site* Whatis:

O *spam* é uma mensagem não solicitada na Internet (...), onde o endereço do destinatário é obtido através de robôs ou fornecido por empresas especializadas em listas de distribuição. (...) O *spam* equivale aos telefonemas de *marketing* ativo, exceto pelo fato de o destinatário pagar em parte pela comunicação, já que todos os usuários da Internet compartilham os custos de mantê-la. Os *spams* normalmente são enviados a milhões de destinatários, na expectativa de que apenas um número insignificante responderá à oferta (WHATIS.COM, 2005).

Para o objetivo – mínimo – definido qualquer amostra de correspondência eletrônica que não contenha somente ocorrências de aulas virtuais e discussões coletivas deve ser suficiente, desde que o *corpus* seja suficientemente grande. Isto significa que a amostra não precisa ser representativa de todas as possíveis famílias de eventos comunicativos. Ela precisa apenas ter tamanho suficiente para conter uma quantidade qualquer de mensagens não solicitadas de tal modo que tenha boa probabilidade de conter pelo menos alguns exemplares de uma nova família de eventos comunicativos, caso a proposição seja válida. Essa simplificação do *corpus* é possível porque se trata apenas de identificar alguns novos eventos – e elaborar algumas hipóteses a propósito de suas convenções. Experimentos posteriores, que visem confirmar a emergência de tais gêneros, precisarão, naturalmente, de amostras mais representativas.

Desse modo, escolhemos como *corpus* de análise nossa própria correspondência eletrônica, recebida no período de 1 a 6 de junho de 2005, que totalizou 42 documentos. A seguir, arbitramos valores numéricos para os vários parâmetros de localização no ciberespaço das ocorrências do *corpus*, conforme exibido na Tabela 7. O procedimento seguinte consistiu em leitura preliminar dos documentos procurando determinar o propósito comunicativo de cada um e eventuais gêneros não-digitais que pudessem de alguma forma ser considerados como constituintes do texto dos documentos. A idéia por trás deste último procedimento parte da premissa de que os novos gêneros surgem reelaborando gêneros existentes. Tais procedimentos visaram tão somente elaborar algumas hipóteses preliminares a propósito do *corpus*, não fazendo parte da análise propriamente dita. Esse procedimento conduziu a uma

categorização preliminar em quatro grupos distintos, que constituiriam nossa primeira hipótese a propósito de diferentes eventos comunicativos.

	Locução	Audiência	Interatividade	Hipertextualidade	Hipermodalidade
0	1-1	Conhecida	Não existem recursos de resposta direta	Não existem <i>links</i> para outros domínios	Não existem semioses não lingüísticas
1	1-n	Qualificada (jornalista, técnico)	Existe pelo menos um recurso	Existe pelo menos um link	Existe pelo menos uma semiose
2	n-n	Construída (consumidor)	Existem vários recursos de resposta	Existem vários <i>links</i> para vários domínios	Existem várias semioses diferentes

Tabela 7: Os parâmetros de localização no ciberespaço das ocorrências do *corpus*.

O próximo procedimento consistiu em inventariar em cada documento a incidência dos parâmetros de localização definidos. Para cada um deles foi atribuído um valor numérico segundo os critérios definidos na Tabela 7. Por fim, foi calculada a média aritmética desses valores numéricos e os resultados foram listados na Tabela 8.

É de se notar que os grupos foram separados também pela origem provável das mensagens e não apenas pelo propósito suposto. A origem do primeiro grupo de mensagens são assessorias de imprensa de empresas de tecnologia, para quem o destinatário é conhecido como jornalista especializado e, portanto, provavelmente, interessado no conteúdo. A origem do segundo são empresas e instituições cuja atividade está relacionada ao perfil profissional do destinatário. O terceiro grupo tem origem incerta, podendo ser considerado um caso clássico de *spam*. O último grupo é composto por mensagens enviadas por pessoas conhecidas. Note-se, por fim, que as cinco últimas colunas receberam os nomes de conveniência da dimensão proposta e correspondem aos parâmetros definidos na Tabela 7, na mesma ordem em que foram ali listados. Como dissemos, as duas primeiras colunas foram preenchidas por hipótese, a partir de análise preliminar do *corpus*, sem consideração pelos parâmetros propostos.

Propósito comunicativo	Constituintes	Docs	Z	X	Y	W_a	W_b
Informar a ocorrência de um fato passível de tornar-se notícia	Press-release, notícia	18	1	1	0.61	1.55	1.44
Convidar para um evento de interesse profissional	Convite, carta profissional	7	1	1	0.57	0.85	1.71
Oferecer um produto ou serviço para aquisição	Mala direta, publicidade, home page	12	1	2	0.41	0.83	1.50
Trocar informações de	Carta pessoal	5	0.25	0	0.00	0.00	0.00

interesse pessoal							
-------------------	--	--	--	--	--	--	--

Tabela 8: Incidência média dos parâmetros considerados na amostra.

A primeira observação que podemos fazer sobre o resultado da medição é que não foi possível identificar no *corpus* distinção inequívoca entre os convites eletrônicos e as malas diretas comerciais. Podemos notar facilmente que a média de todos os parâmetros dessa categoria está dentro da mesma ordem de grandeza. Pela premissa assumida, não deve haver distinção entre os dois grupos; logo, trata-se da mesma família de eventos comunicativos. Provavelmente, trata-se de meras variantes de uma mesma família – o *e-offer* ou, mais convencionalmente, a mala direta eletrônica. O propósito comunicativo deve, portanto, ser o mesmo – o oferecimento de produtos ou serviços para aquisição. É provável que nos convites esse propósito esteja sendo deliberadamente ocultado, de modo a captar a atenção do destinatário. Trata-se aqui do sugerimos na seção 2.1.4. *As mensagens do meio: gêneros do discurso em Maingueneau* a propósito da fórmula de Charaudeau (“Estamos aqui para dizer ou fazer o quê?”). Neste caso, locutor e interlocutor respondem diferentemente à pergunta – e esta ambigüidade serve a uma função claramente retórica. Essa discrepância, contudo, não é suficiente para refutar a hipótese feita, já que somos capazes de distinguir pelo menos dois grupos de documentos claramente. Pelo menos nesse experimento, existe uma correlação definida entre os parâmetros considerados e o propósito comunicativo suposto.

Como se pode notar, a correspondência eletrônica pessoal tende a ser bilateral; quando muito, a mesma mensagem é enviada para um grupo pequeno e definido de destinatários, como seria de se esperar. O mesmo não ocorre com as duas variedades de *spams*, coerentemente com a definição dada.

A variação de uma ordem de grandeza nas medições feitas na dimensão da audiência X deve ser vista com cautela. A efetiva distinção entre o que chamamos uma audiência qualificada e uma audiência genérica é menor que o valor numérico arbitrado faz supor. Em ambos os casos, trata-se de variantes da audiência construída, distintas apenas na quantidade de atributos de qualificação que o remetente supõe no destinatário. Esses atributos – em se tratando das técnicas publicitárias – afetam a eficácia da mensagem, mas não necessariamente suas características. Nessa dimensão, portanto, trata-se de chamar a atenção para a clara distinção entre o grupo de mensagens de natureza pessoal e os demais.

A dimensão da interatividade Y é mais interessante. Na Tabela 6 não hesitamos em designar toda correspondência eletrônica como sendo de interação direta. Ao analisarmos o *corpus* selecionado, porém, notamos que nem toda correspondência tem essa característica. E

mais, na correspondência pessoal verificamos a inexistência de marcas textuais de interatividade. Neste caso, é a própria situação comunicativa a assegurar a interatividade: o endereço de correio foi fornecido espontaneamente, na expectativa de manter correspondência entre interlocutores conhecidos. É natural que a mensagem enviada implique em resposta: trata-se de situação inerentemente conversacional. No mínimo as convenções de etiqueta forçarão uma resposta. Isso pode ser facilmente verificado apenas inventariando a quantidade de *e-mails* pessoais que ficam sem resposta, que deverá ser necessariamente próxima de zero. Já na variedade de *spams* analisada no *corpus* a situação é bem outra: não existe expectativa de conversação por nenhuma das partes; mesmo o remetente espera taxa insignificante de respostas. Aqui a situação comunicativa não é inerentemente conversacional – e, no entanto, essa é a situação que o remetente deseja criar! Isso provoca o aparecimento de marcas visíveis de interatividade, de convites explícitos à resposta: a interatividade se move para a superfície do texto; ela se torna um artifício retórico.

A dimensão da intertextualidade W parece ter a mesma função retórica no caso dos eventos de *e-offer* e *e-release*. Tudo se passa como se as mensagens fossem produtos que, para serem vendáveis, precisassem de embalagem atrativa para captar a atenção do cliente. Isso é esperado nas mensagens de *e-offer*. Porém, sua incidência – elevada – nos *press-releases* eletrônicos convida à reflexão. A retórica da embalagem dirigindo a distribuição da notícia denuncia a falácia do imaginário jornalístico, que procura caracterizar a informação como bem cultural. Essa retórica revela a face de mercadoria assumida pela notícia na sociedade contemporânea – um produto que precisa ser embalado e vendido para circular.

Se o experimento realizado é convincente, o modelo que elaboramos para caracterizar eventos comunicativos no ciberespaço pode ser justificado simplesmente porque foi capaz de prever a emergência de novas classes de eventos ainda não reconhecidas pela comunidade que estuda o problema. É sua produtividade aliada à sua coerência com alguns dos conceitos mais consagrados das principais teorias dos gêneros discursivos que o torna sedutor. E se o modelo for aceitável, ele poderá ser utilizado como argumento em favor da caracterização dos portais de interesse geral como uma classe de eventos comunicativos relevantes para o ciberespaço. Afinal, não tivemos dificuldade alguma em plotar o acesso aos portais no gráfico quadridimensional proposto. No entanto, o modelo elaborado não afirma que o portal é um gênero semiótico; ele apenas caracteriza a situação comunicativa em que o evento “acesso a portal” ocorre e indica que ela é distinta dos demais eventos considerados. Em outras palavras, o passo (iii) que assinalamos logo no início do capítulo precisa ser dado.

2.3.3. O gênero portal de interesse geral

Na seção 2.2. *Canais (marcianos) e veículos de comunicação* procuramos desqualificar a noção de suporte, argumentando que ela não agrega valor ao conhecimento da dimensão contínua entre o canal de comunicação e as mensagens, em se tratando da comunicação humana e dos seus artefatos culturais. E ela não agrega valor porque falha em estabelecer fronteiras discretas nesse *continuum*. Ao longo dessa discussão, assinalamos que inúmeros artefatos culturais participavam tanto da dimensão do transporte quanto da própria dimensão da mensagem; o exemplo dado foi o *livro*, artefato que não é apenas parte da camada de transporte da comunicação, participando igualmente da dimensão das mensagens ao revelar valor de *signo*. Ora, se admitimos a existência de uma dimensão contínua ligando o canal da comunicação humana às suas múltiplas formas de mensagens, admitimos implicitamente que alguns artefatos estão mais próximos da dimensão do transporte que outros. Cabe, então, indagar em que posição desse *continuum* podemos situar os portais de interesse geral.

Em primeiro lugar, é preciso assinalar que se trata de uma categoria de *websites* bem delimitada e não uma ficção conceitual. Basta lembrarmos a pesquisa citada já no primeiro capítulo (RELATÓRIO, 2005): os usuários da Internet reconhecem um subconjunto dos *websites* como algo diferente dos demais, a ponto de terem consciência de que parte do seu tempo na Internet é gasto navegando neles. Além disso, inúmeros mantenedores de *websites* categorizam suas iniciativas como Portais: veja-se, por exemplo, o diretório Cadê (http://br.cade.dir.yahoo.com/Compras_e_Servicos/Internet_e_Web/Portais/), que lista inúmeros desses *sites*, que foram assim classificados por seus próprios mantenedores. A questão, portanto, consiste em determinar se estamos diante de um artefato (multimodal) de linguagem – um dos vários gêneros digitais – ou outra coisa qualquer, como, por exemplo, “um ambiente específico para localizar uma série de informações”.

Qual o propósito comunicativo dos portais de interesse geral? Se estamos diante de um gênero digital, deveríamos ser capazes de aplicar a esse evento comunicativo a fórmula de Charaudeau. Essa é uma pergunta para a qual, provavelmente, encontraremos inúmeras respostas, algumas possivelmente contraditórias entre si. Isso, no entanto, não deve ser um problema novo: se admitirmos, como Bonini que os jornais (e as revistas e os canais de televisão, entre outros “veículos” de comunicação) são um tipo de gênero do discurso (um

hiper-gênero, que inclui outros gêneros na sua constituição), o problema se aplicaria igualmente a eles. Assim, devemos admitir que o propósito comunicativo dos seus vários constituintes deve fazer parte da resposta. Uma possibilidade, portanto, é de que os portais se caracterizem, sob esse ponto de vista, exatamente pela *diversidade de propósitos*. Para essa diversidade poderia ser entendida como *típica* do gênero, ela deve se constituir numa *unidade* de propósitos, de tal modo que, ao nos depararmos com um *website* onde possamos reconhecer esse conjunto de propósitos, devemos reconhecer igualmente um portal de interesse geral. Essa tipificação requer uma longa demonstração, pelo que dedicaremos a ela toda a próxima seção, onde procuraremos determinar o conteúdo temático dos portais de interesse geral. Por ora, basta assinalar que se trata de meio indireto de determinarmos essa unidade de propósitos.

O terceiro aspecto que queremos abordar refere-se aos papéis desempenhados no evento, o que Maingueneau chamou o “estatuto dos parceiros legítimos”. Como vimos no capítulo anterior, plotamos os portais na região n-n (Z_3) da dimensão Z da Locução, obviamente influenciados pela diversidade de gêneros digitais encaixados nele. No entanto, em todos os casos observados, os portais incluem apenas os *chats* entre todos os eventos plotados em Z_3 ; não encontramos nenhum outro evento multilateral nos exemplares do *corpus*. Por esse aspecto, os portais tendem a se confundir com quaisquer *websites*, apenas com a característica de serem multitemáticos. É curioso mesmo que os portais, que tão claramente se empenham em agradar a todos os tipos de usuários de modo a atraí-los para sua órbita, não invistam mais nessa direção. Lembremos que a mesma pesquisa citada dá conta de interesse elevado na navegação por comunidades virtuais, como o Orkut, interesse mesmo mais elevado que nos próprios portais, alguns dos quais se proclamam campeões de audiência. O levantamento do Comitê Gestor da Internet no Brasil (CGI.BR, 2006) indica que 21,97% dos usuários declaram explicitamente a utilização de sites e comunidades de relacionamento (o Orkut é nominalmente citado). É muito provável que esses números só façam crescer. Isso pode ser inferido indiretamente, pela avaliação do crescimento de uso dos *chats* (mais de 28% das atividades listadas pelo CGI.BR). Em junho de 2004 inventariamos 2.694 salas de *chat* abertas no site UOL; em outubro de 2006, esse número cresceu para 3.018 (ver <http://tc.batepapo.uol.com.br>). No mesmo período, o iG disparou de 587 para 1.833 salas (ver <http://wchat.ig.com.br>), refletindo diretamente o crescimento do *site*.

É muito provável que este “desinteresse” num dos mais relevantes aspectos da navegação dos usuários brasileiros esteja relacionado ao motivo porque a MSN encerrou sua manutenção de salas de *chat*: o descontrole. É notório o clamor “público” por medidas de

controle das comunidades do Orkut, como se a resolução da criminalidade passasse pela redução dos direitos de associação e expressão de todos os cidadãos. Um *site* como o Orkut é incontrolável: qualquer usuário pode criar sua própria comunidade; nos portais, apenas os mantenedores podem abrir novas salas de *chat*. Desse modo, a locução multilateral tende a ser controlada nos portais, pelo menos ao nível temático. Esse controle do discurso se estende também para outras locuções incorporadas aos portais, como os *blogs*: todos os portais exigem identificação para a criação dessa nova versão do diário adolescente. O controle aqui visa menos “prevenir” o discurso “criminoso” e mais assegurar a posse da “identidade virtual” do usuário, de modo a melhor direcionar a publicidade e as ofertas de bens e serviços.

Uma hipótese possível para dar conta desses controles é que eles se reflitam no discurso como “auto-censura” da enunciação, como uma uniformidade de discursos que possivelmente se originem numa mesma região ideológica, pelo menos naqueles “canais” de responsabilidade do mantenedor. Essa hipótese pode se articular diretamente ao outro aspecto dos participantes desse evento comunicativo – o que chamamos no capítulo anterior a dimensão X da Audiência. Ali, plotamos os portais na região das audiências construídas. Mas que “construção” é essa?

Como assinalamos muito cedo, os portais procuram abarcar o máximo de interesses variados da audiência de modo a assegurar o máximo de audiência e o maior tempo possível de permanência nos seus domínios, condição necessária para a rentabilidade do empreendimento, cuja receita é baseada principalmente em publicidade. Supomos que isso deve ter um reflexo na construção da audiência, possivelmente reduzida à sua expressão mais simples – a de Consumidor. Isso se reflete diretamente na distribuição temática dos portais: abstrações de público como o Jovem, a Mulher, o Homem, o Gay; itens de consumo de massas como Carro, Diversão, Cinema, Jogos, Beleza, Música; temas genéricos e de largo alcance como Sexo, Empregos, etc.. A distribuição temática dos portais tende a encarar a audiência como abstrações genéricas ou simplesmente como faixas (abstratas) de público orientadas ao consumo de massas; uma hipótese interessante consiste em supor que os portais tendem a categorizar os atributos da sua audiência do mesmo modo que a publicidade: mulher com cerca de trinta anos, de classe média e faixa A de renda e consumo, casada (ou pelo menos candidata ao casamento), etc. A variedade dos atributos esconde a uniformidade do perfil, de modo a melhor captar a variedade de interesses da audiência, condição para a ampliação do leque de anunciantes (este, sim, a verdadeira “audiência” dos portais). Tal como a posição (controlada) da locução, essa categorização da interlocução deve afetar diretamente a construção do discurso permitido nos portais, projetando uma imagem de sujeito

possivelmente típica daquela projetada pelo discurso publicitário. Algumas idéias e exemplos a esse respeito serão abordados na última parte desta dissertação.

Um outro aspecto que consideramos ao incluir os portais de interesse geral na categoria de gênero digital é o que Maingueneau chamou uma “organização textual” (multimodal) própria, facilmente rastreável nas suas instâncias. Podemos abordar as várias características dessa organização através de exemplos, colhidos de seus exemplares, supondo que tais exemplos incidam igualmente numa quantidade suficientemente grande de instâncias. Podemos, ao contrário, eleger como *prototípica* uma instância qualquer, considerada pela comunidade de usuários como *exemplar* do gênero e abordá-la. Essa última abordagem estaria baseada na concepção corrente de “protótipo” entre os estudiosos do gênero, mas teria um sério inconveniente: ela requer uma pesquisa de campo entre os usuários do gênero, utilizando amostra suficientemente abrangente do conjunto universo para ser considerada ergódica, algo fora de cogitação para um pesquisador individual sem financiamento adequado.

Uma outra abordagem mais simples pode, porém, ser adotada, desde que utilizemos uma outra concepção do que seja “protótipo”. Essa concepção, já abordada e colocada em prática no próximo capítulo, incorpora imprecisão suficiente para que possamos *arbitrar* um exemplar imaginário e confrontá-lo contra as instâncias concretas do gênero. Essa concepção leva em consideração que cada usuário poderá responder de modo diferente ao ser indagado a propósito do que seja um exemplar prototípico de uma categoria qualquer; desse modo, devemos pensar no protótipo em termos de *respostas médias* do universo de observadores possíveis, dentre os quais o próprio investigador. Assim, podemos simplesmente imaginar o que seja um exemplar prototípico com base na nossa própria experiência, confrontando-o contra as instâncias do gênero, medindo seu grau de proximidade (ou afastamento) com o modelo. Essa é a abordagem que adotaremos, dada sua simplicidade. Com base nisso, elaboramos um “modelo” do que imaginamos ser a distribuição visual e a composição temática de um portal de interesse geral, visto na Ilustração . Note-se que este modelo foi elaborado *antes* de iniciarmos uma abordagem sistemática do nosso *corpus* de pesquisa, como uma hipótese de trabalho.

Algumas observações concernentes à organização multimodal desse protótipo arbitrário são relevantes. De saída, assinalamos os recursos mobilizados para constituir uma *identidade visual* dos portais. Trata-se de empreendimentos comerciais e, como tais, sujeitos a acirrada competição e que, portanto, precisam oferecer elementos diferenciais para se manterem viáveis. Esse “diferencial competitivo” não se resolve somente no conteúdo temático, incluindo também uma caracterização própria, que os permita ser “lembrados” e identificados

pelos usuários. O “nome” do empreendimento e seu logotipo, as características mais óbvias das marcas comerciais são imediatamente identificados; porém, entre os principais empreendimentos notamos a utilização de uma cor ou combinação predominante de cores vizinhas no espectro como marca identitária. Isso ficou marcado no nosso protótipo pela utilização predominante de tons de verde (sempre contra fundo branco) e é corriqueira a utilização do recurso.

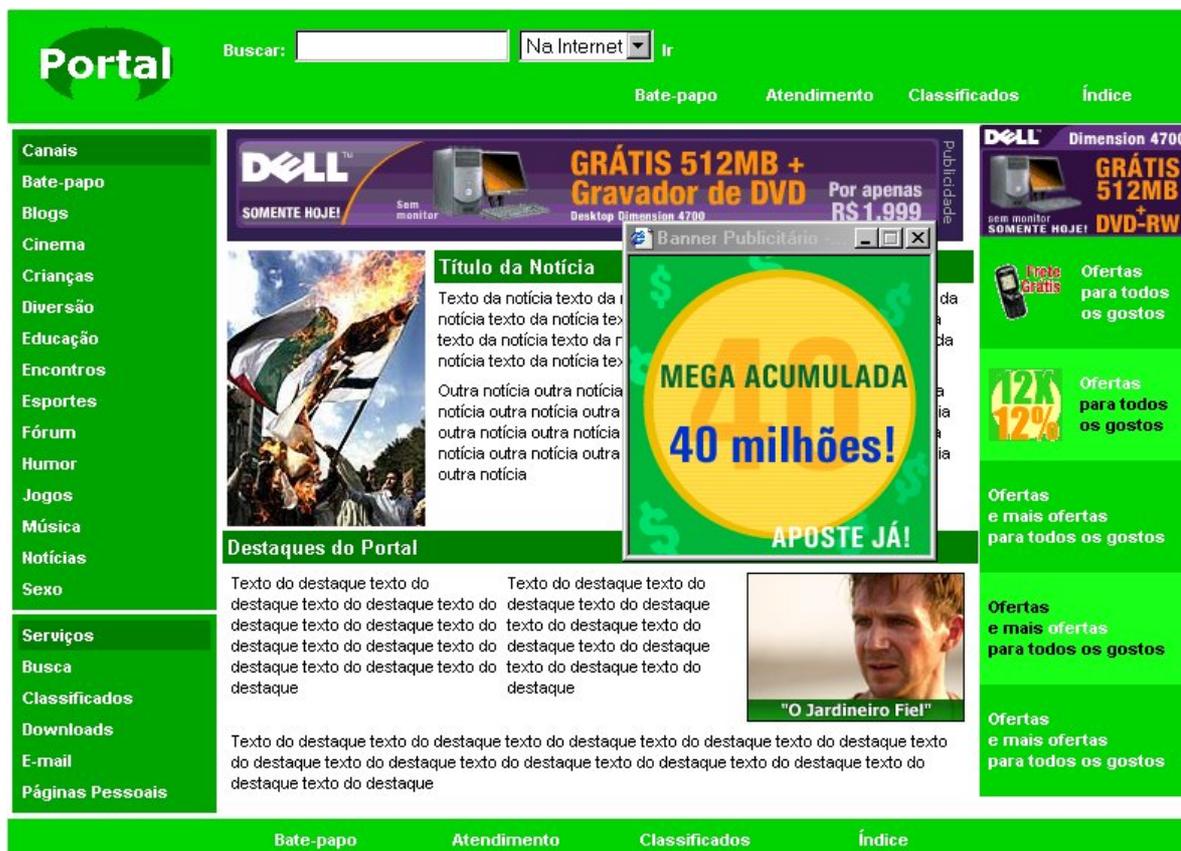


Ilustração 16: Protótipo (intuitivo) de um portal de interesse geral.

Embora a distribuição geográfica (a diagramação) dos elementos composicionais possa ser utilizada como marca identitária, seu uso como diferencial é restringido pela própria finalidade dos portais – um elemento comum a todos eles. Os hábitos ocidentais de leitura de texto e imagem impõem certa distribuição preferencial dos elementos na tela do computador. Assim é que os elementos principais do conteúdo – a oferta de “canais” e serviços privilegiados como diferenciais competitivos – aparecem no topo da tela e à esquerda, certamente na expectativa de que sejam lidos primeiro. Igualmente, o centro da tela é reservado para os elementos cuja leitura possa ser considerada crítica, tanto do ponto de vista da finalidade do *site* (a publicidade) quanto do ponto de vista da “atração” do usuário. Em todas os exemplares observados essa região é reservada para pelo menos um *banner*

publicitário e para o conteúdo mais “atrativo”, geralmente de valor noticioso, normalmente fatos tipicamente publicados em jornais ou revistas, dependendo da “seriedade” que se queira transparecer.

Dois são os elementos que presidem a coesão espacial da primeira página: em primeiro lugar, como já assinalamos, o hábito ocidental de leitura – da esquerda para a direita e de cima para baixo. Mas também um elemento típico do *design* visual: o centro da linha de visada. Estes elementos costumam ser dedicados aos tópicos centrais do conteúdo, geralmente de teor noticioso e jornalístico. Em torno desse tema central se organizam os demais elementos, com destaque para o material publicitário e os elementos temáticos típicos dos portais. Em torno desses elementos de destaque os vários constituintes são dispostos em *colônias*, este interessante conceito devido a Hoey, O pesquisador entende que “uma colônia é um discurso cujos constituintes não têm seu sentido derivado da seqüência onde se encontram” (HOEY, 2001: 71). Discurso aqui é empregado de modo bem distinto do que vimos fazendo, como equivalente de texto ou conjunto de enunciados. A idéia da colônia é simplesmente a de um texto com princípio de coesão não seqüencial, sendo facilmente aplicável ao que vimos assinalando a propósito do *design* visual da primeira página dos portais.

Se internamente as colônias de elementos textuais não têm relação direta entre si, estando relacionados apenas a um elemento de nível superior – seu tema ou assunto, a relação entre as colônias é nitidamente hierárquica. Os espaços que se supõe serem primeiro acessíveis ao olhar do usuário (o centro, a esquerda e o topo) são dedicados aqueles elementos mais atrativos ou à publicidade.

Um outro aspecto que precisa ser considerado, tanto na montagem de cada colônia quanto na sua distribuição na geografia da página, é o uso de modos como cor, imagens, fontes, linhas e outros artificios para assegurar tanto a diferenciação quanto o destaque de cada colônia. O resultado tem a sobrecarga visual e informativa típica de vitrines de produtos para a venda. Tais elementos visam captar a atenção do usuário para cada colônia individual e funcionam, portanto, como um atrator retórico (visual).

Esse conjunto de elementos do *design* temático e visual dos portais não são encontrados tão fortemente conjugados em outras categorias de *websites* – e se constituem, portanto, em importante mecanismo identitário do gênero. Além disso, sua ampla disseminação em todo o *corpus* é um indicador de que tais *websites* devem ser vistos como uma categoria particular de artefatos de linguagem multimodal.

3. ALGUMAS CURVATURAS NO CIBERESPAÇO: OS ATRADORES RETÓRICOS

No último capítulo procuramos posicionar os portais de interesse geral no contexto dos gêneros (semióticos) digitais. Como assinalamos, um dos objetivos dessa operação era justificar a abordagem dessa categoria de *websites* do ângulo dos discursos (multimodais). Esperamos, porém, que o saldo credor dessa discussão tenha sido maior; notadamente, esperamos ter ampliado por pouco que seja nossa compreensão dos gêneros digitais emergentes no ciberespaço. Em especial, gostaríamos de ter adicionado pelo menos um pequeno tijolo ao edifício conceitual com que os pesquisadores têm abordado esse (ainda) novo e importante campo da atividade humana – o ciberespaço. Porém, estamos cientes de que avançamos notavelmente pouco na compreensão do objeto que nos move: tudo o que agregamos a essa compreensão foram algumas poucas observações de caráter geral sobre os objetivos e justificativas desse gênero. É o momento, portanto, de limitarmos nossas observações a um perímetro mais estreito, abordando algumas características dos portais.

Ao introduzir nossa pesquisa, assinalamos que os hábitos de navegação dos usuários conectados à Internet a partir da rede brasileira pareciam mais centrípetos que os dos demais usuários. Sugerimos também que tais hábitos centrípetos estariam em contradição com as tendências centrífugas tipicamente esperadas na *web*, caso assumíssemos o modelo interpretativo do ciberespaço proposto no segundo capítulo. Aquele modelo, como vimos, propunha um ciberespaço cuja topologia poderia ser considerada como de curvatura negativa, aberta, que implicaria em hábitos de navegação centrífugos. Sugerimos, então, que pelo menos parte da explicação para a tendência observada entre os usuários brasileiros poderia ser dada pelo peso dos portais de interesse geral nos seus hábitos de navegação. Isso foi marcado naquela introdução pela sugestão da metáfora da distorção da curvatura do ciberespaço nas vizinhanças dos portais de interesse geral.

Ora, se essa metáfora é mais que especulação divertida, devemos ser capazes de apontar elementos estruturais constitutivos dos portais que sustentem essa tendência. E como, evidentemente, estamos interessados nos portais não como empreendimento, mas como artefato de linguagem (multimodal), é entre os elementos dessa ordem que devemos empreender nossa busca. Assim, cabe a pergunta: o que seduz nos portais? O que torna a sua uma palavra sedutora – ou pelo menos aprisionante? Queremos, portanto, situar alguns dos

seus *atratores retóricos*, aqueles elementos do discurso multimodal que tentam atrair ou aprisionar os usuários nos domínios dos portais. As próximas seções abordarão alguns desses atratores retóricos, em especial aqueles que envolvem recursos típicos da Internet.

3.1. Satisfação garantida ou seu dinheiro de volta: o domínio temático dos portais

Ao discutirmos a noção de gêneros nas diversas tradições de pesquisa, concluímos que o *propósito comunicativo* (para utilizarmos o conceito de Swales, embora a noção esteja presente de uma forma ou de outra também nos demais autores) é decisivo na formulação do problema do ponto de vista dos participantes do evento. Posto o problema no nosso *corpus* em particular, devemos responder à fórmula de Charaudeau já largamente citada: vamos aos portais para fazer o quê? Em outras palavras, devemos determinar de saída o conteúdo temático e composicional dos portais de interesse geral.

A solução do problema deve ser, evidentemente, análoga ao problema geral do reconhecimento de um gênero semiótico. Para retomar sumariamente o problema da categorização, já discutido, assumimos que, embora um gênero seja reconhecido por um conjunto difuso e heterogêneo de atributos constitutivos, funcionais e interacionais, cada um desses atributos pode ser igualmente reconhecido como um conjunto difuso (embora não heterogêneo). Isto é, os atributos do gênero, tomados isoladamente, são também reconhecidos com referência ao seu grau de prototipicidade.

Desse modo, nossa abordagem do problema do conteúdo temático e composicional dos portais deverá ser essencialmente a mesma que adotariamos ao empreender a solução do problema do reconhecimento do gênero em si. Como a abordagem de Swales nesse aspecto foi a conceitualmente mais precisa, nós a tomaremos como ponto de partida.

Como os membros de uma comunidade (discursiva) reconhecem numa dada instância textual um exemplar de um gênero? A resposta de Swales, como vimos, é simples: eles o fazem do mesmo modo como reconhecemos os membros de uma categoria – isto é, eles emitem seu juízo com referência a exemplares prototípicos aprendidos através de exemplos, em oposição à teoria clássica da categorização. Como já sabemos, isto significa que uma categoria não tem fronteiras precisas nem seus membros podem ser inequivocamente

associados a elas no mesmo grau – eles se localizam em posições variáveis em relação ao que se convencionou chamar o *centro prototípico* da categoria. Uma consequência interessante dessa noção é que não podemos definir com precisão um gênero semiótico, do mesmo modo que um membro de uma comunidade discursiva não pode categorizar com precisão um determinado exemplar dele.

Essa imprecisão no conhecimento pode ser associada ao que alguns lógicos chamam *incerteza epistêmica*, relacionada à incerteza referente às crenças e opiniões expressas nas nossas proposições a propósito de fragmentos da realidade. Essa incerteza seria completamente distinta da incerteza associada aos eventos aleatórios, como os jogos de azar ou o movimento das moléculas de gás. Para caracterizar essa distinção, esses lógicos utilizam a noção de *imprecisão* (ou *vagueza*), aplicável ao que chamamos há pouco “incerteza epistêmica”, por oposição à noção de *incerteza* (propriamente dita). Esta última decorreria da nossa incapacidade de prever o comportamento dos eventos num mundo observado qualquer, enquanto que a primeira seria “uma propriedade dos sinais utilizados nos sistemas de representação” (OSTASIEWICZ, 2003: 14) construídos a propósito das observações que podemos realizar nesse mundo possível. Por exemplo: “antes de jogar um dado *eu não sei* que número ocorrerá. Este tipo de incerteza é chamado incerteza aleatória. Por outro lado, antes ou mesmo depois de jogar o dado eu não sei sempre se o lance resultante é *bom*. Trata-se de incerteza epistêmica” (OSTASIEWICZ, 2003: 14).

Nosso argumento é que podemos utilizar essas idéias a propósito da imprecisão e da vagueza do conhecimento, formalizadas na lógica nebulosa, para dar conta da categorização dos gêneros semióticos em geral. No entanto, no presente contexto, em vista do alvo particular que temos em mente, propomo-nos a utilizar a *fuzzy logic* para sugerir respostas ao problema do domínio temático dos portais de interesse geral. Ao invés de inventariarmos todos os assuntos abordados em todos os portais (ou numa amostra ergódica do universo de exemplares) e, por comparação, determinarmos quais os temas necessários e suficientes que precisaríamos encontrar para reconhecer um exemplar típico – a abordagem clássica do problema da categorização, propomos determinar um conjunto nebuloso de assuntos que possam fazer parte em algum grau definido do domínio temático dos portais.

3.1.1. Uma caracterização vaga e imprecisa dos gêneros digitais

Suponhamos que se queira projetar um dispositivo capaz de distinguir instâncias da categoria *xícara* das instâncias de *copo* e *caneca*, e que este dispositivo aprenda a discriminar as categorias através de exemplos. Para simplificar nosso problema, suponhamos também que tudo o que precisamos sejam algumas poucas variáveis e que os exemplos submetidos ao dispositivo tenham a forma lógica da proposição:

$$(x \text{ é pequeno}) \wedge (x \text{ tem alça}) \wedge (x \text{ é cônico}) \cong (x \text{ é xícara})$$

Cada uma dessas propriedades é inerentemente vaga. Por exemplo, como identificar uma alça? A partir de que ponto nas inúmeras possibilidades de uma certa protuberância no recipiente dizemos que ele “tem alça”? Com base nisso, convencionamos por simplicidade alguns valores discretos para dar conta do valor lingüístico de cada predicado, incluindo aí o valor de verdade de cada proposição.

Após uma quantidade suficientemente grande de exemplos, nosso dispositivo possivelmente acabaria por concluir que uma xícara X qualquer poderia ser identificada pela correlação lógica das seguintes proposições restritivas de Zardeh⁶:

$$H_x = \text{normalmente}(X \text{ é pequeno})$$

$$A_x = \text{freqüentemente}(X \text{ tem alça})$$

$$F_x = \text{eventualmente}(X \text{ é cônico})$$

$$U_x = \text{normalmente}(X \text{ é usado para beber café})$$

Note-se que as proposições representam somente restrições probabilísticas, que indicam, por exemplo, com que freqüência uma proposição como (x é pequeno) será considerada verdadeira pelos observadores. Nesse modelo, uma “condição necessária” de uma categoria é apenas um caso especial e deve obedecer a um valor de restrição como em

⁶ Zardeh assinala que “o significado de uma proposição p obtida de uma língua natural pode ser representado pelo que é chamado uma restrição geral de uma variável” (ZARDEH, 2003: 4). Mais especificamente, uma restrição desse tipo é representada pela proposição geral $X \text{ is } r \text{ } R$, onde X é a variável restrita, R é a relação que a restringe e $\text{is } r$ é uma cópula onde r é uma variável cujo valor define o modo pelo qual R restringe X .

sempre(*X é pequeno*), já que uma propriedade que é também uma condição necessária deve *sempre* ocorrer. A questão é que o modelo prescinde de uma propriedade desse tipo! Do mesmo modo, nenhuma propriedade isolada é *suficiente* para a categorização – ela é dada pela correlação lógica de todas as proposições.

Desse modo, argumentamos que é possível caracterizar os gêneros digitais utilizando como parâmetro um número definido de generalizações restritivas da forma proposta pela lógica nebulosa. A correlação de tais proposições poderá ser interpretada como a representação lógica do *protótipo* do gênero. Por outro lado, com base nessa representação, podemos igualmente determinar o grau de pertinência de um exemplar à categoria (isto é, emitir um juízo sobre uma instância textual qualquer) simplesmente determinando a distância lógica entre as proposições a propósito do exemplar e o protótipo representado por aquelas generalizações.

O problema é que o conhecimento das propriedades que podem eventualmente caracterizar um gênero está difuso na comunidade (discursiva) de seus usuários. E mais, esse conhecimento não é o mesmo para todos os membros da comunidade: pelas premissas que assumimos é esperado que os observadores não concordem em seus juízos sobre as proposições a propósito das propriedades dos gêneros. Trata-se daquela incerteza epistêmica de que falamos anteriormente. Desse modo, qualquer que seja a abordagem adotada, seja a observação direta de uma quantidade muito grande de exemplares, seja o inquérito junto a uma quantidade muito grande de usuários do gênero, seja simplesmente arbitrando um conjunto de propriedades com base no nosso próprio conhecimento intuitivo e submetendo-o a experimentos empíricos, devemos nos contentar com a incerteza inerente dos resultados. O melhor resultado possível não será mais que *um* dos muitos resultados *possíveis*.

A próxima questão que devemos abordar refere-se ao *cálculo* da correlação lógica entre as proposições restritivas elaboradas a propósito de um gênero digital qualquer. Como dissemos, é essa correlação entre as propriedades do gênero que o caracteriza como categoria específica. Uma interessante abordagem é sugerida por Siler e Buckley (2005). Nesse modelo, uma quantidade qualquer de observadores é solicitada a emitir um juízo sobre se uma determinada proposição é verdadeira ou falsa, admitindo-se para as respostas somente um desses dois valores, o que não é realista, mas pode servir como ponto de partida. O valor de verdade da proposição é, então, considerado como sendo a média de todas as respostas. À medida que o número de observadores cresce, a granularidade da estimativa diminui até se tornar na probabilidade de que um observador qualquer afirme que a proposição seja verdadeira.

Trata-se agora de abordar a questão de definir o tipo de correlação lógica que podemos estabelecer a propósito das proposições concernentes às propriedades de um gênero digital qualquer. Em primeiro lugar, queremos medir o grau de pertinência de uma dada instância ao gênero ao qual supostamente pertence. Admitamos o caso mais simples, onde todas as propriedades contribuem igualmente para a caracterização do gênero. Neste caso, as p_n proposições que podemos enunciar sobre as n propriedades do gênero são logicamente equivalentes e, portanto, o grau de pertinência da instância deve ser afetado tão somente pelo valor de verdade de cada uma das proposições. Neste cenário, parece-nos plausível o conhecimento intuitivo de que quanto mais propriedades forem satisfeitas pela instância dada (isto é, quanto mais o valor de verdade de suas proposições se aproxima da unidade), mais provável se torna a afirmação de que ela pertence ao gênero. É, portanto, razoável supor que o grau de pertinência da instância ao conjunto (nebuloso) do gênero possa ser dado pela média aritmética do valor de verdade das proposições restritivas de Zadeh elaboradas para dar conta das suas propriedades constitutivas.

Sintetizemos a abordagem com um único exemplo. Suponhamos que estamos avaliando as n instâncias de um gênero digital G , que pode ser caracterizado por p_i proposições restritivas a propósito de suas propriedades constitutivas, conhecidas *a priori*. Admitamos agora que temos acesso a todos os membros de uma comunidade discursiva usuária do gênero. Submetemos, então, a instância i_1 ao primeiro observador e solicitamos que avalie a proposição p_1 concernente a uma propriedade do gênero e repetimos o processo até a i -ésima proposição. A seguir, submetemos esta mesma instância a todos os demais observadores e calculamos a média das respostas dadas a cada proposição. Por fim, repetimos o processo até a n -ésima instância. Um resultado hipotético para essa pesquisa pode ser aquele exibido na Tabela 9.

	i_1	i_2	i_3	i_4	I_5	i_n	
p_1	0,93	0,78	0,84	0,45	0,84	...	0,87
p_2	0,89	0,47	0,80	0,39	0,77	...	0,65
p_3	0,96	0,29	0,63	0,77	0,90	...	0,70
p_4	0,78	0,91	0,54	0,32	0,44	...	0,59
p_5	0,76	0,68	0,18	0,45	0,55	...	0,33
p_i
	0,96	0,62	0,58	0,38	0,71	...	0,11

Tabela 9: Resultado de exemplo da análise das propriedades de G .

Neste exemplo, o total de cada coluna representa a média das respostas de todos os observadores consultados a todas as propriedades e pode ser interpretado como o grau de pertinência da instância ao gênero (ou simplesmente seu grau de prototipicidade). Esse valor pode ser interpretado também como a probabilidade de que um observador responda afirmativamente à pergunta: o texto x é um exemplar do gênero G ? O total de cada linha representa a média das respostas em todas as instâncias para cada proposição e pode ser interpretado como o grau de relevância da propriedade para o gênero. Esse valor pode ser interpretado como a probabilidade de que um observador qualquer responda afirmativamente à pergunta: um gênero G satisfaz à propriedade p_i ? Por fim, o total da última coluna é o produto da conjunção lógica de todas as proposições (obtido por um dos métodos sugeridos por Siler e Buckley). Esse resultado indica que 11% das instâncias pesquisadas podem ser consideradas prototípicas do gênero G ou, por outro lado, que existe probabilidade de 0,11 de que a n -ésima instância investigada seja um protótipo do gênero.

É importante assinalar que a abordagem que estamos propondo para o problema não se constitui num modelo de *caracterização* dos gêneros digitais, já que, como assinalamos, ela pressupõe que suas propriedades constitutivas sejam conhecidas de antemão. Tudo o que o modelo nos permite é *reconhecer* um exemplar de um gênero já conhecido e avaliar seu grau de prototipicidade. No entanto, o modelo pode ser utilizado para *testar* hipóteses a propósito das propriedades de um gênero, elaboradas empiricamente por um observador semi-divino com propósito estritamente acadêmico. Isto pode ser feito porque, como vimos na seção anterior, o modelo permite calcular a *relevância* de uma propriedade.

A idéia é, portanto, elaborar algumas hipóteses preliminares sobre quais as características temáticas esperadas de um portal e submeter o modelo (ou alguma variante dele) a um conjunto de *sites* candidatos a exemplares da categoria. Do resultado obtido, podemos supor que as propriedades que se revelarem mais relevantes para a caracterização (isto é, aqueles temas cujo valor de verdade estiver mais próximo da unidade) provavelmente farão parte do conjunto de propriedades utilizadas pela comunidade discursiva para categorizar o gênero. A utilização repetida dessa abordagem pode então permitir inferir por aproximações sucessivas as características temáticas de um portal. E como se trata de uma abordagem genérica, se bem sucedida, o modelo proposto pode ser utilizado para a caracterização de qualquer aspecto de um gênero digital.

3.1.2. A seleção do *corpus* para a análise

Um modo bastante simples de determinarmos as características desse tipo de *site* seria a observação direta dos portais de maior audiência, já que eles devem ser determinantes para o modo como a comunidade de usuários concebe a categoria. Infelizmente, as pesquisas de audiência realizadas pela Nielsen/NetRatings, associada do IBOPE no Brasil, não estão disponíveis publicamente. Portanto, a determinação do *corpus* de pesquisa deverá ser indireta. Uma abordagem mais exequível seria assumir que qualquer *site* que se apresente como um portal seja aceito como parte do *corpus* de pesquisa. Isso torna fácil o levantamento dos exemplares para análise. O diretório Cadê, propriedade da tradicional ferramenta de busca Yahoo, apresenta um elenco de *sites* nesta categoria (ver http://br.cade.dir.yahoo.com/Compras_e_Servicos/Internet_e_Web/Portais). Como a inscrição nesse tipo de diretório é voluntária (e remunerada), podemos estar certos que os mantenedores dos *sites* querem que eles sejam reconhecidos como portais.

A Tabela 10 lista os 30 *sites* no diretório Cadê no período de 8 e 9 de dezembro de 2005, quando o levantamento foi efetuado. Note-se que apenas os *links* ativos na data da pesquisa estão arrolados na lista do que consideramos o nosso *corpus* básico. Por outro lado, em vista da efemeridade da Internet, é possível que alguns desses *links* possam não estar mais ativos. No entanto, como nossa argumentação refere-se à *população* dos portais, isso não deve afetar significativamente os resultados, desde que não ocorreram mudanças significativas nas expectativas do público e nas concepções dos empreendedores (isto é, sendo mantida a concepção básica do gênero), sendo de se esperar que resultados equivalentes sejam encontrados para *qualquer* amostra.

Site	URL	Site	URL
AOL	http://www.americaonline.com.br	Bem Rápido	http://www.vemnavegar.com.br/
Brasil Online	http://www.bol.uol.com.br/	Cadê	http://br.cade.yahoo.com/
Cidade Internet	http://www.cidadeinternet.com.br	Crazy Mania	http://www.superzap.com/
Eg Interactive Tools	http://www.entreganhe.com.br	Esquinas	http://esquinas.terra.com.br/
Globo.com	http://www.globo.com/	Ibest	http://www.ibest.com.br
iG	http://www.ig.com.br	Inter .net	http://www.br.inter.net
iTudoMais	http://www.itudomais.com.br	LatinLook.com	http://www.latinlook.com/86/business/index.cfm

Matrix 2005	http://www.matrix.com.br	MSN Brasil	http://www.msn.com.br
Mundo Perdido	http://www.mundoperdido.com.br/	O Site	http://www.osite.com.br
POP	http://www.pop.com.br	Portal Free	http://www.portalfree.com.br
Portal Nacional	http://www.portalnacional.com.br/	Post	http://forum.pootz.org/
Rio Brasil	http://www.riobrasil.com.br/	Site do JP	http://www.sitedojp.com/
Sociedade Digital	http://www.sociedadedigital.com.br	Starmedia	http://www.br.starmedia.com/
Terra	http://www.terra.com.br	Terravista	http://www.terravista.pt/
UOL	http://www.uol.com.br	Yahoo Brasil	http://br.yahoo.com/

Tabela 10: Portais ativos listados no diretório Cadê.

Visando ampliar o conjunto universo pesquisado, submetemos a expressão completa “portal Internet” aos dispositivos de busca Google (<http://www.google.com.br>) e Yahoo (<http://br.yahoo.com>), na suposição de que *sites* contendo estas palavras em algum ponto das suas primeiras páginas estariam igualmente se assumindo como portais (ainda que não necessariamente de interesse geral). Do resultado da busca foram ignoradas, naturalmente, as referências aos *sites* já elencados no *corpus* principal, resultando em 27 *sites* adicionais, cujos endereços estão listados na Tabela 11. É de se supor que os resultados obtidos neste grupo sejam diferentes dos obtidos no grupo de referência, sendo esperado que eles tenham grau de pertinência ao gênero menor.

Site	URL	Site	URL
Âmbito Jurídico	http://www.ambito-juridico.com.br/	Arteplástica	http://www.arteplastica.com/portal.htm
CAPES	http://www.capes.gov.br/capes/portal/	Cidadania na Internet	http://www.cidadania.org.br/
Correios	http://www.correios.com.br/	Educacional	http://www.educacional.com.br
Elnet	http://www.elnet.com.br/	Embratel	http://www.embratel.com.br/
Firemasters	http://www.portalfiremasters.com.br/	Google	http://www.google.com.br/
Indústria Virtual	http://www.industriavirtual.com.br	Jus Navigandi	http://jus.uol.com.br/
Oi	http://www.oi.com.br/data/Page/s/30E5C7B0PTBRIE.htm	Piq	http://www.piq.com.br/
Polycol	http://www.artesanatonarede.com.br/	Portal AZ	http://www.portalaz.com.br/
Portal do Espírito	http://www.espirito.org.br/index.asp	Portal do Gaúcho	http://www.portaldogaicho.com.br/
Portal do Governo Brasileiro	http://www.brasil.gov.br/	Portal Tattoo	http://www.portaltattoo.com/

PortalDF	http://home.portaldf.com.br/consultas/	Primeira Mão	http://www.primeiramao.com.br/
Radio Click	http://radioclick.globo.com/	SESC São Paulo	http://www.sescsp.com.br/sesc/home/index.cfm
Telefônica	http://www.telefonica.com.br/	Via	http://portal.via.com.br/index.php
VilaRé	http://www.vilare.com.br/		

Tabela 11: Portais ativos obtidos em consulta aos buscadores Google e Yahoo.

Num último esforço de ampliação do *corpus*, consultamos também a categoria Acesso Gratuito no diretório Cadê, notando que não poucos dos sites de provedores ali listados compunham igualmente as listas anteriores. Alguns sites estavam inscritos também na categoria de Portais. Desse modo, as novas entradas foram incorporadas ao *corpus* estendido, sendo tratadas nos levantamentos efetuados em conjunto com a lista anterior. A lista de *sites* é exibida na Tabela 12. Com isso, nosso *corpus* de análise totalizou 63 *sites* Internet que, de algum modo evidente, se assumem como portais. Ainda que tratemos todos do mesmo modo, nossa análise procurará abordar o *corpus* em dois grupos, o primeiro sendo considerado o grupo de referência.

Site	URL
Click21	http://www.click21.com.br/
Miracula	http://www.miracula.com.br
iTelefônica	http://itelefonica.terra.com.br/
Supersol	http://www.supersol.com.br/
Tutopia	http://www.tutopia.com.br/

Tabela 12: Portais adicionais obtidos no diretório Acesso Gratuito do Cadê.

O tipo de experimento que queremos conduzir requer um grupo de controle; em outras palavras, devemos igualmente aplicar o modelo a um grupo de *sites* que *não* se apresentam como portais. Neste caso, é esperado que as propriedades que venham a ser avaliadas não se mostrem relevantes para a caracterização desses *sites*; deverá ser igualmente esperado que tais instâncias tenham baixo grau de pertinência ao conjunto universo dos portais. Isso é necessário para que possamos determinar, em confronto com os resultados apurados no grupo de controle, que os eventuais resultados positivos obtidos no *corpus* não sejam acidentais.

Para grupo de controle qualquer *site* que não se assuma como um portal deve ser suficiente. O conjunto universo de candidatos, portanto, é muito grande, exigindo algum critério de seleção. O ideal, naturalmente, é que esse critério conduzisse a um conjunto representativo do universo de *websites* destinados ao público brasileiro. Há inúmeras

abordagens que podem ser tentadas, com diferentes graus de validade estatística. Como buscamos tão somente um grupo de controle, decidimo-nos pelo critério mais simples, cuja validade estatística é, seguramente, duvidosa. Para isso inventariamos as diferentes categorias de *sites* listadas no diretório Cadê, que apresenta a distribuição quantitativa de entradas listada na Tabela 13, onde as subcategorias não estão representadas. Na seleção, as categorias Internet, Regional, Saúde e Fontes de Referência foram ignoradas, simplesmente porque abrangem uma quantidade de *websites* incontrolável. Caso uma subcategoria seja responsável por mais de 80% das entradas da categoria, ela foi selecionada; caso contrário, uma das subcategorias foi selecionada ao acaso. Por fim, para cada subcategoria escolhida foi selecionada ao acaso uma entrada para cada 40 mil ou fração, resultando em 15 *sites* escolhidos para grupo de controle.

Categoria	Entradas	Categoria	Entradas
Artes e Cultura	41.157	Compras e Serviços	51.878
Notícias	5.477	Business-to-business	75.161
Lazer	33.717	Finanças	532
Esportes	14.107	Governo	2.313
Educação	4.299	Informática	6.219
Ciência	18.074	Internet	4.455
Sociedade	82.769	Regional	305.023
Saúde	4.968	Fontes de Referência	634

Tabela 13: Distribuição de entradas por categoria no Cadê.

O elenco de *sites* arbitrariamente escolhidos para grupo de controle é apresentado na Tabela 14. Com isso, obtivemos um conjunto de 78 *sites* Internet, aos quais submetemos o modelo elaborado, operação realizada ao longo do mês de dezembro de 2005. Novamente, é de se assinalar que a execução do mesmo modelo a este conjunto de *websites* em outros momentos do tempo não deve alterar significativamente o resultado. Se isto ocorrer, o modelo poderá ser considerado aceitável para a finalidade à qual se propõe.

Site	URL	Site	URL
Tears for Fears	http://www.decadadamusica.hpg.ig.com.br/interna17.htm	APM	http://www.escolacatavento.com.br/
ProArte	http://www.proartegaleria.com.br/	Semp Toshiba	http://www.semptoshiba.com.br/
Jornal da Secretária	http://welcome.to/jornalsecretaria	Alfredo Rocha	http://www.alfredorocha.com.br/

Cartoon Brasil	http://www.cartonbrasil.com.br/	Grandes Aventuras	http://www2.uol.com.br/voltadasamericas/index.shl
Pladin	http://www.planetadinheiro.com.br	Paintball	http://www.riopaintball.com.br/
C & S	http://www.ces.com.br/	VOD	http://vod.prav.unisinos.br/
UnipDireito	http://groups.msn.com/unipdireito/_homepage.msnw?pgmarket=pt-br	Fantasma	http://sites.uol.com.br/danielbo/
Polo	http://www.terenos.com/		

Tabela 14: *Sites* utilizados como grupo de controle.

3.1.3. As propriedades temáticas dos portais – uma primeira hipótese

Embora o *press-release* de divulgação da pesquisa já citada (RELATÓRIO, 2005) não indique o critério utilizado para categorizar os *sites* acessados pelos usuários da sua amostra, é razoável supor que os chamados “portais de interesse geral” estejam representados na amostra ilustrada na Tabela 10. Se cruzarmos esta amostra com a pesquisa de audiência da NetRatings (<http://www.netratings.com>) em nove outros países que não o Brasil, notaremos a presença da AOL e da MSN entre os *sites* mais acessados pelos usuários da Internet. Portanto, a categoria utilizada pelo IBOPE deve estar relacionada aos *sites* da amostra.

De uma forma geral, trata-se de empreendimentos comerciais que baseiam sua expectativa de receita em dois negócios bem definidos, como já assinalamos: fornecimento de acesso à Internet e venda de espaço publicitário, onde, muito provavelmente, àquele está a serviço deste ao proporcionar a quantidade de usuários constantes necessários à valorização do espaço. Portanto, para serem viáveis, tais *sites* precisam fixar usuários nos seus domínios, sejam assinantes, sejam quaisquer outros. Isso é feito através do fornecimento de conteúdo amplo e bastante variado. O UOL, por exemplo, “está organizado em 42 estações temáticas, com mais de mil diferentes canais de notícias, informação, entretenimento e serviços” (UOL, 2006). A variedade do conteúdo está diretamente relacionada à variedade do público que os portais pretendem atingir: a própria natureza da publicidade na Internet, que exige milhões de visualizações (*page loads*) para ser eficaz, requer que se atinja um público tão vasto quanto possível. Além disso, esse conteúdo deve de algum modo corresponder às expectativas dos usuários; eles devem satisfazer às suas necessidades de informação e entretenimento, garantindo sua permanência nos limites dos portais.

Esse raciocínio nos fornece um bom ponto de partida para procurarmos algumas das características dos portais – pelo menos aquelas características temáticas típicas desse gênero. Podemos simplesmente procurar determinar quais são os interesses e expectativas dos usuários (brasileiros) da Internet, aquilo que os move em direção ao ciberespaço, e postular que tais interesses devem ser atendidos pelo espectro temático dos portais.

O melhor lugar para começarmos a procurar por tais interesses são as pesquisas da NetRatings. O *press-release* já citado apresenta pesquisa que procurou determinar o que os usuários fazem na Internet, o que foi feito determinando-se a razão entre o tempo gasto pelos usuários nas várias categorias pesquisadas e o tempo total de uso da Internet. O resultado representa o percentual do tempo total gasto pelos usuários em cada categoria, conforme exibido na Tabela 1, apresentada na seção *Os portais: a restauração da cultura do espetáculo*, na introdução e são uma boa representação dos interesses dos usuários, podendo, portanto, ser utilizados para determinar algumas das características temáticas esperadas dos portais de interesse geral.

Ao considerarmos aquela Tabela, supomos que por “comunidades” certamente se conceba *sites* especializados em grupos de interesse e relacionamentos interpessoais como o Orkut; este tipo de promoção dos relacionamentos pessoais tradicionalmente se faz através de salas de *chats*, como ou sem a utilização de software específico. Mais recentemente temos observado igualmente a disseminação de *sites* de encontros e a popularização dos *blogs*. Estes últimos têm sido utilizados como pontos de encontro e de troca da informações entre os usuários da Internet. O Orkut, por exemplo, é um dos grandes fornecedores de ambientes de *blog*, daí incluirmos esse gênero na grande categoria “comunidades”.

A categoria “classificados e leilões” está relacionada a *sites* especializados como o E-Bay e, no Brasil, o Arremate. No entanto, temos observado que inúmeros *sites*, em especial aqueles destinados ao “grande público”, incorporam seções específicas destinadas à oferta de produtos para venda, um tipo especializado de publicidade, além de fornecerem serviços de classificados. Assim, podemos incluir essa categoria nos “temas” tipicamente esperados nos portais.

Desse modo, a pesquisa do IBOPE nos permite elaborar uma primeira hipótese sobre o conteúdo temático típico para os portais de interesse geral, a saber: inclusão de seções específicas para *chats*, *blogs*, jogos, busca, e-mail e produtos para venda, incluindo todas as categorias cujo interesse dos usuários brasileiros é notável na pesquisa. Obviamente, pelo que discutimos anteriormente sobre a “missão” dos portais, a publicidade pode e deve ser considerada uma característica típica dos portais.

Um outro modo de determinarmos os interesses dos usuários seria avaliarmos as expressões de busca mais utilizadas nos dispositivos, na suposição de que elas exprimem de maneira direta os interesses dos usuários – e satisfazer esses interesses é a finalidade básica dos portais. Infelizmente, os dispositivos mais populares (o Google, o Yahoo e a MSN) não publicam estatísticas a propósito de suas buscas. Durante alguns meses e enquanto se utilizou de um dispositivo próprio, o Terra (<http://www.terra.com.br/busca>) publicou uma lista dessas. Ainda que pudesse estar desatualizada em face do interesse atual (sua última publicação data de 2 de outubro de 2004), ela poderia ser utilizada como hipótese. No entanto, os editores do portal confessadamente censuravam a lista, eliminando termos considerados “irrelevantes”. Embora tenhamos uma boa idéia do que sejam esses termos “irrelevantes”, é conveniente buscarmos uma outra fonte de informações, pelo menos a título de comparação.

Encontramos uma possível alternativa no pequeno buscador Exploora (<http://www.exploora.com.br>). Durante um certo tempo, esse buscador – fruto de pesquisa de estudantes de graduação em computação, que propunha critérios de avaliação da intenção manifesta na expressão de busca distintos dos convencionalmente utilizados – publicou uma lista irrestrita sobre o interesse dos seus usuários, exibida na Tabela 15. Esse dispositivo nunca foi popular; portanto, a lista não pode ser considerada representativa do universo de usuários brasileiros. No entanto, como estamos interessados apenas numa hipótese, ela pode ser utilizada para comparação com a lista do Terra, ainda que de modo cauteloso.

1	Blog erótico	2	Fotos sexo	3	Travestis	4	Amadoras
5	Sexo	6	Música	7	MP3	8	Revista
9	Pornô	10	Jornal	11	Sylvia Saint	12	Chat
13	Detran	14	Prefeituras	15	Downloads	16	Viviane Araújo
17	TV	18	Rádio	19	Jogos	20	Blogs

Tabela 15: Expressões mais utilizadas nas buscas do Exploora em julho de 2004.

Se compararmos essa lista com aquela publicada no Terra, notaremos extensa concordância entre ambas, se excluirmos o interesse manifesto no Exploora por temas relacionados à sexualidade. Logo, a lista pode ser considerada minimamente representativa. Assim, essa pesquisa nos permite ampliar a lista de temas de interesse dos usuários, que possivelmente deveriam compor o conteúdo temático típico dos portais, a saber: sexo, música, revista, jornal, personalidades, *downloads*, TV e rádios.

3.1.4. A metodologia de coleta e análise dos dados

Isso nos basta como uma primeira aproximação do problema. A Tabela 16 lista todas as propriedades temáticas sugeridas para a avaliação dos portais. Supomos que as linhas da tabela representem as restrições às quais deve se submeter um *site* para ser considerado um portal. Para cada um dos elementos temáticos devemos buscar seções (isto é, um conjunto de páginas) cujo conteúdo temático possa ser identificado como um membro do conjunto T . Isto é, estamos supondo que a função de pertinência $\mu_t(x)$ de uma propriedade ao tema T é dada pela proposição:

$$(i) \quad \mu_t(x) = \exists x \mid x \text{ is } t, t \in T$$

Elementos temáticos	Elementos temáticos
Chat	Busca
Blog	E-mail
Sexo	Downloads
Música	Jornal
Personalidades	Revista
TV	Produtos para venda
Rádios	Publicidade
Jogos	

Tabela 16: Primeira hipótese de propriedades temáticas dos portais.

Na prática, estaremos avaliando tantas proposições $x \text{ is } t$ quantos forem os membros do conjunto T para obter a função de pertinência $\mu_t(x)$ de x em t . Para obtermos o valor de verdade da proposição (i) devemos encontrar, na primeira página do portal, um *link* cujo texto seja uma palavra do domínio vocabular da propriedade que estamos avaliando e que aponte para um conjunto de páginas tal que possamos afirmar que se trate de uma seção permanente do portal. Como se pode notar facilmente, a proposição (i) é vaga e altamente imprecisa; aliás, não podemos sequer afirmar que se trate de proposição atômica, onde não seja possível decompô-la em outras proposições. No entanto, para o propósito que nos move – conjecturar sobre as propriedades temáticas de um portal – ela deve ser suficiente, principalmente porque está destinada a ser avaliada por dispositivos humanos, habituados a raciocinar com base em proposições vagas e imprecisas.

Uma perda de confiabilidade ocorrerá por conta de uma outra decisão que fomos obrigados a tomar. Vimos na seção anterior que o modelo de Siler e Buckley utilizado como base para o modelo de reconhecimento de instâncias de gêneros semióticos requer que as várias proposições restritivas a propósito das propriedades da categoria sejam submetidas a uma quantidade de observadores no mínimo representativa do universo de usuários. Evidentemente, numa pesquisa acadêmica sem financiamento próprio nem tempo suficiente como a nossa, isso não é viável. Por outro lado, submeter o modelo a um conjunto arbitrário de observadores sem ter em vista a representatividade estatística da amostra, um direcionamento viável nesse momento, aumentaria enormemente o esforço gasto sem, contudo, agregar confiabilidade aos resultados. Portanto, decidimo-nos por uma outra simplificação, postulando a possibilidade de um observador ideal – nós mesmos – cujas respostas presumivelmente serão próximas o suficiente de uma amostra ergódica para tornar confiáveis os resultados. Isso, evidentemente, é falso. Novamente, nas presentes condições é o melhor que podemos obter. Sacrificaremos, portanto, a precisão em benefício da viabilidade prática, submetendo o modelo a um único observador.

Portanto, o levantamento se dará submetendo cada uma das proposições a um observador ideal para todas as instâncias do *corpus*, admitindo-se tão somente dois valores de verdade para cada juízo. Com isso, será possível calcular a relevância R de cada propriedade, utilizando para isso a média aritmética das respostas obtidas em todas as i instâncias. Para cada instância, será possível também calcular sua prototipicidade P , através da média aritmética das respostas obtidas em todas as suas n propriedades. Em outras palavras, queremos obter o resultado:

$$(ii) R_t = \sum \mu_t(x) / i$$

$$(iii) P_t = \sum \mu_{tc}(x, y) / n$$

Tanto o cálculo (ii) quanto (iii) deverão obter valores numéricos. No entanto, estamos mais particularmente interessados no valor lingüístico de verdade das proposições simplesmente porque ele corresponde melhor ao modo dos usuários categorizarem. Como vimos anteriormente, não dizemos que uma xícara tem probabilidade 0.7 de ter alça; nos simplesmente afirmamos que ela *normalmente* tem alça... Assim, precisamos encontrar um meio de *difundir* (*fuzzify*) o resultado dos cálculos de relevância e prototipicidade. Como se verá na análise, a difusão dos valores de relevância e prototipicidade nos permitirão tirar

interessantes conclusões a propósito do *corpus*, conclusões que não necessariamente seriam obtidas com os valores numéricos.

O valor lingüístico de verdade de uma proposição é sempre *convencional*; isto significa que podemos arbitrar um conjunto discreto de valores de verdade com qualquer quantidade de elementos. Para simplificar nossos cálculos, definiremos o seguinte conjunto de valores de verdade para cada proposição:

$$V = \{ \text{falso, pouco verdadeiro, mais ou menos verdadeiro, muito verdadeiro, verdadeiro} \}$$

Podemos convencionar que um valor de pertinência entre 0 e 0.5, exclusive, seja considerado “pouco”; mas quão “pouco”? E sobretudo como exprimir o conhecimento de que o “pouco” representado por 0.1 é menor que o “pouco” representado por 0.2? Além disso, a medida que o valor de “pouco” decai, aumenta o de “médio”; afinal, algo que é “pouco” num grau, digamos, de 0.4 é “médio” em algum outro grau. O mesmo vale para o valor de “muito”.

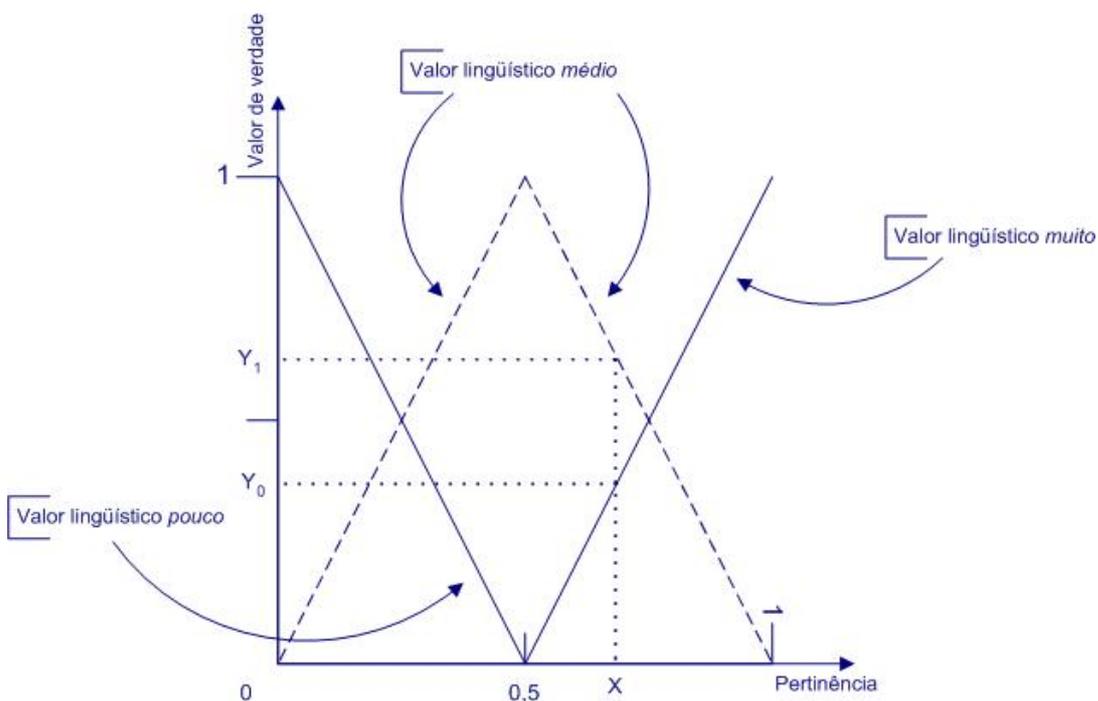


Ilustração 17: Convenções para os valores lingüísticos de verdade.

Uma boa maneira de convencionar um valor para “pouco” é arbitrar para ele uma função contínua decrescente maior que 0 e menor que 0.5; reciprocamente, “muito” pode ser arbitrado como uma função contínua crescente maior que 0.5 e menor que 1. Um valor “médio” (algo “mais ou menos” verdadeiro) pode ser definido por uma função complexa, crescente entre 0 e 0.5 e decrescente entre 0.5 e 1, exclusive, como pode ser visto na

Ilustração . Nesta figura, os valores de pertinência 0 e 1 estão, naturalmente, reservados para os valores lingüísticos “falso” e “verdadeiro”, respectivamente. Essa abordagem tem a vantagem de ser relativamente “intuitiva”, além de permitir um refinamento crescente. Podemos simplesmente aumentar a granularidade dos valores intermediários, utilizando funções complexas como aquela proposta para “médio” na Ilustração .

Como se pode notar pelo gráfico, nosso problema é calcular os valores para “pouco”, “médio” ou “muito”, dado um valor de pertinência. No exemplo da Ilustração , o valor de pertinência x (que pode representar tanto um valor de relevância quanto de prototipicidade) implica em dois componentes y_0 e y_1 para seus valores lingüísticos de “muito” e “médio”, respectivamente.

Por exemplo, se $x = 0.6$, segue-se que “médio” receberá um valor 0.8 e “muito”, um valor 0.2, correspondendo ao conhecimento intuitivo que 0.6 é tanto “médio” quanto “muito” (numa escala entre 0 e 1), mas está mais próximo dos valores medianos. Isso pode ser interpretado da seguinte forma: dado um valor x de pertinência, existe uma probabilidade p_{po} de este valor ser considerado “pouco”, uma probabilidade p_{me} de ser considerado “médio” e p_{mu} de ser considerado “muito”. Há inúmeras maneiras de calcular os valores do eixo y do gráfico dado um valor no seu eixo x ; a mais simples decorre da geometria do próprio gráfico e é ilustrada na Tabela 17.

Valor lingüístico	$x \leq 0.5$	$x \geq 0.5$
Pouco	$1 - 2x$	0
Médio	$2x$	$1 - (2x - 1)$
Muito	0	$2x - 1$

Tabela 17: Cálculo do valor lingüístico de verdade.

3.1.5. Primeira avaliação da amostra – a relevância das propriedades

O *corpus* completo, incluindo o grupo de controle, foi submetido ao modelo de reconhecimento durante o mês de dezembro de 2005. A média das respostas obtidas para cada proposição restritiva em todas as instâncias, isto é, a medida da relevância das propriedades sugeridas, pode ser vista na Tabela 18, ordenada em ordem crescente pela coluna Pertinência. A coluna Referência lista a média dos resultados obtidos considerando-se apenas os sites

listados no diretório Cadê como portais; a coluna Estendido considera apenas a média dos *sites* obtidos em consulta ao Google e ao Yahoo e nos *sites* adicionais obtidos no diretório Acesso Gratuito do Cadê. A coluna Controle refere-se aos resultados obtidos apenas considerando-se o grupo de controle. Por fim, a coluna Pertinência lista a média entre as colunas Referência e Estendido e pode ser considerada a medida da relevância das propriedades no *corpus*.

Elementos temáticos	Referência	Estendido	Pertinência	Controle
Personalidades	0.20	0.14	0.17	0.00
TV	0.20	0.14	0.17	0.00
Rádios	0.20	0.19	0.20	0.00
Revista	0.27	0.17	0.22	0.00
Sexo	0.43	0.10	0.27	0.00
Blog	0.50	0.17	0.34	0.00
Chat	0.53	0.21	0.37	0.00
Música	0.53	0.26	0.39	0.13
Downloads	0.50	0.53	0.51	0.07
Jogos	0.67	0.39	0.53	0.00
E-mail	0.73	0.39	0.56	0.00
Produtos para venda	0.50	0.63	0.56	0.00
Jornal	0.63	0.51	0.57	0.07
Publicidade	0.73	0.60	0.66	0.00
Busca	0.83	0.60	0.71	0.13

Tabela 18: Relevância das propriedades arbitradas no *corpus*.

A primeira observação que podemos fazer refere-se ao fato de os valores obtidos na avaliação do *corpus* de referência foram, em geral, significativamente maiores que os obtidos do *corpus* entendido. Isso seria de se esperar. Como dissemos, o *corpus* de referência foi obtido de lista explícita de portais, onde a inscrição é voluntária e remunerada. Espera-se, portanto, que os *sites* ali inscritos pretendam apresentar-se mais fortemente como portais que aqueles com apenas fazem referência à expressão “portal Internet” na sua página principal. Uma possível consequência para esse fato é a sugestão de que as propriedades arroladas sejam realmente uma boa hipótese para as características dos portais. Isso pode ser corroborado pela comparação dos resultados com a coluna Controle. É notável aí o valor de pertinência 0 ou muito próximo de 0 para quase todos as propriedades do domínio temático. Isso, porém, ficará mais claro se considerarmos o valor lingüístico de verdade para a medida da relevância das propriedades temáticas, conforme exibido na Tabela 19.

Tema	Pouco	Médio	Muito	Tema	Pouco	Médio	Muito
Personalidades	0.66	0.34	0.00	Downloads	0.00	0.97	0.03
TV	0.66	0.34	0.00	Jogos	0.00	0.94	0.06
Rádios	0.61	0.39	0.00	E-mail	0.00	0.87	0.13
Revista	0.56	0.44	0.00	Produtos para venda	0.00	0.87	0.13
Sexo	0.47	0.53	0.00	Jornal	0.00	0.86	0.14
Blog	0.33	0.67	0.00	Publicidade	0.00	0.67	0.33
Chat	0.26	0.74	0.00	Busca	0.00	0.57	0.43
Música	0.21	0.79	0.00				

Tabela 19: Valor lingüístico de verdade da relevância das propriedades.

A premissa que adotamos simplifica bastante a interpretação dos resultados exibidos na tabela. Assumimos que os temas mais freqüentemente oferecidos pelos portais são aqueles que provavelmente mais mobilizam os usuários. Assim, a resposta à questão – vamos aos portais para fazer o quê? – pode ser traduzida como se segue. Os usuários utilizam os portais:

- i *Freqüentemente* para realizar buscas, conhecer produtos, ler notícias, procurar ofertas e enviar *e-mails*;
- ii *Normalmente* para jogar, realizar *downloads* de aplicativos, ouvir música, participar de *chats*, interagir com *blogs* e realizar fantasias sexuais;
- iii *Às vezes* para ler notícias de variedades, ouvir rádios, conhecer as últimas novidades da TV e descobrir as mais recentes fofocas sobre as personalidades mais conhecidas da mídia.

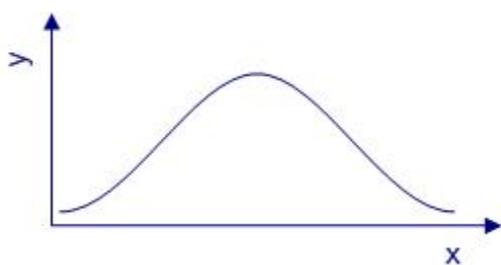


Ilustração 18: Uma curva de distribuição normal.

Um outro argumento em favor da caracterização adotada pode ser posto. Em se tratando de estudo estatístico de uma população de *sites* que esperamos ser representativa pelo menos da categoria dos portais de interesse geral, seria de se esperar que a distribuição das suas propriedades obedecessem à distribuição padrão esperada de uma curva normal, conforme mostrado na

Ilustração . Este tipo de curva é encontrado quando analisamos determinadas variáveis em populações normais. É, por exemplo, esperado que, na análise da variável “idade” de uma amostra de população humana “saudável”, a distribuição da quantidade de membros pela

variável obedeça ao padrão da curva normal. No nosso caso, isso corresponderia à suposição intuitiva de que a maioria das propriedades se concentrassem nas faixas de valores médios, deixando os extremos para quantidades menores de ocorrências.

Uma distribuição normal de valores lingüísticos de verdade corresponderia, por exemplo, a um conjunto de valores distribuídos de modo bem próximo ao gráfico teórico exibido na Ilustração , com ocorrências mais ou menos uniformemente distribuídas ao longo da curva esperada do gráfico. No caso dos dados exibidos na Tabela 19, a distribuição dos valores de verdade é vista na Ilustração , uma curva muito próximo da curva teórica esperada para uma distribuição normal, salvo por uma ligeira distorção do valor lingüístico de verdade para “muito”. Essa distorção não é suficiente para invalidar a análise: ela decorre simplesmente da escolha do próprio *corpus*.

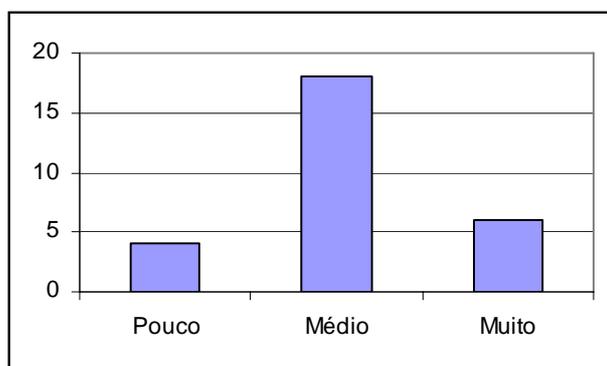


Ilustração 19: Distribuição dos valores predominantes de verdade.

O que discutimos é suficiente para concluirmos, ainda que alguma cautela, que há boas indicações de que as propriedades selecionadas possam ser utilizadas efetivamente para caracterizar o domínio temático dos portais de interesse geral. Algumas delas – como, por exemplo, rádios, personalidades, televisão e revista – talvez não precisem ser consideradas porque bem pouco incidentes no *corpus*. Outras, contudo, podem ser buscadas a partir da observação das regularidades eventualmente existentes nessa ou em outra amostra que venha a ser eventualmente encontrada. Porém, pelo menos nesse momento e a essa altura da pesquisa, o que já encontramos é suficiente para caracterizarmos preliminarmente essa interessante categoria de *sites* da Internet.

3.1.6. Segunda avaliação – a prototipicidade da amostra

Trata-se agora de avaliar o *corpus* quanto à prototipicidade de cada uma das instâncias. Como dissemos anteriormente, a idéia é calcular a distância lógica entre uma instância dada e um protótipo ideal do domínio temático dos portais, definido arbitrariamente como um exemplar cuja avaliação de suas propriedades temáticas seja igual à unidade. Essa operação deverá equivaler a uma afirmação a propósito do grau de pertinência de uma instância ao conjunto universo do gênero, pelo menos quando considerado o seu domínio temático. A Tabela 20 ilustra o resultado do levantamento no *corpus* selecionado.

Corpus de referência	Pertinência	Corpus estendido	Pertinência
UOL	1.00	iTelefônica	0.93
Terra	1.00	Oi	0.73
Globo.com	0.93	Click21	0.53
iG	0.87	Piq	0.47
AOL	0.93	Âmbito Jurídico	0.47
MSN Brasil	0.80	Via	0.40
POP	0.73	Elnet	0.40
Ibest	0.73	Gaúcho	0.40
Inter .net	0.67	Supersol	0.40
Matrix 2005	0.67	VilaRé	0.27
Yahoo	0.60	SESC	0.27
Cidade Internet	0.60	Radio Click	0.27
O Site	0.53	Tutopia	0.27
iTudoMais	0.53	Firemasters	0.20
Mundo Perdido	0.47	Primeira Mão	0.20
Starmedia	0.47	Polycol	0.20
CrazyMania	0.47	Jus	0.20
Terravista	0.53	AZ	0.20
Sociedade Digital	0.47	Educacional	0.20
Portal Free	0.40	Cidadania	0.20
BOL	0.47	Correios	0.13
BemRápido	0.27	PortalDF	0.13
Pootz	0.27	Portal do Espírito	0.13
Portal Nacional	0.20	Google	0.13
JP	0.13	Miracula	0.13
Cadê	0.13	Tattoo	0.07
Esquinas	0.00	Arteplástica	0.07
Rio Brasil	0.07	CAPES	0.07

Eg	0.00	Brasil	0.07
LatinLook	0.00	Embratel	0.00
		Indústria	0.00
		Telefônica	0.00

Tabela 20: Prototipicidade das instâncias da amostra.

Como se pode facilmente notar, apenas o UOL e o Terra incluíam, quando da realização do levantamento, todos os temas escolhidos em seu universo de canais e serviços. Porém, os portais Globo.com, AOL, o iG e possivelmente o MSN Brasil alcançaram alto grau de prototipicidade de domínio temático. O significado disso é que, possivelmente, se os usuários com um mínimo de experiência nos portais forem indagados do que seja o gênero, há boa probabilidade que respondam com o exemplo de uma dessas instâncias. Naturalmente, deve-se levar em consideração que avaliamos somente um dos aspectos constitutivos de um gênero digital (o seu domínio temático, o seu propósito, etc.); outros aspectos deveriam ser considerados para que pudéssemos fazer previsões mais precisas sobre o comportamento dos usuários. Do mesmo modo, o resultado obtido pelo grupo de controle, ilustrado na Tabela 21, corrobora a hipótese de que as propriedades foram selecionadas apropriadamente.

Instância	Pertinência	Instância	Pertinência
Fantasmas	0.20	APM	0.00
Pladin	0.07	C&S	0.00
Tears for Fears	0.07	VOD	0.00
ProArte	0.07	Semp Toshiba	0.00
Paintball	0.00	Polo	0.00
Jornal da Secretária	0.00	Alfredo Rocha	0.00
Carton Brasil	0.00	UnipDireito	0.00
Grandes Aventuras	0.00		

Tabela 21: Prototipicidade do grupo de controle.

Um outro aspecto que devemos considerar é, novamente, a distribuição dos valores lingüísticos de verdade obtidos na avaliação. Como se pode notar na Ilustração , tais valores parecem razoavelmente obedecer à distribuição normal esperada, salvo por uma razoável distorção dos valores lingüísticos de verdade para “pouco”. Tanto esses resultados quanto os discutidos na seção anterior parecem sugerir que a distribuição de instâncias num gênero digital qualquer deve obedecer à curva de distribuição normal, com a maioria de exemplares situados em valores medianos de prototipicidade e relevância média das propriedades.

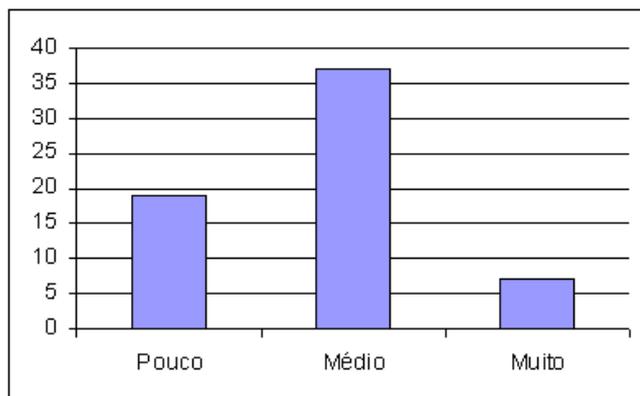


Ilustração 20: Outra distribuição dos valores predominantes de verdade.

Ainda que as limitações da pesquisa não nos autorizem qualquer otimismo, é razoável admitirmos que tanto a seleção do *corpus* quanto a das propriedades constitutivas dos portais foram adequadas. Isso significa que o prosseguimento da pesquisa deve se concentrar nos exemplares mais fortemente prototípicos do *corpus*, na expectativa de encontrar novas e mais

relevantes características para essa categoria de sites. Podemos concluir também que o modelo de prototipicidade baseado na lógica nebulosa se revelou adequado para a descrição das características constitutivas de um gênero digital, podendo ser aplicado a outros domínios. Além disso, o modelo foi capaz de revelar também um possível padrão para a distribuição dos exemplares de uma categoria. É certamente um desenvolvimento interessante observarmos outras categorias, visando determinar se todas obedecem ao mesmo padrão de distribuição.

O que vimos discutindo corrobora a hipótese de que aqueles interesses que mais mobilizam os usuários em direção à Internet são os mesmos privilegiadamente atendidos pelos portais de interesse geral. Isso é coerente com a “missão” primária desses empreendimentos: fazer dinheiro através da publicidade, o que, nesse meio em particular, requer público vasto e diversificado. Dito de outro modo, o conteúdo temático diversificado e explicitamente orientado às demandas prévias dos usuários funcionam como o primeiro e mais crítico atrator retórico dos portais. Trata-se de “potes de mel” destinados a atrair os usuários para a órbita desse aglomerado de *sites*. No entanto, como todos os “potes de mel”, ele é apenas um chamariz para a presa: outras estratégias retóricas devem ser utilizadas para *fixar* os usuários na teia da aranha. É de algumas dessas estratégias que trataremos nas próximas seções.

3.2. O controle da palavra do outro nos dispositivos de busca

No capítulo 1. *O ciberespaço: interação à velocidade da luz* procuramos interpretar o ciberespaço como um território social avesso ao controle, onde a liberdade poderia se

expressar plenamente, para o bem ou para o mal. Argumentamos que essa tendência é sustentada pelas próprias características tecnológicas da Internet, em especial pelo fato de se tratar de uma rede sem centro e virtualmente ilimitada, o que inviabiliza a adoção de mecanismos de controle eficazes. No entanto, assinalamos também a persistência de uma contra-tendência centrada nas tentativas de controle da circulação irrestrita da informação. Evidentemente, esse esforço de controle do discurso tem inúmeras vertentes com graus variados de eficácia. Uma das mais eficazes pela sua viabilidade consiste, como assinalamos naquela ocasião, em exercer algum tipo de controle nesses *centros de informação* que são os dispositivos de busca. Naquela ocasião abordamos apenas os controles externos mais grosseiros representados por restrições governamentais a determinadas buscas ou *sites*.

O controle exercido nos dispositivos de busca, pelas razões apontadas, é um mecanismo de controle do discurso que procura censurar a enunciação por meios indiretos, controlando a palavra do outro: ao restringir o acesso a documentos que incorporam temas “proibidos”, o que se quer é restringir a vigência no discurso de múltiplas vozes, perigosas porque diferentes, substituindo-as pelo discurso *monótono*, obliterando-as para o território do não-dito (porque não-lido). É essa uniformidade de um discurso plano o que se quer alcançar com a censura. No entanto, os mecanismos que abordamos sucintamente naquele capítulo sempre se revelaram de curto alcance a longo prazo, sendo, portanto, ineficazes. Muito mais eficaz que a censura é a auto-censura, sobretudo quando assumida como expressão da liberdade para a palavra do outro...

Nesta seção abordaremos o funcionamento desses mecanismos de auto-censura incorporados aos dispositivos de busca, assumindo que eles têm como efeito não apenas o controle do próprio discurso, como também o do discurso do outro – o usuário que busca – ao delimitar (e, implicitamente, *limitar*) o universo de discursos a que temos acesso no ciberespaço. No segundo capítulo abordamos o papel desempenhado pelos dispositivos de busca no acesso à informação na Internet, papel indispensável desempenhado na prática por apenas três buscadores – o Google, o Yahoo e o MSN, com 49,2%, 23,8% e 9,6% das 213 milhões de buscas diárias, respectivamente (ver SULLIVAN, 2006). O universo alvo dessas buscas é muito grande: em novembro de 2004, o Google anunciava que já havia cerca de 8 bilhões documentos indexados em suas bases de dados (SULLIVAN, 2004). Mesmo supondo que os índices nas bases de dados de todos os demais dispositivos estejam contidos neste conjunto, é fácil notar que sem o auxílio dos search engines seria impraticável o acesso a qualquer parcela de documentos disponíveis na *web*. Portanto, do ponto de vista prático, o acesso à informação na Internet é controlado por esses três dispositivos. O que sustentamos é

que tais dispositivos controlam não apenas o *acesso* à informação, mas também, em grande medida, sua própria *produção*. Isso decorre da concepção de linguagem com a qual estamos privilegiadamente trabalhando – a de Bakhtin.

Na tradição bakhtiniana todo discurso é orientado à audiência, da conversa cotidiana ao poema lírico: como já assinalamos, “o traço essencial (constitutivo) do enunciado é o seu *direcionamento* a alguém, o seu *endereçamento*” (BAKHTIN, 2003: 301). Isto significa que determinadas suposições sobre o interlocutor afetam as escolhas disponíveis para o locutor, determinando seu próprio enunciado. Este seria um traço de toda interação comunicativa. Ora, se supusermos que a natureza artificial da inteligência dos dispositivos de busca não afeta a interação comunicativa com eles, que tem as mesmas características básicas de interação do mesmo tipo entre seres humanos, estaremos admitindo também que tanto o enunciado que elaboramos numa busca quanto a resposta dada pelo dispositivo são afetados por certas suposições sobre as partes envolvidas. Na resposta a uma indagação, o dispositivo humano elabora uma hipótese sobre o que o locutor quer de fato saber e responde com base nela; na elaboração dessa hipótese é levado em consideração tanto o conhecimento de mundo disponível quanto uma concepção prévia do próprio enunciador da pergunta (a audiência da resposta). Assim, um dispositivo computacional que simule em alguma medida o interlocutor humano, ao elaborar sua resposta, deve levar em consideração uma concepção prévia do que o enunciador da pergunta deseja saber.

Essa hipótese tem pelo menos uma decorrência que pode ser facilmente testada por experimentos: diferentes dispositivos de busca, se programados com base em diferentes concepções da audiência, devem responder com resultados distintos às mesmas buscas, independente dos “conhecimentos” disponíveis (a base de dados consultada). Isso significa que a resposta fornecida pelo dispositivo não é “neutra”: ele não responde com o que queremos saber, mas com aquilo que ele mesmo entende que queremos saber. Essa é uma das hipóteses que submetemos a um teste prático, cujos resultados serão discutidos nesta seção. Isso, porém, tem outras implicações mais importantes para o problema do controle do discurso.

Se “todo enunciado concreto é um elo na cadeia da comunicação discursiva de um determinado campo” (BAKHTIN, 2003: 296), meu enunciado de hoje é também uma resposta a um conjunto de enunciados de outrem, com quem dialoguei ontem. Ora, ao controlar em grande medida as vozes com quem sou capaz de dialogar na Internet, os dispositivos de busca, a longo prazo, são capazes de controlar também, em alguma medida, o meu próprio discurso. Se o ciberespaço fosse o único espaço social de interação discursiva disponível, os resultados

das buscas nos *search engines* abarcariam praticamente *todos* os enunciados com os quais sou capaz de dialogar... cremos que essa é uma razão muito boa para tornar essencial o estudo dos dispositivos de busca do ponto de vista discursivo.

Por fim, queremos assinalar que dedicar toda uma seção aos dispositivos de busca numa dissertação sobre os portais de interesse geral foi plenamente justificado no capítulo anterior. Dentre os “tópicos de conversação” selecionados como próprios de serem abordados pelos usuários na sua interação com os portais, a busca de informações foi considerada na análise o tópico mais relevante, com 83% de incidência no *corpus* de referência. Pelas premissas ali adotadas, um observador qualquer teria a probabilidade de 0,43 de responder que a busca de informações é *frequentemente* um dos “tópicos de conversação” dos usuários com os portais de interesse geral. A questão que se põe, então, é indagar *como* os portais de interesse geral incorporam o diálogo com os dispositivos de busca no seu próprio universo discursivo. Em outras palavras, trata-se de perguntar se o resultado de uma busca num mesmo dispositivo é de algum modo afetado pelo fato de ser realizada no interior de um portal.

3.2.1. Diferentes concepções dos dispositivos de busca

De uma forma geral, todos os dispositivos de busca utilizam a mesma estratégia para construir suas bases de dados (seu “conhecimento de mundo”): eles incorporam um robô que varre sistematicamente os endereços de rede da Internet em busca de servidores que disponibilizem documentos mediante requisição, tais como servidores *web*, *newsgroups*, servidores de arquivos, etc. Esses robôs compõem o chamado módulo de *crawling* do dispositivo. Isso significa que dispositivos de busca com a mesma capacidade computacional à disposição dos seus robôs, dado um período de tempo suficientemente longo, tenderão a dispor das mesmas bases de dados para suas consultas. Uma implicação disso é que é de se esperar que o resultado das buscas em dispositivos equivalentes não seja afetado pela base de dados disponível para suas consultas; esse resultado deve ser somente uma função dos critérios utilizados para sua classificação (indexação) e ordenação – e são esses critérios que darão conta da concepção de audiência programada naqueles dispositivos.

Qualquer um que já tenha utilizado um *search engine* saberá que o crítico nesses dispositivos não é o acervo – geralmente envolvendo centenas de milhões de documentos –

mas o critério de ordenação e apresentação dos resultados, isto é, a classificação da *relevância* das saídas. Uma simples busca pode envolver milhares de documentos como saída – e obviamente ninguém é capaz de inspecionar todos. Provavelmente, a maioria das buscas é concluída após uma fração desprezível das suas saídas ter sido inspecionada, o que significa que a maioria delas deve se resumir às primeiras páginas do resultado. Em pesquisa recente, a iProspect (IPROSPECT, 2006) concluiu que 88% dos usuários abandonam uma busca após consultar apenas a terceira página de resultados e 41% deles a abandonam já após a primeira página!

Isso implica que o conteúdo do acervo dos dispositivos é irrelevante para o resultado concreto da interação de busca não apenas porque diferentes dispositivos deverão ter o mesmo acervo à disposição ao longo do tempo, mas principalmente porque a quantidade de documentos coerente com a expressão buscada é sempre muito maior do que a nossa capacidade de inspecioná-los. Desse modo, o fator mais relevante na busca deve ser o critério de classificação do acervo e de ordenação das saídas para apresentação. É aqui que os projetistas de sistemas devem fazer algumas suposições críticas sobre o que os usuários querem obter quando interrogam o dispositivo em suas buscas no ciberespaço.

A idéia na origem dos dispositivos de busca é coletar, armazenar e eventualmente localizar o que se convencionou chamar “documentos relevantes”, isto é, aqueles documentos disponíveis na Internet nos quais o usuário da busca está interessado. Essa é uma tarefa impraticável, ainda que teoricamente possível: a quantidade de documentos disponíveis na Internet é muito grande e cresce continuamente. Além disso, uma fração não desprezível desses documentos é continuamente removida e atualizada. Esses fatores aliados à limitada capacidade computacional dos dispositivos fazem com que o conjunto de respostas de uma busca deva ser diferente em algum grau daquele conjunto (teórico) de documentos relevantes.

Seja R o conjunto de documentos relevantes no universo da Internet e $|R|$ o número de elementos em R ; seja D o conjunto de documentos obtidos por um dispositivo de busca e $|D|$ o número de elementos em D . O alvo do processo de busca deve ser, então, obter o conjunto intersecção de ambos, isto é: $R_d = R \cap D$. Chama-se *cobertura* a fração C de documentos relevantes recuperados na busca, isto é, a razão entre $|R_d|$ e $|R|$; chama-se *precisão* a fração P de documentos recuperados que é relevante, isto é, a razão entre $|R_d|$ e $|D|$. Como não é praticável obter $C = 1$, o objetivo dos projetistas desses dispositivos é aumentar o valor da precisão P para valores tão próximos da unidade quanto possível. Duas estratégias são aqui utilizadas. A primeira delas é de natureza estritamente tecnológica: a indexação dos documentos deve ser tal que assegure ao máximo possível que os diferentes assuntos

abordados num documento que eventualmente possam ser relevantes numa busca estejam acessíveis durante o processo. Como resultado, os dispositivos refinam continuamente sua capacidade de indexação, não apenas de modo a abarcar *todo* o conteúdo de um documento, independente do seu formato, como também de modo a catalogar elementos do “contexto” do documento, incluindo coisas como o diretório virtual onde reside, o nome do documento, documentos vizinhos, etc. A maior parte das pesquisas nesse domínio acadêmico visa elevar e acelerar essa capacidade de indexação e a localização dos documentos catalogados.

Uma outra estratégia, porém, é também adotada, embora tenhamos encontrado notavelmente poucos estudos sobre o assunto: trata-se de elevar a capacidade do dispositivo de determinar *o que* o usuário está buscando. Essa estratégia permite que os dispositivos selecionem do acervo de indexação somente aquelas entradas *relevantes* para o usuário. Quanto maior a capacidade de o dispositivo interpretar essa *intenção* do usuário, mais preciso será o resultado de P . Mesmo isso pode ainda resultar num valor muito elevado para D , mesmo que a razão R_d seja ótima (igual à unidade). Desse modo, os dispositivos devem igualmente classificar o *grau* de relevância de um documento e ordenar sua saída de acordo.

Esta estratégia é, certamente, mais relevante para uma pesquisa em Lingüística, já que é neste ponto que os projetistas dos dispositivos devem elaborar uma *hipótese interpretativa* sobre o *sentido* (expressão aqui empregada no “sentido” bakhtiniano, conforme discutimos no capítulo correspondente) da pergunta elaborada pelo usuário no diálogo de busca. É com base nessa hipótese interpretativa que os dispositivos avaliam as entradas encontradas nos seus índices e ordenam o resultado das buscas.

O critério clássico de relevância se baseia, grosso modo, no seguinte raciocínio. Suponhamos que o documento a qualquer aborde *principalmente* o tema x e *marginalmente* o tema y , enquanto que o documento b faça o inverso. Esse conhecimento deve ser expresso de algum modo no dispositivo, de tal modo que um usuário que esteja pesquisando x receba a antes de receber b , já que, para o assunto da busca, aquele documento deve ser (supostamente) *mais relevante* que este. No caso da busca y deve ocorrer precisamente o inverso.

O critério clássico para classificação e ordenação dos resultados das buscas parte, então, da premissa de que a incidência da expressão de busca num documento é indício do interesse do usuário. Assim, a taxa de incidência num documento e a exatidão da expressão são os critérios principais utilizados no ordenamento. Quanto mais ocorrências da expressão buscada (e quanto maior sua exatidão) forem encontradas num documento, maior será a probabilidade dele ser ordenado nas primeiras posições apresentadas na saída. Esse critério elementar é refinado continuamente pelos dispositivos dedicados à busca, incluindo elementos

“contextuais” e heurísticas complexas para a tomada de decisão classificatória; no entanto, a idéia geral é (quase) sempre esta. Duas são as hipóteses (relacionadas) utilizadas para dar conta da “intenção” do usuário:

- i Quando interrogamos um bibliotecário (o equivalente pré-ciberespaço do dispositivo de busca) a propósito de um livro, nós geralmente nos referimos ao “assunto” principal abordado no texto;
- ii Há boa probabilidade de que o tema de um documento se exprima através da incidência do mesmo vocabulário no corpo do texto.

Por exemplo, este capítulo é dedicado aos dispositivos de busca; até este ponto, a palavra-tema se repetiu quinze vezes, enquanto que os temas relacionados (portais de interesse geral, ciberespaço e Internet) foram repetidos cinco, seis e dez vezes, respectivamente. Naturalmente, a hipótese (ii) nem sempre é boa. Por exemplo, o tema de um livro que contenha somente receitas de cozinha pode ser expresso por vocábulos como “culinária”, “gastronomia”, “receitas”, etc. e, ainda assim, não conter em todo o documento uma única ocorrência dessas expressões! Por conta disso, a medida que ganham experiência, os próprios usuários percebem que os dispositivos de busca respondem segundo a hipótese (ii) e se adaptam isso. Eles passam também a perguntar não por documentos relacionados ao “tema”, mas por documentos que *contenham* uma determinada expressão de busca.

Um critério simples de classificar a relevância, elaborado com base nessa hipótese interpretativa, poderia ser, por exemplo, determinar a razão entre a expressão tema e a quantidade total de palavras no documento. No nosso caso (até o parágrafo anterior contamos 2.348 palavras), a razão da relevância dos temas catalogados – dispositivos de busca, portais de interesse geral, ciberespaço e Internet – seria, respectivamente, 0.00638, 0.00212, 0.00255 e 0.00425, o que parece consistente com a estrutura temática deste fragmento de documento. Poderíamos, por exemplo, supor que existe boa probabilidade do tema do documento ser expresso na primeira linha (o título) e ponderar o cálculo de acordo; se arbitrarmos peso três para essa condição, o valor calculado de relevância naquela seção do documento da expressão “dispositivos de busca” saltaria para 0.01916. Esse critério pode ser refinado continuamente através de novas hipóteses sobre como os produtores de texto exprimem do ponto de vista vocabular o assunto dos seus documentos, agregando características típicas da Internet. Por exemplo, é altamente provável que um *site* cujo nome de domínio seja <http://www.searchengine-watch.com> se refira a dispositivos de busca! Isso pode se refletir como valor mais alto na ponderação das nossas entradas de índice.

Esse critério tem um problema: ele *não* reage como um bibliotecário humano culto e bem informado. Suponhamos que estamos utilizando um dispositivo de busca construído com base no critério descrito no parágrafo anterior para o cálculo da relevância dos documentos. Para uma busca pela expressão “ciberespaço” ele devolveria o presente texto antes do artigo *A emergência do ciberespaço e as mutações culturais* (ver <http://empresa.portoweb.com.br/pierrelevy/aemergen.html>), de Pierre Lévy, cujo cálculo de relevância para essa expressão de busca teria sido de 0.00110. É de se presumir que um bibliotecário humano, culto e bem informado, não respondesse desse modo, mas devolvesse o texto de Lévy em primeiro lugar. Afinal, este bibliotecário hipotético provavelmente imaginaria que estamos em busca de idéias de pesquisadores importantes e seu conhecimento de mundo o forçaria a se perguntar: mas afinal, quem é mais importante, um obscuro pesquisador júnior de uma universidade brasileira ou o filósofo francês Pierre Lévy?

Foi raciocinando deste modo que alguns pesquisadores de Stanford (Larry Page, Serguei Brin, Rajeev Motwani e Terry Winograd – os dois primeiros os fundadores do Google) elaboraram há alguns anos um critério distinto, batizado como *PageRank*. A idéia central daqueles pesquisadores é buscar um meio de determinar a “importância” de um documento na *web* consistente com o raciocínio daquele bibliotecário hipotético. Para isso, eles partiram da premissa de que um dos critérios acadêmicos de avaliação dos pesquisadores – a contagem de citações recebidas – seria aplicável para o cálculo da “importância” de um documento na *web*, tirando proveito da própria estrutura do hipertexto, baseada em *links* entre páginas. Para eles, “páginas [*web*] para as quais muitos *links* apontam são, em geral, mais ‘importantes’ que páginas com poucos *links*” apontando para elas – do mesmo modo que “a simples contagem de citações tem sido utilizada para especular sobre os futuros ganhadores do Prêmio Nobel” (PAGE, 1998: 3). Pelo critério, a importância de uma página se propaga do mesmo modo que uma citação elogiosa a outro feita por um pesquisador muito importante. À época, esses pesquisadores determinaram, por exemplo, que a página principal do Yahoo era (então) o documento para o qual mais *links* apontavam. Pelo critério adotado, eles admitem então que “se uma página tem um *link* saindo da *home page* do Yahoo, ele pode ser apenas um *link*, mas trata-se de um muito importante”, já que “esta página deveria receber um *ranking* mais alto que muitas páginas com mais *links* [apontando para elas], mas oriundos de lugares mais obscuros” (PAGE, 1998: 3).

Algumas implicações desse critério de relevância serão discutidas mais adiante. Por ora, basta assinalar que, embora não seja dito no artigo citado, parte-se da premissa de que o usuário está interessado primariamente nos *sites* mais populares que atendam ao critério de

busca. Neste caso, as saídas passam a ser classificadas e ordenadas segundo a taxa de referência ao documento feita por outros documentos. Quanto mais *links* forem encontrados em *sites* diferentes apontando para um determinado documento, maior a probabilidade deste último aparecer nas primeiras saídas do resultado da busca. Trata-se aqui de determinar a audiência de um *site* a partir das referências feitas a eles por outros *sites* e devolver ao usuário aqueles documentos de maior audiência que atendam à expressão de busca utilizada.

3.2.2. Primeiro experimento: o diferencial das hipóteses interpretativas

Pelo rápido sumário de cada um dos critérios, é fácil concluir que eles devem retornar resultados de busca substancialmente distintos, já que correspondem a hipóteses interpretativas a propósito do interesse dos usuários substancialmente distintas. Isso pode ser facilmente corroborado por um experimento bastante simples. O modo mais simples – investigar o modo como os vários dispositivos de busca foram projetados – está inacessível por força dos segredos industriais. Em vista disso, precisaremos adotar técnicas de investigação de caixas pretas⁷.

Em primeiro lugar, espera-se que as buscas retornem quantidades semelhantes de documentos em dispositivos de poder computacional equivalente simplesmente porque os acervos tendem a ser os mesmos ao longo do tempo, desde que os processos de catalogação (indexação) e cálculo da relevância não estejam relacionados entre si. Assim, espera-se que a quantidade de documentos retornados pelas as buscas realizadas num dado dispositivo seja equivalente à quantidade retornada pela mesma busca em outro. É razoável supor que os dispositivos dedicados tenham acervo maior que aqueles incorporados a *sites* com outros propósitos comerciais, já que aqueles precisam dedicar maior poder computacional à tarefa. Isso deve implicar em quantidades diferentes de documentos retornados em cada caso.

⁷ Uma caixa preta é simplesmente um dispositivo cujo mecanismo de funcionamento está inacessível. O observador tem acesso somente às entradas e saídas da caixa, sendo capaz de manipular umas e observar outras de modo a deduzir do comportamento controlado da máquina o seu funcionamento interno.

Se isso se revelar verdadeiro para qualquer busca realizada, então será possível demonstrar que o critério de classificação e ordenação, projetado com base numa determinada concepção da audiência dos *sites*, é a única função capaz de afetar o resultado das buscas. Desse modo, para qualquer busca realizada, é esperado que dispositivos projetados com concepções de audiência equivalentes devolvam seu resultado na mesma ordem de apresentação. E como o acervo à disposição dos dispositivos deve ser o mesmo em todos os casos, existe probabilidade significativa de incidência das mesmas saídas em todas as buscas. Em outras palavras, as buscas nesses dispositivos terão alto grau de *redundância*. Conseqüentemente, as buscas realizadas em dispositivos com outra concepção produzirão resultados significativamente distintos daqueles. Como sabemos que o Google é o único dispositivo que implementa um critério cuja concepção de audiência é substancialmente distinta dos dispositivos clássicos, é esperado que as buscas realizadas naquele *site* produzam resultados diferentes dos demais.

Para o experimento foram selecionados todos os dispositivos de busca dedicados com presença em quaisquer das pesquisas de audiência da Nielsen//Net-Ratings referentes a 10 países cujas buscas possam se restringir a *sites* brasileiros. Com base nesse critério, foram selecionados os dispositivos Altavista, Google e Yahoo. A seguir selecionamos os dispositivos de busca incorporados aos grandes portais de propósito geral orientados ao público brasileiro; neste caso, foram considerados apenas os portais com dispositivos próprios à época do experimento, excluindo-se aqueles que se utilizam dos dispositivos dedicados. Com isso, selecionamos os dispositivos da MSN, do portal UOL e, finalmente, do iG.

Foram realizados apenas dois experimentos, já que se buscava somente uma validação inicial das hipóteses. No primeiro deles, selecionamos uma expressão de busca possivelmente popular. Para isso, recorreremos à lista de expressões mais utilizadas no buscador do Terra entre 8 de Fevereiro a 2 de Outubro de 2004. Como qualquer expressão de busca deve retornar os mesmos resultados relativos, a seleção foi arbitrária do ponto de vista do experimento.

Supomos também que a maioria dos usuários – mesmo quando não encontrar resultados coerentes com o objeto do seu interesse – se satisfaz com as primeiras páginas do produto da busca, simplesmente porque a quantidade de saídas é muito grande, em conformidade com a pesquisa iProspect já citada. Assim, analisamos somente as vinte primeiras saídas de cada busca. As entradas selecionadas foram catalogadas por domínio (nome de servidor) e não por URL; documentos oriundos do mesmo domínio foram considerados uma única saída, na suposição de que estes documentos são apenas locuções diferentes da mesma voz discursiva e não caracterizam resultados efetivamente distintos para o usuário da busca.

Num primeiro momento, simplesmente inventariamos a quantidade total de documentos devolvidos nas duas buscas, cujo resultado é ilustrado na Tabela 22. Os resultados são razoavelmente discrepantes em relação às expectativas. Assim, não podemos considerar que o experimento corrobora a hipótese acerca da quantidade total dos documentos obtidos. Artigo publicado no *site* Search Engine Watch, já citado (SULLIVAN, 2004), dá conta que o Google tem indexados aproximadamente o dobro de documentos indexados pelo Yahoo e o MSN. Assim, seria esperado que na busca pela expressão “dia do índio” aquele dispositivo obtivesse mais documentos que estes, o que de fato não ocorreu. Isto pode significar tanto que os dispositivos endereçam documentos diferentes, isto é tenham acervos diferentes, quanto que o critério de relevância empregado afeta *também* o universo de documentos obtidos. Em qualquer caso, devemos interpretar que o conjunto D de documentos obtidos em resposta à busca seja diferente em cada caso, cabendo-nos explicar a discrepância encontrada. A primeira hipótese poderá ser verificada na próxima avaliação; se ela for verdadeira, espera-se que os resultados das buscas efetuadas no MSN sejam distintos também das efetuadas no Yahoo e não somente das efetuadas no Google. Se isto não ocorrer, é provável que a segunda hipótese seja a melhor.

Expressão	Execução	Dispositivos dedicados			Portais de conteúdo		
		Altavista	Yahoo	Google	iG	UOL	MSN
Dia do índio	12/05/2005	59.500	62.700	24.900	4.360	1.209	4.308
Sincretismo	06/06/2005	93.100	26.600	97.000	1.175	4.181	11.704

Tabela 22: Quantidade total de documentos devolvidos em cada experimento.

Na segunda parte do experimento, passamos a contar a incidência das mesmas saídas em dispositivos diferentes. Foram consideradas apenas as saídas que apareceram em buscas em mais de um dispositivo; todas as demais foram desprezadas como acidentais. As saídas reincidentes numa mesma busca foram igualmente ignoradas. Calculamos, então, a percentagem em face do total das saídas incidentes por dispositivo. O resultado, exibido na Tabela 23, exprime a probabilidade de ocorrência de uma mesma saída em dispositivos distintos dada uma mesma expressão de busca. Este cálculo pode ser considerado a medida da *redundância* das saídas dos dispositivos.

URL	Altavista	Yahoo	Google	iG	UOL	MSN
Dia do índio	25.35	25.35	5.63	7.04	12.67	23.94
Sincretismo	25.39	23.80	4.76	11.11	14.28	20.63

Tabela 23: Taxa de incidência de saídas diferentes por dispositivo.

As pesquisas realizadas no Altavista e no Yahoo conduzem a produtos idênticos para todos os efeitos práticos, mas não as realizadas no Google. Este dispositivo produz, invariavelmente, saídas inteiramente distintas dos demais. Isso conduz à conclusão de que o critério de ordenação utilizado é substancialmente distinto de todos os demais dispositivos e corrobora (pelo menos preliminarmente) a hipótese de que a concepção de audiência implícita nessas heurísticas afeta significativamente o produto das buscas. No entanto, a equivalência das saídas do dispositivo do portal MSN às dos dispositivos do Altavista e Yahoo é discrepante em face das suposições iniciais: ela indica que a suposição de que suas consultas são feitas contra acervos menores é provavelmente falsa. É possível que essa equivalência indique, ao contrário, que o acervo deve ser o mesmo nos três casos. Aliás, isso é consistente com as observações, em artigo citado (SULLIVAN, 2004), o *site* Search Engine Watch alerta que o MSN e o Yahoo proclamam acervos equivalentes, com 4.2 e 5 bilhões de documentos indexados, respectivamente. O fato é que as duas estratégias de ordenação apresentadas não são os únicos fatores que intervêm nas buscas. Esses são apenas os fatores mais salientes. Além disso, é provável que os dispositivos não apliquem *primeiro* o critério de relevância para *depois* ordenar o resultado. Como o acervo é sempre muito grande, as estratégias de busca devem ser otimizadas quanto aos índices consultados.

Independente das reservas que possamos ter quanto aos resultados dos dois experimentos, eles não podem ser considerados conclusivos. Como se tratava apenas de validar as hipóteses formuladas, isso não pode ainda ser considerado um problema. No entanto, tais resultados deverão ser obrigatoriamente confrontados contra novos experimentos. Naturalmente, os resultados de apenas dois experimentos devem ser vistos com cautela, não podendo ser considerados conclusivos tanto para corroborar quanto para refutar as hipóteses levantadas, requerendo novos experimentos. No entanto, eles indicam que, pelo menos, as hipóteses merecem ser investigadas. Além disso, fornecem pistas que nos devem permitir inserir determinadas interações com dispositivos computacionais entre as práticas discursivas emergentes no ciberespaço, com lugar assegurado entre as práticas decorrentes de interações estritamente humanas.

3.2.3. Segundo experimento: em busca da palavra do Mesmo

É preciso afirmar inequivocamente que tudo indica que o critério de relevância do Google é o melhor disponível, não tanto pelas suas qualidades intrínsecas, mas pela palavra final dos usuários das buscas: a cada duas buscas realizadas no mundo, uma é feita neste dispositivo, a outra ficando por conta de pelo menos treze dispositivos, para contarmos somente os mais populares. Mesmo sem conhecer a natureza dessa diferença, os usuários devem ter escolhido o Google por força da consistência do produto das buscas realizadas no dispositivo com seus próprios interesses. Note-se que não é possível atribuir essa predominância simplesmente a um maciço investimento de capital: o Google nasceu pequeno, como “mera” pesquisa acadêmica – e só recebeu investimentos de capital *depois* de se tornar um sucesso de público. Apesar dessa maciça adesão, precisamos considerar as características intrínsecas do seu critério de relevância e indagar o que resulta dele: quais suas implicações para a *prática social* em torno de um dos eventos comunicativos mais comuns no ciberespaço e para os hábitos de navegação dos usuários da Internet. A nosso ver, as premissas assumidas por Page e colegas implicam em alguns problemas importantes.

Em primeiro lugar (e esse ponto é essencial para o restante do raciocínio), é preciso considerar que o cálculo da relevância relativa dos vários documentos que atendam a uma expressão de busca, ainda que imaginado somente como critério de *ordenação* dos resultados, é, na prática, um critério de *seleção* de que subconjunto daquele conjunto D de documentos obtidos será, de fato, exibido ao usuário. Já dissemos que o produto da maioria das buscas, por envolver geralmente milhares de documentos, é impossível de ser processado pelos seres humanos. E os projetistas dos dispositivos sabem disso! Por exemplo, a busca pela expressão “theory of games” no Google resultou, em 29 de outubro de 2006, em 404 mil documentos; no entanto, apenas 710 foram exibidos pelo dispositivo. A mesma expressão no Yahoo, resultou em 120 mil documentos encontrados, mas apenas mil exibidos! Isso é consistente com o comportamento de fato dos usuários. Como já assinalamos, apenas 12% dos usuários insistem na busca além da terceira página de resultados (IPROSPECT, 2006: 5). E mesmo que essa minoria de usuários pacientes insista na sua pesquisa, os dispositivos não lhe permitirão seguir além de um certo ponto arbitrário. Portanto, o cálculo da relevância relativa (não importando segundo qual hipótese foi elaborado) é utilizado de fato para a *seleção* do que será exibido ao usuário e não meramente para sua ordenação.

Ora, como vimos, a hipótese do Google é baseada no critério acadêmico de contagem de citações recebidas para julgamento da importância de um pesquisador. Admitamos, por hipótese, a validade desse critério no domínio para o qual elaborado. Trata-se de um julgamento pelos próprios pares do pesquisador. Embora esse julgamento implícito nem sempre possa ser “justo”, ao longo do tempo as melhores pesquisas acabam por ser reconhecidas pelo menos por parte da comunidade acadêmica. Pensemos no caso de Galileu, quase condenado à fogueira por seus pares, um destino que – pelo menos metaforicamente – ameaça todas as idéias revolucionárias: se hoje fizer minhas as críticas dos seus detratores, provavelmente serei eu o condenado à fogueira... Ocorre que no domínio acadêmico, a contagem de citações é apenas um critério de “ordenamento”, podendo servir para determinar o próximo ganhador do Prêmio Nobel, mas não sendo utilizada para decidir que pesquisas serão ou não submetidas à comunidade acadêmica. Esta dispõe de inúmeros mecanismos para assegurar a circulação de novas idéias, mesmo que produzidas por pesquisadores obscuros, oriundos de instituições igualmente obscuras, em qualquer parte do mundo. Além disso, essa comunidade é suficientemente pequena para que tais mecanismos sejam viáveis: sempre posso saber com pouco esforço o que outros pesquisadores em qualquer parte do mundo no meu domínio de interesses estão publicando, já que são relativamente poucas e facilmente acessíveis as fontes de informação a consultar.

Isso não ocorre na Internet, quando a consideramos isoladamente. Aquilo que a torna avessa aos controles e à censura – o fato de ser uma rede sem centro virtualmente ilimitada – é também aquilo que restringe o *acesso* à informação publicada: *tudo* pode estar *disponível*, mas *pouco* está efetivamente *acessível*. A circulação de idéias em grande escala na Internet depende consideravelmente dos dispositivos de busca – e se estes *selecionam* mais que ordenam o que estará acessível aos usuários, funcionam na prática como mecanismos de *bloqueio* à circulação de algumas informações, *restringindo* o que os usuários podem conhecer. Por si só (e independente de qualquer ação institucional), os dispositivos controlam, *censuram* a circulação da informação na Internet – e *o que* eles censuram está, como vimos, diretamente relacionado às hipóteses interpretativas adotadas para elaborar o critério de cálculo da relevância relativa dos documentos buscados.

Como vimos, o critério do Google é baseado na *audiência* dos *sites*: ele valoriza as páginas na proporção direta da quantidade de *links* que apontam para ela, propagando esse valor para as páginas “citadas”. Isto significa que uma busca no Google retorna *somente* os *sites* mais populares para uma dada expressão de busca. Ora, idéias novas e revolucionárias não podem ser populares simplesmente porque ainda são novas e revolucionárias... Essas são

as idéias *censuradas* pelo Google! Acrescente-se a isso aquelas idéias não tão novas e reacionárias, mas impopulares (admitamos que elas existam), sobre um determinado assunto e concluiremos que o Google é avesso aos direitos de expressão das minorias. A “missão” do Google – “organizar a informação mundial e torná-la universalmente acessível e útil” (GOOGLE, 2006) – precisa ser vista sob esta ótica menos otimista. Nosso argumento é que a primeira parte da missão é verdadeira, já que “organizar” implica em *controlar*. No entanto, há que se questionar a autenticidade do restante da proposição, em particular seu advérbio...

É fácil imaginar um experimento visando determinar se estamos ou não sofrendo de uma Síndrome de Cassandra ideológica. Não é necessário imaginarmos idéias novas e revolucionárias ou simplesmente impopulares, pesquisando-as no Google. Aquela empresa disponibiliza uma versão *desktop* do seu dispositivo, o Google Desktop (ver <http://desktop.google.com>), destinada a uso em computadores pessoais, para indexação e recuperação de arquivos ali armazenados. Embora inúmeras características de implementação sejam, obviamente, diferentes do dispositivo Internet, o que nos interessa – o critério de relevância – é igualmente implementado nesta versão. Isso nos permite realizar experimentos laboratoriais inteiramente controlados, experimentos impraticáveis nas condições reais de uso da Internet, em especial numa base de oito bilhões de documentos.

Suponhamos que estamos na metade do Século XVI e que um certo Nicolau Copérnico, obscuro astrônomo polonês, acaba de publicar um curioso estudo intitulado *De revolutionibus orbium coelestium* no hipotético *website* <http://www.frombork.edu> na não menos hipotética Internet da Renascença. Imaginemos também que isso é tudo o que, naquele momento, os astrônomos têm à disposição para divulgar suas idéias. Como se trata de um novo estudo de um pesquisador obscuro de uma universidade obscura com idéias excessivamente diferentes daquelas admitidas não apenas pelos seus pares, mas por todos os usuários da Internet Renascentista, nenhum outro *website* se deu ao trabalho de indicá-lo com um *link* de qualquer tipo. Estamos supondo também, para simplificar o problema, que o *website* <http://www.malleusmaleficarum.org>, àquele tempo, não publicasse listas exaustivas de candidatos à fogueira, mas apenas daqueles efetivamente incinerados para edificação dos fiéis; não fosse assim e Copérnico teria sido candidato ao prêmio de popularidade da Inquisição. Como não existe nenhum *link* apontando para ele, o dispositivo de busca <http://www.googlorum.com> classifica o documento com a menor relevância possível. O único modo de alguém interessado nos movimentos celestes localizar esse documento no Googlorum é digitar uma combinação de palavras que só exista naquele texto e em nenhum outro com relevância maior, o que não é plausível, a menos que o usuário tenha dons de

clarividência, caso em que não teria necessidade de utilizar o dispositivo de busca. Num conjunto universo de pelo menos oito bilhões de documentos e milhões de *websites*⁸, a probabilidade de um usuário, navegando ao acaso, localizar um documento específico sem a utilização de um dispositivo de busca é, para todos os efeitos práticos, desprezível. Nesse cenário, o documento de Copérnico está irremediavelmente inacessível – e nós condenados a continuar acreditando ser a Terra o centro do universo, circundada por uma esfera de estrelas fixas.

Com o Google Desktop, a plausibilidade desse cenário é facilmente verificável. Para isso, construímos uma rede de *websites* com hiperdocumentos no formato HTML armazenados em disco, com as seguintes características primárias:

- i Cinco deles, denominados *fanzine_?.htm*, são documentos extraídos de um *site* dedicado a um jogo para computadores bastante popular que utiliza personagens históricos, dentre eles um certo Nicolau Copérnico.
- ii Três deles, denominados *note_?.htm*, são notas biográficas sobre o astrônomo Nicolau Copérnico, extraídas da Wikipedia.
- iii Um deles, cujo arquivo recebeu o nome de *wikipedia_whois.htm*, contém um outro fragmento da mesma nota biográfica encontrada na Wikipedia.
- iv Um deles, de nome *wikipedia_ideas.htm*, é um verbete sobre a obra *De revolutionibus*, de Copérnico, também extraído da Wikipédia.
- v O último é um fragmento do texto original de Copérnico, encontrado na Internet, que recebeu o nome de arquivo de *_revolutionibus.htm*.

Os títulos dos documentos foram escolhidos ao acaso e visam tão somente melhor visualização dos resultados das buscas. Além disso, todos os documentos contêm a expressão a ser utilizada na busca, “Nicolaus Copernicus”. É de se notar também que estes documentos foram confeccionados em computador diferente daquele que seria utilizado no experimento, visando eliminar quaisquer “contaminações” do ambiente onde seria realizado o experimento. Finalmente, construímos uma rede de *links* ligando os vários hiperdocumentos, construindo

⁸ Em janeiro de 2006, haviam registrados somente no Brasil cerca de um milhão de domínios, conforme o CGI.BR (2006).

um hipertexto e deixando o documento com o texto original de Copérnico fora dele, conforme *graphos* exibido na Ilustração 21.

O cálculo da relevância definido para o Google indica que o documento `wikipedia_whois.htm` deveria receber o mais alto valor de relevância relativa, simplesmente porque é o mais citado; já o documento `wikipedia_ideas.htm`, deveria receber um valor intermediário porque, apesar de receber uma só citação, ela é feita pelo documento de maior relevância. Todos os demais documentos devem receber valor de relevância relativa muito próximos entre si, senão iguais.

O que se espera de um experimento que tente simular o pesquisador júnior Galileo

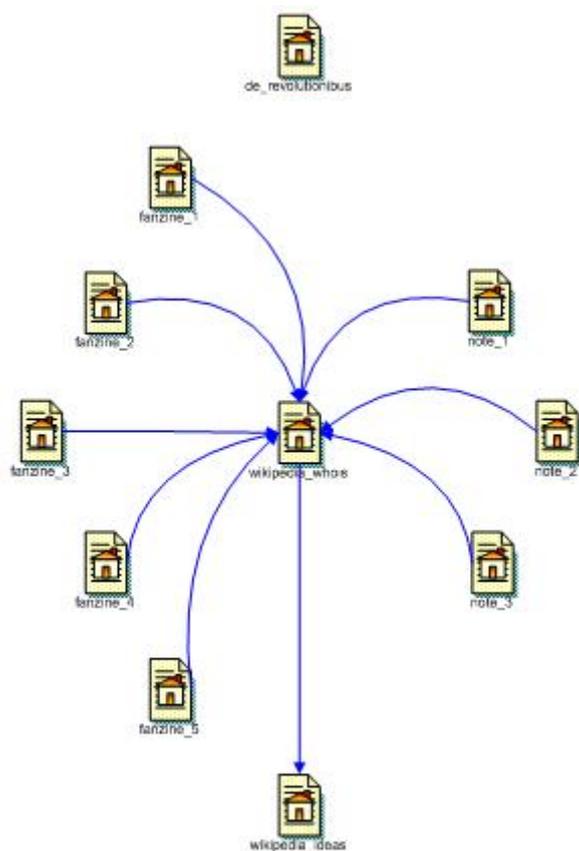


Ilustração 21: A estrutura de *links* dos hiperdocumentos do experimento Copérnico.

Galilei em sua busca intelectual no ramo da astronomia utilizando nosso hipotético dispositivo `http://www.googlorum.com` é que, para qualquer histórico de buscas com uma mesma expressão, os documentos `wikipedia_whois.htm` e `wikipedia_ideas.htm` são sempre devolvidos primeiro e sempre nesta ordem. Como o documento `de_revolutionibus.htm` recebeu o menor valor possível para a relevância e o conjunto universo de documentos obtidos numa busca qualquer é muito grande (uma busca no Google pela expressão “Nicolaus Copernicus” retornou, em 29 de outubro de 2006, 437 mil documentos), ele *nunca* será retornado em qualquer histórico de buscas.

Em vista do ambiente estar (quase que) inteiramente sob controle do observador, o protocolo da experiência pode ser bastante simples e se inicia após a instalação do Google Desktop e a completa indexação dos documentos armazenados no computador de simulação da Internet. Antes de qualquer outro passo precisamos executar a busca (pela expressão “Nicolaus Copernicus”), visando determinar o estado corrente do computador onde será realizado o experimento. Todos os documentos obtidos nesta busca serão, naturalmente, retornados nas buscas subseqüentes e deverão ser removidos dos resultados, já que não afetam o experimento. No nosso experimento, foram encontrados oito documentos nessa busca, que

foram ignorados em todas as avaliações posteriores. Porém, é de se notar que, em todas as buscas realizadas, tais documentos foram sempre devolvidos ao final da lista de resultados, o que corrobora nossa decisão de ignorá-los na avaliação do experimento. Somente após este passo, os documentos do experimento foram instalados, cada um deles num diretório de arquivos diferentes, visando melhor assegurar a simulação da Internet. Após esse passo, aguardamos tempo suficiente para que os arquivos instalados fossem corretamente indexados pelo Google. Após esses procedimentos de controle, executamos a busca alvo do experimento, ilustrada na Tabela 24.

	Título da página	Localização do documento
1	Nicolaus Copernicus	C:\Temp\web_11\wikipedia_whois.htm
2	De revolutionibus orbium coelestium	C:\Temp\web_10\wikipedia_ideas.htm
3	About Copernicus	C:\Temp\web_09\note_3.htm
4	About Nicolaus Copernicus	C:\Temp\web_08\note_2.htm
5	Nicolaus Copernicus	C:\Temp\web_07\note_1.htm
6	De Revolutionibus	C:\Temp\web_01\de_revolutionibus.htm
7	Game Spot	C:\Temp\web_06\fanzone_5.htm
8	Game Center	C:\Temp\web_05\fanzone_4.htm
9	Civilization The Game	C:\Temp\web_04\fanzone_3.htm
10	Civilization IV	C:\Temp\web_03\fanzone_2.htm
11	Civilization IV Fanatics' Center	C:\Temp\web_02\fanzone_1.htm

Tabela 24: Resultado da primeira busca do experimento Copérnico.

O resultado está rigorosamente dentro do ordenamento esperado. A posição do documento de *de_revolutionibus.htm* após os vários documentos *note_?.htm* é irrelevante, já que podemos atribuir esse fato a fatores acidentais. Afinal, as saídas de mesmo valor de relevância relativa precisam, ainda assim, ser ordenadas de alguma forma. Uma outra explicação plausível é a utilização subsidiária dos critérios clássicos de relevância. Isso, porém, não afeta as conclusões que queremos ter para o experimento: num universo de 437 mil documentos, o texto original de Copérnico estaria inacessível.

O passo seguinte do experimento foi abrir diretamente o arquivo *fanzone_1.htm*, o de menor relevância na busca. Isso foi feito cinco vezes e visava determinar o efeito dos *hits* realizados diretamente num documento independente dos resultados das buscas. Com isso, pretendíamos determinar se a *audiência* independente de um *website* afetaria a relevância de um documento, caso pudesse ser captada pelo dispositivo de busca. O resultado é visto na Tabela 25. Note-se que a relevância do documento aumentou.

Este resultado deve ser visto com cautela e considerado válido tão somente para a versão *desktop* do dispositivo. Isto porque para obter o mesmo efeito na Internet, um dispositivo de busca deveria ser capaz de analisar tráfego de rede independente dos documentos indexados. Isto pode ocorrer (o Google nada nos informa), mas não nos parece plausível, em vista do gigantismo da tarefa. O que queremos ressaltar com esse passo do experimento é sua consistência com o critério primário de cálculo da relevância relativa: o rato do nosso laboratório se comporta sempre como esperado... Mais uma observação precisa, porém, ser feita. Note-se a sexta saída da busca. Trata-se do documento que criamos para registrar cada um dos resultados do experimento que, obviamente, foi indexado pelo Google e, como continha inúmeras referências a Nicolau Copérnico, foi devolvido nas buscas. É um caso de Princípio de Incerteza de Heisenberg aplicado à nossa pesquisa: não podemos realizar qualquer experimento sem que os eventos que queremos medir sejam afetados de alguma forma pelos meios utilizados como medição no próprio experimento.

	Título da página	Localização do documento
1	Nicolaus Copernicus	C:\Temp\web_11\wikipedia_whois.htm
2	De revolutionibus orbium coelestium	C:\Temp\web_10\wikipedia_ideas.htm
3	About Copernicus	C:\Temp\web_09\note_3.htm
4	About Nicolaus Copernicus	C:\Temp\web_08\note_2.htm
5	Nicolaus Copernicus	C:\Temp\web_07\note_1.htm
6	results.doc	E:\results.doc
7	De Revolutionibus	C:\Temp\web_01\de_revolutionibus.htm
8	Civilization IV	C:\Temp\web_03\fanzone_2.htm
9	Civilization IV Fanatics' Center	C:\Temp\web_02\fanzone_1.htm
10	Game Spot	C:\Temp\web_06\fanzone_5.htm
11	Game Center	C:\Temp\web_05\fanzone_4.htm
12	Civilization The Game	C:\Temp\web_04\fanzone_3.htm

Tabela 25: Resultado de uma busca captando audiência independente.

O próximo passo do protocolo foi navegar, a partir da página de resultado da busca anterior, para o documento *note_1.htm*, o de menor relevância no grupo, visando simular os *hits* num documento qualquer a partir das buscas realizadas no dispositivo. Esse passo, em conjunto com o seguinte, descrito no próximo parágrafo, visa determinar se buscas bem sucedidas no dispositivo afetam o resultado das buscas subseqüentes, isto é, se o dispositivo ignora ou não suas próprias páginas, realimentando o processo. A partir da mesma página de busca, efetuamos cinco *hits* no documento alvo. A seguir, efetuamos nova busca, que obteve rigorosamente o mesmo resultado exibido na Tabela 25, referente à busca anterior.

A seguir, executamos cinco vezes a mesma busca, continuando o experimento anterior, para avaliar mais diretamente a influência do próprio processo de busca no cálculo da relevância. Ora, o produto das buscas é armazenado em *caches* temporários, visando acelerar o processo de devolução dos documentos na Internet. Esses *caches* são compostos por hiperdocumentos válidos e são considerados pelo mecanismo de indexação, em especial porque contêm *links* – e *links* são a pedra de toque do dispositivo. É esperado que isso afete o experimento, como de fato ocorreu e pode ser comprovado pelos dados listados na Tabela 26. Como se pode notar, o documento *note_1.htm* saltou para a primeira posição da relevância (se ignorarmos o fruto espúrio do nosso próprio experimento), num processo de realimentação positiva. Novamente, isso é coerente com a filosofia de projeto do dispositivo: documentos muito encontrados em buscas pela comunidade de usuários constituem um índice de audiência que o buscador pode captar facilmente. Isso significa que o próprio dispositivo contribui para a popularidade dos *websites* mais populares, realimentando o processo.

	Título da página	Localização do documento
1	Results.doc	E:\results.doc
2	Nicolaus Copernicus	C:\Temp\web_07\note_1.htm
3	Nicolaus Copernicus	C:\Temp\web_11\wikipedia_whois.htm
4	De revolutionibus orbium coelestium	C:\Temp\web_10\wikipedia_ideas.htm
5	About Copernicus	C:\Temp\web_09\note_3.htm
6	About Nicolaus Copernicus	C:\Temp\web_08\note_2.htm
7	De Revolutionibus	C:\Temp\web_01\de_revolutionibus.htm
8	Civilization IV	C:\Temp\web_03\fanzone_2.htm
9	Civilization IV Fanatics' Center	C:\Temp\web_02\fanzone_1.htm
10	Game Spot	C:\Temp\web_06\fanzone_5.htm
11	Game Center	C:\Temp\web_05\fanzone_4.htm
12	Civilization The Game	C:\Temp\web_04\fanzone_3.htm

Tabela 26: Resultado de busca com realimentação positiva.

É preciso assinalar que os resultados rigorosamente dentro do previsto neste experimento não significam que acrescentamos algum conhecimento ao problema dos dispositivos de busca: eles decorrem claramente da concepção do Google. O experimento Copérnico apenas descreve um cenário plausível para uma das implicações da hipótese interpretativa sobre os interesses dos usuários no diálogo de busca. São essas implicações que importam. Se as premissas com as quais trabalhamos são corretas, a censura exercida pelo dispositivo de busca mais utilizado no mundo implica em dirigir a produção do discurso para um coro de múltiplas vozes que, no entanto, cantam em uníssono, sem tons discordantes, sem

contrapontos nem vozes dissonantes. O que a concepção do Google acaba por produzir é o discurso do Mesmo no discurso do Outro – as idéias da multidão, nunca das minorias, repetidas por todos.

Esta é, certamente, uma situação limite, até porque a Internet fornece um conjunto de outros mecanismos de acesso à informação que permitem a circulação de idéias novas, embora não na mesma escala e com a mesma abrangência dos dispositivos de busca. No entanto, o que queremos assinalar é essa possibilidade – assustadora – implícita naquela ferramenta que muitos parecem considerar o oráculo dos novos tempos. Se assim for, a sacerdotisa só interpreta uma única profecia – a das vozes em uníssono. Infelizmente, nada temos a contrapor a esta concepção para o cálculo da relevância, em particular quando todos parecem satisfeitos com ela...

3.2.4. Terceiro experimento: o super-controle da palavra do outro nos portais de interesse geral

Resta uma indagação importante no presente contexto: os resultados das buscas são afetados pela presença dos portais? Em outras palavras, os dispositivos de busca incorporados aos portais são utilizados como atratores retóricos? A expectativa é de que a resposta seja afirmativa: trata-se de um uso evidente de algumas características desses dispositivos. Além disso, já vimos a importância das primeiras páginas dos resultados da busca para o comportamento dos usuários. O relatório da pesquisa já citada assinala que “é vital para os profissionais de *marketing* assegurar que seus *websites* sejam encontrados na primeira página do resultado das buscas, ou pelo menos nas três primeiras páginas desse resultado” (IPROSPECT, 2006: 12). Essa pesquisa incluiu também a seguinte questão: “encontrar uma empresa listada nos primeiros resultados de uma busca me leva a concluir que a empresa é líder do seu mercado”. Chamados a concordar ou não com esta afirmação, apenas 25% dos consultados discordaram! Portanto, a primeira página do resultado de uma busca é um local privilegiado para os empreendimentos na Internet. Além disso, assumimos que os portais têm como estratégia primária manter os usuários nos seus domínios. Ora, as buscas são eventos que potencialmente remetem os usuários para fora dos domínios do dispositivo (e dos

portais). Portanto, é possível que os portais incluam estratégias para assegurar a alta relevância de páginas de seus domínios nos dispositivos cujos serviços fornecem.

Se as buscas são afetadas pela proximidade dos portais, é de se esperar que uma busca realizada num dispositivo incorporado a um portal resulte em saídas diferentes da mesma busca quando realizada no *site* original do dispositivo. E se este recurso é utilizado como atrator retórico, é de se esperar que essa diferença aponte para os domínios do portal, isto é, espera-se que as saídas diferentes daquelas encontradas no dispositivo original correspondam a *links* para documentos armazenados no próprio portal e não fora dele. Essas hipóteses podem ser facilmente testadas já que a maioria dos principais portais (aqueles considerados como mais prototípicos da categoria no capítulo anterior) não utiliza dispositivos próprios. A exceção é o MSN Brasil, que, portanto, não poderá ser avaliado. Para esse experimento utilizaremos somente os cinco portais mais prototípicos segundo a avaliação realizada na seção 3.1.6. *Segunda avaliação – a prototipicidade da amostra* e ilustrada na Tabela 20. Esses portais utilizam dispositivos dedicados – o Yahoo e o Google – para as buscas dos seus usuários. Assim, o experimento será realizado no *corpus* exibido na Tabela 27, onde são indicados os dispositivos de referência.

Portal	URL do dispositivo	Buscador
UOL	http://busca.uol.com.br	Yahoo
Terra	http://buscador.terra.com.br	Google
Globo	http://zoom.globo.com	Yahoo
iG	http://farejador.ig.com.br	Yahoo
Pop	http://www.pop.com.br/pesquisa	Google

Tabela 27: *Corpus* do experimento com os dispositivos de busca dos portais.

O *corpus* será dividido em dois grupos: o dos portais que utilizam o Yahoo (UOL, Globo e iG) e os que utilizam o Google (Terra e Pop). Como os dispositivos são diferentes, não poderemos comparar os resultados *entre* os grupos, apenas aqueles obtidos *dentro* do mesmo grupo. A hipótese do experimento consiste em esperar que, se um usuário solicitar uma expressão passível de ser encontrada num dos documentos mantidos por um dos canais do portal ou de seus parceiros de negócio, o dispositivo de busca deverá responder fornecendo privilegiadamente essas páginas, independente do que forneceria na ausência do portal. Desse modo, para cada um dos portais mais prototípicos, selecionamos arbitrariamente expressões que de algum modo pudessem estar associadas a temas de interesse de parcelas dos usuários, conforme Tabela 28.

Expressão de busca	Portal	URL
Modelos da temporada	UOL	http://fashionprofile.uol.com.br/noticias.php?id=10883
Go-go boy	UOL	http://gay.uol.com.br/
Monteiro Lobato	Globo	http://lobato.globo.com/
Celulite	iG	http://delas.ig.com.br/beleza/
Namorar alguém mais jovem	Terra	http://mulher.terra.com.br/
Petkurt	Terra	http://www.terra.com.br/jovem/

Tabela 28: Expressões de busca utilizadas no experimento.

Naturalmente, não estamos certos de que tais expressões possam ser encontradas também nos demais portais; isso, porém, não é indispensável para as características do experimento. O experimento consistiu no seguinte protocolo muito simples, aplicado em separado aos dois grupos de portais:

- i Submeter a expressão de busca ao dispositivo de referência (Yahoo ou Google); o resultado dessa busca é considerado o resultado de referência;
- ii Submeter a mesma expressão aos dispositivos de cada um dos portais do grupo, resultado que será confrontado ao de referência;
- iii Repetir o passo (i) até que não existam mais expressões selecionadas.

Os resultados obtidos no passo (i) receberam um valor normalizado igual à unidade. Para os resultados obtidos no passo (ii), contamos a incidência das mesmas saídas obtidas no passo (i) e dividimos esse valor pela quantidade total de documentos retornados na busca de referência; o resultado desse cálculo representa a *redundância* da saída quando comparada à saída de referência. Como as buscas são realizadas no mesmo dispositivo e contra o mesmo acervo de documentos indexados, espera-se que o valor da redundância em todos os casos sejam sempre igual à unidade – caso a busca *não* seja afetada pela proximidade dos portais. O cálculo resultante é mostrado na Tabela 29, para os portais do grupo Yahoo, e na Tabela 30, para os portais do grupo Google.

Expressão	Yahoo	UOL	Globo	iG
Modelos da temporada	1,00	0,20	0,30	0,00
Go-go boy	1,00	0,70	0,50	0,00
Monteiro Lobato	1,00	1,00	0,70	1,00
Celulite	1,00	0,80	0,90	0,90

Tabela 29: Taxa de redundância das saídas dos portais em comparação com o Yahoo.

Os resultados são equívocos, como se pode notar. No grupo Yahoo, a expressão “modelos da temporada” foi aquela cujos resultados mais foram afetados pela proximidade dos portais. Além disso, ficou evidente que, pelo menos para algumas das expressões selecionadas, o iG é o portal cuja presença mais afeta o resultado das buscas: em metade dos casos não houve coincidência alguma com a busca no Yahoo.

Expressão	Google	Terra	Pop
Namorar alguém mais jovem	1,00	0,50	0,60
Petkurt	1,00	0,80	0,90
Celulite	1,00	0,70	0,80

Tabela 30: Taxa de redundância das saídas dos portais em comparação com o Google.

Os resultados do grupo Google são ainda mais discrepantes em relação aos esperados, ainda que contenham elementos que ajudem a corroborar as hipóteses. Apenas a expressão “namorar alguém mais jovem” afetou significativamente o resultado das buscas. Além disso, os resultados obtidos no Terra indicam que ele afeta apenas moderadamente as buscas. Nas duas tabelas, os valores em negrito indicam o portal onde a expressão foi selecionada. Pela premissa adotada, era de se esperar que, em todos os casos, esse fosse o *menor* valor da linha, o que ocorreu em apenas três dos sete casos. Isso não é suficiente para corroborar a hipótese.

Ainda que equívocos, os resultados nos permitem inferir que a proximidade dos portais afeta de *algum modo* o resultado das buscas, embora não em todos os casos e sempre. A questão que se põe agora é verificar se o *efeito* causado por essa proximidade é o de um atrator retórico, isto é, se a diferença obtida no resultado das buscas privilegia os documentos armazenados nos próprios domínios dos portais. Para determinar isso quantitativamente apenas contamos em cada produto obtido no passo (ii) a incidência de *links* para documentos do próprio portal. O valor obtido representa o que chamamos a *taxa de fechamento* da busca. No caso do grupo Yahoo, o resultado é mostrado na Tabela 31.

Expressão	UOL	Globo	iG
Modelos da temporada	0,90	0,00	1,00
Go-go boy	0,00	0,00	1,00
Monteiro Lobato	0,10	0,10	0,00
Celulite	0,10	0,00	0,00

Tabela 31: Taxa de fechamento dos portais do grupo Yahoo.

Esses resultados são significativos principalmente quando comparados à Tabela 29. Notamos que no caso do iG *todos* os resultados das buscas para as expressões “modelos da temporada” e “go-go boy” apontam para o próprio portal e *não* são obtidos no dispositivo de referência. Isso significa que o iG é o portal que mais se utiliza da busca como recurso para não permitir que os usuários saiam dos seus domínios. Nos demais casos, o resultado foi significativo somente para a expressão “modelos da temporada” quando buscada no UOL. A determinação desse efeito no caso do grupo Google é ilustrada na Tabela 32. Como se pode notar, pudemos constatar o efeito de atrator retórico somente na expressão “namorar alguém mais jovem” quando buscada no Terra.

Expressão	Terra	Pop
Celulite	0,00	0,00
Namorar alguém mais jovem	0,50	0,00
Petkurt	0,00	0,00

Tabela 32: Taxa de fechamento dos portais do grupo Google.

Note-se que não estamos preocupados com as causas tecnológicas desse comportamento observado em alguns portais – o UOL, o iG e o Terra. Tais causas podem ser, por exemplo, inteiramente acidentais – ou pelo menos os administradores desses portais podem estar inconscientes delas. A questão não é essa. O que viemos fazendo foi avaliar a influência dessas características possivelmente “acidentais” na prática concreta dos portais, procurando determinar se elas de alguma forma afetam os hábitos de navegação dos usuários. Visto desse modo, tais características tecnológicas deixam de ser acidentais, tornando-se práticas sociais compatíveis com o comportamento dos agentes envolvidos.

Infelizmente, os resultados obtidos não foram suficientemente convincentes para fornecermos uma resposta conclusiva às indagações feitas. Talvez tenhamos escolhido as expressões erradas; talvez simplesmente tenhamos elaborado hipóteses improdutivas. No entanto, esse pequeno experimento foi sugestivo: ele indica que (pelo menos) em alguns casos e (pelo menos) para algumas expressões de busca, a proximidade dos portais pode funcionar como atrator retórico, curvando o ciberespaço, fechando-o, de modo a manter os usuários em sua órbita. Fomos capazes de concluir pelo menos que existe uma probabilidade diferente de zero que isso ocorra – o que é pouco, mas pode não ser o bastante para simplesmente falsear as hipóteses iniciais.

3.3. O labirinto do surfista

Na seção 1.5. *O hipertexto e as teias do ciberespaço* assinalamos a tendência centrífuga dos usuários brasileiros da Internet e atribuímos boa parte desse comportamento à “distorção” da topologia do ciberespaço causada pelos portais de interesse geral. Supondo que esta “distorção” é intencional, imaginamos que ela seja causada por um conjunto de estratégias retóricas que seduzam ou forcem os usuários para permanecer nos domínios do portal, de modo a favorecer a comercialização de espaço publicitário, o que chamamos *atratores retóricos*. Ao longo deste capítulo, discutimos duas dessas estratégias: um domínio temático voltado para a satisfação do mais variado leque de interesses dos usuários e um deliberado – ainda que pequeno – fechamento do hipertexto na sua região mais centrífuga – o dispositivo de busca.

Ora, essa estratégia de fechamento deve se refletir diretamente na estrutura do hipertexto tal como produzido e mantido no interior dos portais de interesse geral. Podemos elaborar algumas hipóteses simples para dar conta desse fechamento do hipertexto. Em primeiro lugar, é previsível que os *links* nos portais apontem principalmente para páginas armazenadas em “canais” situados nos seus próprios domínios. E por “principalmente” queremos dizer que a quantidade de *links* apontando para fora dos domínios do portal deve ser menor que a quantidade correspondente num *site* qualquer, utilizado como grupo de controle. Além disso, é provável que a quantidade total de *links* nos documentos dos portais sejam menor que a contida em documentos que não pertençam a portais, independente da direção para onde apontam esses *links*. Isso significa que prevemos um “hipertexto degenerado” nas proximidades dos portais.

3.3.1. A retórica dos *links*: o hipertexto degenerado

Essas hipóteses são relativamente simples de verificar: simplesmente varremos todos os documentos de um portal, contando os *links* de cada um e verificando aqueles que apontam

para fora dos domínios do portal. Isso deve ser feito igualmente num conjunto qualquer de *sites*, utilizado como grupo de controle. Uma primeira dificuldade são os volumes em face da capacidade computacional disponível para um pesquisador individual sem financiamento próprio (nem tempo suficiente): a operação envolve centenas de milhares de documentos. A segunda dificuldade é que não é possível elaborar um *algoritmo* inequívoco para avaliar se um *link* aponta ou não para um *site* no domínio de um portal: alguns dos “canais” dos portais são reconhecidos enquanto tais tão somente por uma logotipo ou alguma indicação puramente visual do seu “hospedeiro”. Isso significa que o dispositivo de avaliação deveria incorporar *heurísticas*⁹ bastante complexas e difíceis de desenvolver. Novamente, essa dificuldade é insuperável para a presente pesquisa. Nossa decisão foi por simplificar o problema.

Em primeiro lugar, construímos um programa equivalente a um *crawler* simplificado de um dispositivo de busca, capaz de varrer sistematicamente todos os documentos no formato HTML de um portal, a partir da sua *home page*. Isso significa que desprezaremos uma razoável quantidade de documentos passíveis de serem exibidos num *browser* Internet e inspecionados por seres humanos, documentos possivelmente com *links*. Como a quantidade total de documentos a serem inspecionados (várias dezenas de milhares) é muito maior que a quantidade de documentos que serão ignorados (provavelmente algumas poucas centenas), essa escolha não deve afetar o resultado do ponto de vista estatístico.

A seguir, cada documento é varrido em busca de *links*. A estratégia aqui foi a mais simples possível: pesquisamos o documento em busca dos comandos HTML padrão de exibição de *links*. Com isso, uma vez mais ignoramos chamadas a *links* através de outras estratégias, como as possíveis sob várias linguagens de *script*. Novamente, nossa expectativa é que sejam ignorados relativamente poucos *links*, já que não é muito comum em páginas de conteúdo estático a construção de *links* dinâmicos através de linguagens de *script*. Supomos que a “lei” dos grandes números trabalhe em nosso favor: como esperamos alguns milhões de *links* na varredura, alguns poucos milhares não deverão afetar o resultado dos cálculos.

⁹ Um algoritmo é um procedimento ou fórmula que resolve um problema. Em computação, isso significa um programa que sempre resolve um determinado problema. Nesse domínio, ao contrário, não é garantido que uma heurística resolva o problema, já que o programa que a implementa simplesmente aplica a ele um conhecimento derivado da experiência.

O dispositivo de varredura realiza duas operações básicas: ele inventaria a quantidade de *links* incidentes no documento e armazena os *links* que apontam para documentos dentro dos domínios dos portais, para que sejam posteriormente inspecionados pelo dispositivo. Aqui a solução para a avaliação dos *links* foi efetivamente simplória. De uma forma geral (mas não sempre), os “canais” de um portal são caracterizados como um sub-domínio do domínio de hospedagem do portal. Por exemplo, o canal dedicado à mulher no portal Terra é acessado através da URL <http://mulher.terra.com.br>. Isso significa que apenas inspecionando a URL de um *link* é possível identificar, em muitos casos, se o endereço acessado está ou não dentro dos domínios dos portais. Assim, o algoritmo de avaliação simplesmente declara como de *link* interno uma URL que contenha na sua composição o nome de domínio do portal; caso contrário, ela é declarada um *link* externo.

Trata-se de um algoritmo excessivamente tosco: ele deve considerar como externos uma quantidade estatisticamente significativa de *links* internos. Como o erro introduzido na contagem desvia o cálculo *contra* a nossa hipótese, optamos por esse algoritmo falho, na suposição de que resultados favoráveis na presença de tais desvios ajudam a corroborar as hipóteses. Além disso, caso nosso dispositivo armazene as URLs consideradas como domínios externos, análises posteriores podem aumentar a precisão do cálculo, se necessário.

A limitada capacidade computacional disponível para a pesquisa nos levou a selecionar para a pesquisa somente três portais, aqueles considerados como os mais prototípicos na seção 3.1.6. *Segunda avaliação – a prototipicidade da amostra*: inspecionamos apenas o UOL, o Terra e o Globo. O processo de varredura operou pelos “níveis” do portal, em ondas sucessivas, considerando-se a *home page* (o documento inicial) como o nível 0. Os documentos apontados pelos *links* encontrados nesse nível 0 foram considerados o nível 1 da busca, e assim por diante. Novamente, nossas limitações nos levaram a inspecionar (incompletamente) apenas os documentos até o nível 4. Como grupo de controle selecionamos um único *site*, novamente por limitações de recursos: o *site* dedicado ao Projeto Gutenberg (<http://www.gutenberg.org>), um interessante projeto colaborativo destinado a digitalizar obras literárias de domínio público. Por fim, apenas as URLs apontando para documentos no formato HTML foram inspecionadas: os *links* que apontavam para imagens, vídeo, arquivos de som e outros recursos multimodais de diferentes tecnologias foram contados, mas não inspecionados. Com isso, durante alguns dias, o programa inventariou 318.418 URLs e 130.759 documentos dentro do critério adotado, com a distribuição mostrada na Tabela 33.

Portal	URL	URLs	Documentos
UOL	http://www.uol.com.br	113,454	53,078
TERRA	http://www.terra.com.br/capa/	51,949	41,747
GLOBO	http://www.globo.com.br	96,462	7,495
Projeto Gutenberg	http://www.gutenberg.org/	56,553	28,439

Tabela 33: Documentos inspecionados na observação.

O resultado inicial da varredura nos permitiu concluir imediatamente que o “nível” de localização do documento na hierarquia do *site* não tem correlação alguma com as características de navegação do hipertexto: a média de *links* (externos ou internos) por documento não obedeceu a um padrão reconhecível em qualquer dos níveis. Os resultados obtidos apenas para os *links* externos pode ser visto na Tabela 34. Isso significa simplesmente que, a medida que a navegação progride, a probabilidade de um usuário navegando ao acaso “escapar” dos domínios dos portais não é alterada. Utilizando uma metáfora, podemos dizer que, uma vez fisdado na rede retórica, o peixe permanece fisdado...

Nível	UOL	Terra	Globo	Gutenberg
0	0.00	4.00	0.00	8.00
1	8.17	1.72	1.79	9.33
2	6.86	1.60	2.77	6.28
3	8.03	1.27	2.69	26.40
4	6.08	6.05	3.23	–

Tabela 34: Média de *links* externos por página em cada nível da amostra.

Os totais obtidos na varredura podem ser vistos na Tabela 35. A coluna Domínios inventaria a quantidade total de *diferentes* URLs apontando para domínios externos obtida na busca. Em vista da pobreza do algoritmo de avaliação dos *links*, URLs como http://www.meusite.com.br e http://meudominio.meusite.com.br foram consideradas como pertencentes a domínios distintos, o que não é verdade. Apesar desse desvio, o resultado nos permite concluir que, dos portais considerados, o Globo é o mais centrípeto – e sua diferença em relação aos demais portais é significativa. Essa relação é confirmada se considerarmos a quantidade total de URLs que apontam para esses domínios, o que é exibido na coluna LE, que inventaria a quantidade de *links* externos encontradas na massa total de documentos.

Portal	Documentos	Domínios	LE	LI	L/DOC	LE/DOC	LE/LI
UOL	53,078	13,158	368,465	3,153,263	82.4440	5.8272	0.0707
Terra	41,747	4,740	209,521	2,667,426	71.3352	2.9283	0.0410

Globo	7,495	476	22,598	545,562	71.2510	2.0976	0.0294
Gutenberg	28,439	318	698,575	7,633,122	153.4648	12.5024	0.0815

Tabela 35: Resultados brutos da inspeção

Aparentemente, o *site* utilizado como grupo de controle é mais centrípeto que o Globo, se considerarmos isoladamente os valores da coluna Domínios. No entanto, isso não é verdadeiro, caso comparemos esse resultado com a quantidade de documentos encontrados em ambos, significativamente maior no caso do *site* do projeto Gutenberg, apesar dele ser dedicado a um único tema, ao contrário do Globo. Além disso, a quantidade total de *links* externos e internos do Projeto Gutenberg é significativamente maior que a dos demais *sites*. Isso é o suficiente para corroborarmos (provisoriamente) a hipótese do hipertexto degenerado, o que fica ainda mais evidente se compararmos os valores da coluna L/DOC, que lista a quantidade média de *links* por documento, isto é, a operação $(LE + LI) / \text{Documentos}$. Assumimos que a quantidade de *links* (externos e internos) encontradas no *site* do Projeto Gutenberg representa de alguma forma o “padrão” do hipertexto: é isso o que significa um grupo de controle. Se isso é verdade, é fácil concluir que os portais, além de centrípetos, degradam o hipertexto: a trajetória resultante configura um espaço fechado e pouco diversificado, mesmo nos seus próprios domínios. Portanto, é possível concluir que a organização do hipertexto nos portais pressupõe um usuário surfista e não um usuário navegador...

No entanto, o resultado obtido na coluna LE/LI, que representa a razão média entre os *links* externos e internos por documento e exprime a probabilidade de um usuário navegando ao acaso “escapar” dos domínios do portal, não foi rigorosamente o esperado. O resultado obtido para o UOL não é significativamente menor que o do grupo de controle, como seria de se esperar. Esse resultado, porém, é significativo para o Terra e o Globo. É provável que a fragilidade do algoritmo de avaliação dos *links* esteja interferindo com o resultado.

Para procurar eliminar esse efeito, inspecionamos visualmente os diferentes domínios encontrados no primeiro nível dos portais, procurando determinar quais eram efetivamente os domínios externos. Assim, navegamos para todos os domínios elencados como externos nesse levantamento e buscamos verificar visualmente se eles se consideravam parte do portal ou um *site* independente. De uma forma geral, encontramos inúmeros casos de *sites* que exibiam claramente sua filiação ao portal, utilizando principalmente uma estratégia visual – a exibição da barra de navegação do portal hospedeiro. Com isso, ao invés de 170 diferentes domínios externos encontrados na busca de primeiro nível no UOL, encontramos apenas 84 *sites*

realmente sem filiação ao portal. Esse número baixou no Terra de 23 para 9 domínios diferentes, enquanto que no Globo ele foi reduzido de 12 para 9. Fizemos, então, a seguinte suposição: imaginamos que essa diferença se propagaria linearmente para todos os demais níveis; além disso, imaginamos também que essa diferença também se propagaria linearmente para as URLs de citação, isto é, que os *links* considerados externos obedeceriam à mesma proporção encontrada para a diferença entre os domínios. Com isso, elaboramos um novo cálculo, “ajustando” os resultados obtidos para reduzir a fragilidade do algoritmo de avaliação. O resultado está ilustrado na Tabela 36.

Portal	Documentos	Domínios	LE	LI	L/DOC	LE/DOC	LE/LI
UOL	53,078	6,500	186,443	3,335,285	82.4440	2.9486	0.0358
Terra	41,747	4,740	209,521	2,667,426	71.34	2.9282	0.0410
Globo	7,495	357	5,650	562,511	71.2510	0.5244	0.0074
Gutenberg	28,439	318	698,575	7,633,122	153.4648	12.5024	0.0815

Tabela 36: Resultado ajustado da inspeção

Com essa operação de ajuste, a razão entre *links* externos e internos passa para valores dentro dos parâmetros esperados: ela é significativamente maior no grupo de controle. Além disso, todos os demais resultados continuam compatíveis. Isso nos leva a concluir que, pelo menos provisoriamente, as “previsões” feitas foram corroboradas: os portais se estruturam como um hipertexto degenerado e fechado em direção ao seu próprio centro.

3.3.2. A retórica da publicidade – surfando no UOL

Nossa hipótese condutora, da qual inferimos as hipóteses testadas na seção 3.3.1. *A retórica dos links: o hipertexto degenerado*, consiste em postular uma estratégia retórica destinada a manter os usuários nos domínios dos portais, de modo a favorecer a comercialização do espaço publicitário. Como pudemos constatar, isso se reflete diretamente na estrutura navegacional do hipertexto construído no interior dos portais. Ora, esse é apenas um dos elementos constitutivos do *design* multimodal (no sentido próprio do termo emprestado por Kress e Van Leeuwen). É de se supor que todos os demais elementos do *design* estejam igualmente a serviço dessa retórica.

Em primeiro lugar, é preciso assinalar que em nenhum dos exemplares da amostra encontramos *links* para fora dos domínios dos portais na sua primeira página (*home page*), ao contrário do grupo de controle, onde encontramos oito *links*. A *home page* dos portais parece funcionar, portanto, como o centro de gravidade retórico daquele domínio. Desse modo, a análise apenas dessa página pode ser suficiente para inferirmos os principais elementos do *design* multimodal envolvidos nessa estratégia retórica.

The screenshot shows the UOL homepage layout. At the top, there's a navigation bar with various services. Below it, a search bar and a main content area with news articles. The right side features several promotional banners for products like Dell laptops and Submarino MP3 players. The page is densely packed with text and images, typical of a news portal from that era.

Ilustração 22: A *home page* do UOL em 20 de novembro de 2006.

Consideremos a Ilustração . Essa página, aqui exibida incompletamente, contém 221 *links* (todos internos, naturalmente); desses, 121 são *links* destinados à navegação para os diversos canais e serviços disponíveis, 56 são *links* apontando para itens de publicidade e venda de produtos e apenas 39 (menos de 18%) referem-se a itens de informação relacionados ao “conteúdo” do UOL. Do ponto de vista da disposição visual, esse conteúdo ocupa pouco menos da metade da área disponível, enquanto que a publicidade ocupa $\frac{1}{3}$ da área disponível, incluindo a área mais “nobre” da página – aquela diretamente em frente à linha de visão do

usuário. Essa organização impede que o usuário se concentre exclusivamente no conteúdo do portal, assegurando que sua atenção seja dirigida em alguma medida para a publicidade.

A questão que se põe agora é indagar a respeito da eficácia dessa estratégia. Suponhamos um usuário que se comporta como um surfista aleatório, navegando ao acaso nas páginas do portal. Esse surfista hipotético não tem preferências nem interesses, não tomando decisões com base neles, sendo, portanto, incapaz de elaborar uma estratégia de navegação no portal. Para ele, cada item de informação (*link*) tem a mesma probabilidade de ser selecionado. Assim, ele tem probabilidade de 0.25 de selecionar um *link* publicitário, o que significa que uma vez a cada quatro ele será atraído para um item de informação publicitária.

Suponhamos agora que esse usuário tenha interesses e preferências relativamente bem definidos, o que corresponde, no nosso modelo, a uma probabilidade maior de selecionar um item de informação correspondente a esses interesses e preferências, mas não à certeza na seleção. Ora, é fácil verificar que, quanto *mais diversificado* esse interesse, *menor* é a redução daquela probabilidade de 0.25 do surfista selecionar um item de informação publicitária. O caso ideal de diversificação de interesses corresponde a um surfista quase aleatório, que não tem interesse específico em publicidade. Mesmo neste caso, a probabilidade de seleção de um item de informação publicitária tem valor muito próximo àquele previsto para o surfista aleatório. Por outro lado, se o interesse do usuário for tal que a probabilidade de seleção de um dos itens disponíveis chegue muito próximo à certeza, o *design* multimodal do portal é ineficaz, já que sua atenção nunca ou quase nunca se fixará nos itens de informação publicitária.

Isso significa que o *design* dos portais foi projetado para atender a usuários com interesses muito diversificados, cujas preferências por um ou outro item de informação sejam relativamente equivalentes. Usuários com fortes preferências e interesses muito específicos não são atendidos pelo *design* portais. Por outro lado, a relevância da publicidade é tal que usuários avessos a ofertas de produtos e serviços tenderão a ser afastados dos portais e não atraídos para eles. Em outras palavras, os portais (pelo menos no caso do UOL, mas é fácil verificar essa vocação nas demais instâncias do nosso *corpus*) são projetados para *consumidores*. Idealmente, tais consumidores devem ter interesses e preferências bastante diversificados.

4. AS VOZES DO MESMO – INCURSÕES NOS DISCURSOS DOS PORTAIS

A essa altura temos uma boa noção de alguns aspectos da retórica dos portais, pelo menos daqueles diretamente vinculados às estratégias de atração e manutenção dos usuários dos domínios dessa categoria de *sites*. No entanto, muito pouco sabemos sobre os *discursos* incorporados aos portais. Sem isso, a caracterização desse novo gênero do universo digital será necessariamente superficial. No entanto, em vista do gigantismo da tarefa (dezenas ou possivelmente centenas de milhares de documentos estão de algum modo associados aos portais), ela é impraticável. Assim, devemos limitar nossa pesquisa a alguns poucos aspectos do problema.

Porém, a seleção desses poucos aspectos é também problemática, em vista da enorme variedade temática desses *sites*, que por si só implica inesgotável variedade de abordagens. Um aspecto, porém, tem estado no centro das nossas preocupações ao longo dessa dissertação: o problema do controle do discurso, da voz do dissenso, dos contra-discursos cuja aparição é facilitada pela própria arquitetura tecnológica da Internet. Esse controle se exerce tanto pelo que o discurso *cala* quanto pelo que *deixa falar*. Esse, portanto, é o mote que nos interessa aqui.

Onde aplicá-lo, porém? Uma escolha é bastante óbvia. Na seção 3.3.2. *A retórica da publicidade – surfando no UOL* assinalamos que, em 20 de novembro de 2006, apenas 18% dos itens de informação da *home page* daquele portal correspondiam ao “conteúdo” assumido como sua finalidade, pelo menos a julgar pela sua própria avaliação (ver UOL, 2006). Desses, a metade correspondiam a conteúdo tipicamente encontrável nos jornais, vinculados tanto à versão on-line do jornal Folha de São Paulo quanto ao seu próprio canal de notícias (<http://noticias.uol.com.br>). Trata-se, portanto, do segundo mais importante foco de interesse isolado daquele *site*, ultrapassado apenas pela informação de natureza estritamente publicitária. Isso nos fornece um primeiro ângulo de abordagem do problema que nos move.

4.1. A retórica da verdade no relato jornalístico

Em 16 de dezembro de 2005, durante a realização de reunião plenária da Organização Mundial do Comércio, um grupo de pessoas que realizavam manifestação de protesto relacionada ao evento nas suas imediações acabou por se ver em confronto com a polícia de Hong Kong, resultando em inúmeras prisões e vários feridos. A questão de determinar quais as relações mantidas por este relato com um conjunto de eventos ocorridos numa certa cidade da China está longe de ser trivial. E mais: este pequeno relato não constitui o único possível de ser enunciado a cerca daqueles eventos. De fato, o incidente foi testemunhado por vários correspondentes de diversas agências noticiosas, cujos relatos foram publicados em diferentes órgãos de imprensa em todo mundo. Na verdade, o relato proposto no parágrafo anterior é inteiramente baseado em algumas dessas publicações: o relato publicado no portal Terra, o publicado no portal UOL e o publicado no portal Globo On-line, os primeiros baseados diretamente no despacho da agência EFE e o último, em vários despachos.

É notável nos vários relatos arrolados o fato de todos serem constituídos por proposições a propósito de um recorte da “realidade”; todos são compostos por sentenças assertivas a propósito de algum aspecto do mundo. Portanto, podemos por o seguinte problema: qual o referente das sentenças componentes dos relatos? Qual a relação entre o referente dessas sentenças e os fatos ocorridos em Hong Kong? Que relações os diferentes relatos mantêm entre si?

4.1.1. Referente, relato, objetividade

Começemos por colocar o problema do ponto de vista da própria imprensa, tal como exposto em alguns manuais destinados aos próprios jornalistas. Para esses manuais, “o que se busca é a notícia: o fato comprovado, relevante e novo”, pois “fatos são mais fortes do que declarações” (FOLHA DE SÃO PAULO, 2001: 27). Mais do que um relato, a notícia é vista como o próprio fato que relata. E para que isso seja possível, os manuais aconselham: “faça textos imparciais e objetivos. Não exponha opiniões, mas fatos, para que o leitor tire deles as

próprias conclusões” (O ESTADO DE SÃO PAULO, 1990: 18). Um tal programa só é alcançável se a linguagem na qual se baseiam os relatos jornalísticos for transparente, isto é, puder servir como “veículo” objetivo para os fatos narrados. Isto torna possível a proposta de que o jornalista seja um observador privilegiado dos fatos e seu mediador com o leitor. Ou, novamente no dizer desses manuais, “o jornalista funciona como intermediário entre o fato ou fonte de informação e o leitor” (FOLHA DE SÃO PAULO, 2001: 16). Esse programa se funda em alguns pressupostos básicos, a saber:

- i Os fatos são objetivos, isto é, os mesmos para todos os observadores;
- ii A verdade está nos fatos e não nos relatos, que apenas a transmitem;
- iii É possível utilizar a linguagem como um veículo transparente dos fatos.

Esses pressupostos estão notavelmente próximos da vulgarização das concepções da lógica clássica, tal como formulada por Frege. Para o filósofo, devemos assumir que todas as sentenças assertivas tenham um referente e que este referente seja a própria realidade (no caso, os fatos ocorridos em Hong Kong). Ao discutir o problema da significação, Frege assinala que “um nome próprio (palavra, sinal, combinação de sinais, expressão) exprime seu sentido e designa ou refere-se à sua referência”, entendendo por “referência” (do sinal) um objeto determinado, uma entidade “objetiva, na medida em que pode servir a vários observadores” (FREGE, 1978: 66), pré-existente à linguagem que o designa. O referente do sinal, o que é designado por ele, não é, portanto, o conceito. Este, o pensamento, numa sentença assertiva completa, “deve ser considerado como seu sentido” (FREGE, 1978: 67).

Para o enfoque fregeano, uma sentença como “Policiais e manifestantes enfrentam-se



 Policiais e manifestantes enfrentam-se em Hong Kong

Ilustração 23: Ilustração do relato do Terra

em Hong Kong”, que legenda a foto publicada no portal Terra (ver Ilustração), constitui-se num pensamento completo (expresso pelo sentido daqueles sinais) que designa determinados eventos ocorridos em Hong Kong. Tais eventos são *objetivos*, estão (ou pelo menos estavam no momento de sua ocorrência) disponíveis para qualquer observador. Portanto, para essa concepção a linguagem – ou pelo menos aquele subconjunto da linguagem composto por sentenças assertivas completas – lida com fatos passíveis de observação e não com conceitos ou

pensamentos. Estes são apenas o *sentido* dessas sentenças.

Essa distinção é crítica para o pensamento de Frege porque o centro de suas preocupações está no *valor de verdade* das proposições e este deve estar no mundo, nos objetos, e não no pensamento. Frege entende que a verdade não é uma relação entre sujeito e predicado, já que estes são “partes do pensamento; eles estão no mesmo nível no que tange ao conhecimento” (FREGE, 1978: 70). Isto leva Frege a reconhecer que o *valor de verdade* de uma sentença é a sua referência. Numa concepção como essa, “os juízos podem ser encarados como uma trajetória de um pensamento para seu valor de verdade” (FREGE, 1978: 70). Em outras palavras, se na cidade da China conhecida por todos (no ocidente) como Hong Kong um grupo de pessoas investidas pelo estado com o poder de polícia e um outro grupo de pessoas que manifestavam sua oposição à reunião da OMC entraram em confronto, ambos como agentes desse confronto, em 16 de dezembro de 2005, se este fato observável ocorreu, podemos então emitir o juízo de que a sentença da Ilustração , “Policiais e manifestantes enfrentam-se em Hong Kong [em 16 de dezembro de 2005]” é verdadeira.

Mas a sentença da Ilustração nos permite inferir mais do que simplesmente a união das proposições “policiais enfrentam manifestantes” e “manifestantes enfrentam policiais”. Ela parece sugerir que *ambas* as partes atacaram a outra ao mesmo tempo, sem que se possa inferir a existência de um agressor inicial ao qual o outro simplesmente *reagiu*. No entanto, a fotografia diz algo que não está dito na legenda, possivelmente contraditando-a. Dentre as centenas de fotos disponíveis para captar o fato, os editores *selecionaram* uma delas – e talvez a única – onde é a polícia a ser agredida e não os manifestantes.

Nossa cultura está habituada a considerar a fotografia jornalística como um “instantâneo” da realidade. De um modo geral, consideramos que a realidade “é definida com base na correspondência entre a representação visual de um objeto e o que normalmente vemos deste objeto à vista desarmada” (KRESS & VAN LEEUWEN, 1996: 163). Como já assinalamos na seção 2.1.3. *A multimodalidade e a produção do sentido em Kress e Van Leeuwen*, esta definição “naturalizada” da realidade é, contudo, fundada sobre a capacidade de resolução de cores e imagem da fotografia em 35 milímetros típica do cinema sob a tecnologia atual – e certamente não era a definição da pintura realista pré-lumière. Essa definição depende de um conjunto de fatores convencionais – perspectiva, diferenciação cromática, temas, etc. No entanto, estamos tão habituadas a ela que “esquecemos” seu caráter convencional e a assumimos como a própria “realidade” representada pela fotografia.

Um outro tipo de atitude diante do objeto pode ser percebida na fotografia publicada no Terra. Neste caso, o responsável pela publicação da fotografia se compromete com uma

determinada interpretação dos eventos – aquela onde a polícia é agredida pelos manifestantes de Hong Kong. O interessante é que este compromisso *não* se observa na frase usada como legenda: ambos realizam enunciados distintos. Como o *efeito de realidade* da fotografia é maior que o da linguagem verbal, o compromisso com uma *interpretação* dos fatos passa despercebido, tomado pelo leitor desatento como a própria *realidade*. Em outras palavras, na fotografia do Terra, a combinação entre imagem e texto se utiliza de um artifício *retórico* destinado a convencer o leitor de uma interpretação dos eventos.

Relações distintas entre imagem e texto são utilizadas na página do Globo (ver Ilustração); no entanto, elas conduzem a uma atitude geral diante dos eventos bastante próxima. A imagem aqui “fala” um discurso próximo à legenda do Terra, com policiais e manifestantes se enfrentando. No entanto, o texto assume imediatamente o compromisso carente na fotografia, quando assinala que “manifestantes entram em choque com a polícia”. Essa sugestão é corroborada por outra sentença (elíptica) encontrada na mesma Figura: “protesto e revide”. Essa sentença não parece ser simplesmente uma conjunção de duas proposições simples. A escolha vocabular sugere que “protesto” é um ato agressivo, conferindo à sentença características de uma implicação lógica. Tudo se passa como se fotografia e texto, nos dois *sites* citados, fossem os enunciados de um argumento – o de que a polícia *foi atacada* pelos manifestantes e simplesmente *reagiu*.

Impasse na OMC e violência do lado fora

Manifestantes entram em choque com a polícia para invadir reunião. Encontro de ministros termina amanhã, com pouca chance de acordo bom para o Brasil



Imagens dos choques



O que está em jogo?

> Protesto e revide



Ilustração 24: Ilustração do relato do Globo.

É fácil notar que os três relatos colhidos na Internet (mas *não* o “contra-relato” que construímos a partir deles) concordam numa *atitude interpretativa* básica, que corrobora o argumento expresso no parágrafo anterior: os manifestantes se “uniram para atacar” e, com isso, “romperam” a barreira policial; os “protestos (mas não a ação policial) foram violentos”;

a polícia lutou apenas para “evitar a invasão”. A Tabela 37, com extratos de todos os relatos, ilustra esse ponto.

Mas por que não é possível dizemos que essas “atitudes” não exprimem o que de fato ocorreu? Simplesmente porque podemos imaginar um conjunto de sentenças que afirmem precisamente o contrário (isto é, podemos imaginar uma proposição que afirme que “os policiais atacaram os manifestantes”) e, dados somente os relatos, não temos como emitir conclusivamente um juízo de verdade sobre suas proposições dentro da perspectiva fregeana. Para essa perspectiva, o fato de todos os observadores citados *concordarem* implicitamente com a negação dessa proposição, *não pode afetar seu valor de verdade*. Lembremos que, para Frege (como também para os manuais jornalísticos), o valor de verdade de uma proposição está no referente e *não* no sentido; logo, podemos imaginar um conjunto de condições tais que todos os observadores concordem com uma proposição e ela, ainda assim, seja falsa! Aliás, é precisamente isso que inferimos a propósito da discussão fregeana sobre as expressões “estrela da manhã” e “estrela da tarde”: até que os astrônomos “descobrissem” que seu referente era o mesmo e um planeta (e não uma estrela), todos os observadores concordavam com as proposições contrárias a essas!

	Sentença	Fonte
1	Um grupo de manifestantes conseguiu romper hoje o cerco policial	Terra
2	os ativistas se uniram para atacar as barreiras policiais	Terra
3	os violentos protestos ocorridos às portas do centro de convenções	UOL
4	forças de segurança lutavam para evitar a invasão do centro de convenções por manifestantes	Globo

Tabela 37: Manifestantes atacam a polícia em Hong Kong.

A questão em jogo aqui envolve a noção fregeana de “observador” (neutro) e sua concepção implícita de “conhecimento como descoberta” (da verdade). Para que seja possível emitir um juízo, a “trajetória de um pensamento para seu valor de verdade” deve ser conduzida por um observador capaz olhar o referente “objetivamente”, em si mesmo, independente das proposições através das quais o designamos. Nesse contexto, uma sentença verdadeira é uma sentença passível de ser enunciada por esse observador privilegiado, capaz de se libertar da “imperfeição da linguagem”, capaz de se tornar o sujeito de uma “linguagem logicamente perfeita”, onde cada sinal “designa, de fato, um objeto”, onde “nenhum sinal seja introduzido como nome próprio sem que lhe seja assegurada uma referência” (FREGE, 1978: 76). Para esse sujeito ideal, observador neutro da “realidade”, capaz de emitir juízos perfeitos sobre as proposições, o conhecimento é simples processo de descoberta de uma verdade que é

propriedade das coisas e não do modo como compreendemos as coisas. Para esse sujeito da Verdade, a linguagem precisa ser “um meio para alcançarmos uma verdade que está fora da linguagem, o que nos permite falar objetivamente sobre o mundo e, conseqüentemente, adquirir um conhecimento seguro sobre ele” (OLIVEIRA, 2000: 27).

Suponhamos que eu me chame H.G. Wells, invente um dispositivo capaz de viajar no tempo (e no espaço) e me transporte para a Hong Kong de 16 de dezembro de 2005. Suponhamos também que eu me coloque em posição tal que me torne observador privilegiado daqueles eventos e, com isso, constate que, de fato, foi a polícia quem atacou os manifestantes e não o contrário. Com isso, eu enuncio a sentença referida anteriormente (“os policiais atacaram os manifestantes”) e a incluo num relato publicado, digamos, pela hipotética Agência Green Peace de Notícias. Imaginemos agora que o leitor é o *webmaster* de um dos portais escolhidos para compor nosso *corpus* e precisa escolher quais dos relatos (contraditórios) utilizar na sua própria matéria. Como decidir qual das duas proposições é verdadeira? Em outras palavras, é possível emitir um juízo perfeito sobre os conhecimentos expressos nos vários relatos? A resposta da vertente de Frege seria, certamente, *não*, simplesmente porque o evento singular designado pelos relatos está inacessível – mesmo para um observador neutro. Se tudo o que temos à disposição é a linguagem utilizada para designar o objeto, não é possível emitir esse juízo perfeito – mesmo no quadro das concepções fregeanas. Com isso, somos levados, fregeanamente, a concluir que a linguagem não pode ser “um meio para alcançarmos uma verdade que está fora” dela... Isso nos leva a concluir também que o programa dos manuais de jornalísticos por “textos imparciais e objetivos” é, para dizer o mínimo, inalcançável.

4.1.2. Pressuposição e argumentação

O problema é que, no mundo jornalístico, tudo o que temos à disposição são os próprios relatos – e vimos que apenas com eles somos incapazes de emitir juízos conclusivos a propósito dos fatos. Desse modo, o editor (nem tampouco o leitor) não tem como avaliar a “veracidade” de um relato para publicá-lo: ele deve simplesmente *acreditar* nela. Trata-se de artigo de fé: supomos que o jornalista é apenas um “mediador” entre fatos e leitores, um observador neutro e privilegiado da realidade e que sua linguagem seja apenas um veículo

transparente dessa realidade. A alternativa a essa fé é assumir que todo relato, mesmo aqueles compostos apenas por sentenças assertivas, é argumentação.

Ao falarmos da atitude geral dos autores dos relatos perante os eventos narrados, estávamos na verdade assinalando o fato de que suspeitávamos de um certo *uso retórico* daquelas “sentenças assertivas”. Nossa “suspeita” nasce do fato de podermos conceber um relato diferente dos recolhidos nos portais tal que não entre em contradição direta e inequívoca com eles e, ainda assim, implique numa atitude distinta. Foi o caso do relato de segunda mão com o qual iniciamos esta seção. O que distingue esse relato hipotético dos demais? Simplesmente o fato de termos procurado, deliberadamente, “neutralizar” as sentenças quanto à afirmação de se os manifestantes atacaram ou não a polícia: nós não *nos comprometemos* com essa afirmação nem com sua negação. Seria igualmente possível construir o relato com base no compromisso oposto ao dos relatos, afirmando, por exemplo, a proposição “a polícia enfrentou os manifestantes” *sem afirmar jamais a proposição* “os manifestantes enfrentaram a polícia”. E mesmo neste caso os eventos poderiam ser reconstruídos de modo similar.

Nossa suposição é de que esses diferentes compromissos se originam em atitudes diversas previamente assumidas pelos observadores e se materializam em *propósitos retóricos* distintos. Se não existem observadores privilegiados e, portanto, sujeitos descomprometidos, como questionamos anteriormente, então toda locução implica numa *atitude* perante o referente. Isto significa que as “representações individuais” de que fala Frege (FREGE, 1978: 64) não podem ser simplesmente um fenômeno marginal à linguagem através da qual pretendemos falar do mundo.

No caso que estamos analisando, esse compromisso se manifesta primeiramente sob a forma de escolhas: embora, como vimos, seja possível relatar o evento sem marcar o ataque dos manifestantes às forças policiais, todos os “observadores” escolheram afirmar precisamente isto. Essa escolha fica bastante evidente na comparação dos relatos sobre a quantidade de feridos: entre 41 feridos (a julgar pelo relato do UOL), o Terra preferiu ressaltar no seu relato um único policial, mencionando-o *antes* dos demais. Do mesmo modo, o Globo On-line preferiu descrever o conflito como o enfrentamento dos manifestantes com a polícia (e não o contrário), ressaltando os ativistas como agentes do conflito. Isto está igualmente evidente na chamada incluída na Ilustração, quando escolheu-se marcar a ação da polícia como “revide”, uma reação à ação dos manifestantes.

Tais escolhas marcam um enunciador comprometido com determinada versão dos eventos, que procura construir uma imagem da polícia local como mera vítima de

manifestantes agressivos. Uma outra estratégia retórica utilizada para dar credibilidade a essa versão pode ser vista na sentença 3 da Tabela 37: trata-se de uma afirmação sobre o local onde ocorreu o incidente e, portanto, atende às convenções de “objetividade” do relato jornalístico (informar o que houve, quando e onde). No entanto, este enunciado assume como pressuposto um outro enunciado, que afirma que “o protesto foi violento”. Novamente, a escolha de um “modo de apresentação” (para retornarmos à expressão fregeana) particular. Que os eventos devem ter sido violentos não resta dúvida, pelo menos a julgar pela quantidade de feridos. No entanto, o “observador” escolheu fazer uma afirmação endereçada aos manifestantes e não, por exemplo, à polícia, como, por exemplo, na sentença hipotética “a ação policial violenta ocorrida às portas do centro de convenções”. Poderíamos dizer que esta última sentença seria plausível no contexto dos relatos, tanto pela quantidade de feridos entre os manifestantes quanto pelas “armas” utilizadas por eles (tambores e telefones celulares) para agredir a pacífica força policial munida de escudos, cassetetes elétricos, gás lacrimogêneo, gás pimenta e jatos d'água. Como vemos, os próprios relatos possibilitam um compromisso distinto do efetivamente assumido.

Ao trazer para a cena do relato esse enunciador comprometido como sujeito de uma pressuposição, o jornalista do UOL se utiliza de um artifício retórico simples e eficaz: assumir como dado, como fato, algo que precisa ser corroborado. Tal a força retórica da pressuposição: assumir como premissa o resultado de um raciocínio, sugerir que um argumento seja, na verdade, um fato.

Um terceiro componente desse jogo retórico consiste na própria estrutura do gênero discursivo: trata-se de um relato jornalístico. Por um lado, as convenções do gênero assumem a forma de uma narrativa – e esperamos que narrativas se refiram a eventos; por outro lado, como se trata de gênero jornalístico – esperamos igualmente que tais eventos sejam “reais”. Desse modo, as *expectativas do leitor* acabam por desempenhar um papel nesse drama retórico: o leitor tende a assumir tais expectativas como “pressupostos”. Na presença de um “leitor desarmado” isso tende a produzir um efeito de convencimento, uma “aura de verdade” que investe os enunciados do relato.

4.1.3. Protesto é guerra

Até agora, discutimos o que chamamos o *compromisso retórico* dos relatos em elementos marginais à trama textual e o fizemos a partir do questionamento do referente dos seus enunciados. No entanto, ele também se manifesta no núcleo conceitual dos relatos, nas imagens que organizam o discurso. A abordagem desse novo problema requer, porém, uma rápida conceituação preliminar.

Estudos relativamente recentes nos domínios da Semântica Cognitiva têm trazido a noção de metáfora para o centro das preocupações. Ao contrário da visão clássica do problema, a metáfora nessas pesquisas não é vista como uma “anomalia” no ramo principal da linguagem, como “uma espécie de agregado decorativo à linguagem comum, um dispositivo retórico para ser usado em certos momentos para obter certos efeitos” (SAEED, 2000: 346). Ao contrário, esses cognitivistas sustentam que a metáfora está no próprio ramo principal da linguagem, disseminada amplamente na própria linguagem quotidiana, embora admitam necessariamente a existência originária de conceitos não metafóricos. Elas são, contudo, o principal meio disponível para conhecermos o mundo e comunicarmos nossa experiência, permitindo-nos entender “um domínio da experiência em termos de outro” domínio (SAEED, 2000: 347).

Ao longo da nossa história cognitiva (e lingüística), aprendemos, a partir da experiência sensorial na interação com o mundo em que vivemos, alguns conceitos básicos, cristalizados sob a forma de “esquemas imagéticos” fundamentais e categorias elementares. É aplicável aqui o exemplo fornecido por Oliveira, numa longa citação: “a criança se move várias vezes em direção a certos alvos. Desses movimentos, emerge um esquema imagético sinestésico (uma memória de movimento) em que há um ponto de partida do movimento, um percurso e um ponto de chegada”. O esquema imagético produzido por essa experiência fundamental “ancora o significado de nossas expressões lingüísticas sobre o espaço” (OLIVEIRA, 2003: 35). Como essas experiências fundamentais são comuns aos membros da espécie e da comunidade, os significados podem ser compartilhados e, portanto, comunicados sem ambigüidades fundamentais. Dessa concepção segue-se imediatamente que os significados lingüísticos não são arbitrários, mas *motivados* pelas experiências sensoriais e interações com o mundo. Não menos imediata mais igualmente óbvia é a implicação de que “a forma [lingüística] deriva da significação” (OLIVEIRA, 2003: 34).

A medida que somos submetidos a novas experiências e a conceitos e categorias cada vez menos apreensíveis diretamente por essas experiências, utilizamos os mecanismos metafóricos para estender os conceitos e significados apreendidos anteriormente para esses novos domínios de experiência. A metáfora é, portanto, “um processo cognitivo que permite mapearmos esquemas, apreendidos diretamente pelo nosso corpo, em domínios mais abstratos, cuja experimentação é indireta” (OLIVEIRA, 2003: 36). Assim é que, a partir do esquema básico de movimento, podemos apreender um domínio abstrato de experiência indireta como a “vida”, que pode ser universalmente compreendida como uma viagem levando do nascimento à morte, desde as manifestações lingüísticas “elevadas” como a poesia (é irresistível aqui lembrarmos o célebre *nel mezzo del camin di nostra vita*) até às expressões quotidianas como “ela tem um bebê a caminho”.

Naturalmente, a generalidade de tais conceitos torna sua abrangência universal, tornando possível utilizá-los para abordar desde o discurso das ciências matemáticas até o discurso político. Sabemos, se assumirmos tais premissas, que os esquemas imagéticos básicos devem ser universais porque derivam de experiências igualmente universais, pelo menos no escopo das experiências (universalmente) compartilhadas por uma comunidade (particular). Queremos, porém, determinar como pode se dar a utilização desses esquemas universais no contexto de experiências, vivências, conceitos e categorias que *não são compartilhadas por todos os membros de uma comunidade*, mas apenas por alguns. Em outras palavras, queremos determinar se esses pressupostos contribuem para a compreensão das *ideologias*, esses modelos conceituais compartilhados por alguns grupos sociais, mas não por outros.

De uma forma geral, o domínio conceitual da guerra e dos assuntos militares tende a ser apreendido em termos dos esquemas imagéticos de *container* e força. Expressões como “perímetro de segurança”, “forçar a transposição” (de um obstáculo natural ou artificial), “fogo de barragem”, “cerco”, “assédio”, “irrompimento das linhas” (inimigas), “bolsão defensivo”, “envolvimento das forças” (inimigas ou amigas), “tropas de ocupação”, “posto de fronteira”, etc. são de utilização corriqueira tanto no jargão profissional e especializado, como em relatos jornalísticos a propósito da guerra. A aparente universalidade dessas metáforas provavelmente decorre da aparente universalidade (e permanência) da experiência bélica. Se nos é permitida uma pequena especulação paleolingüística, elas possivelmente nasceram com a própria linguagem, no momento em que algum dos nossos ancestrais antropóides decidiu que era mais simples recorrer a paus e pedras contra o grupo vizinho que procurar novos territórios de coleta e caça ou fabricar suas próprias ferramentas. Pelo menos essa é uma das

narrativas possíveis contadas por crânios esmagados e perfurados encontrados em inúmeros sítios paleontológicos.

Esse esquema imagético aplicado às relações (conflituosas) entre grupos sociais é igualmente aplicado às relações (possivelmente conflituosas) *dentro* do mesmo grupo social. É também esperada a utilização das metáforas do domínio conceitual estritamente militar quando a “questão social” se torna “caso de polícia”. O inventário a seguir, composto por 40% do *corpus* (mais especificamente daquela parcela do relato diretamente relacionada ao incidente envolvendo a polícia e os manifestantes), é ilustrativo da utilização desses esquemas conceituais aplicados indiscriminadamente a conflitos militares e sociais:

- i Manifestantes rompem cerco perto do local da reunião da OMC.
- ii Um grupo de manifestantes conseguiu romper hoje o cerco policial em torno do centro de convenções de Hong Kong.
- iii Os ativistas se uniram para atacar as barreiras policiais que mantêm vazias as ruas nas proximidades do edifício.
- iv Eles obrigaram dezenas de agentes a se dispersarem protegidos com escudos e armados com cassetetes elétricos.
- v A proteção dos acessos ao edifício está sendo reforçada com vários agentes com máscaras antigás, escudos antidistúrbios e cassetetes elétricos.
- vi Mais de 900 manifestantes foram detidos hoje pela Polícia de Hong Kong após os violentos protestos ocorridos às portas do centro de convenções onde ocorre a reunião da Organização Mundial do Comércio (OMC).
- vii Manifestantes entram em choque com a polícia para invadir reunião.
- viii Forças de segurança lutavam para evitar a invasão do centro de convenções por manifestantes recorrendo a jatos de água e sprays de gás de pimenta.
- ix A polícia de Hong Kong deteve mais de 900 pessoas para controlar os protestos nos arredores do centro de convenções.
- x Os detidos, mantidos sob custódia policial num perímetro guardado, na rua mesmo, serão levados para uma prisão no centro da cidade.

Esse inventário das metáforas conceituais ocorrentes nos relatos, por si só, nada nos informa a propósito do problema que queremos encontrar, isto é, sua utilização ideológica.

Afinal, a expressão “invadir (e ocupar) a fábrica” é utilizada tanto pelos operários revoltados quanto pela polícia convocada para repelir a invasão... No entanto, dois lados se confrontam na guerra: amigos e inimigos. No caso da “questão social”, é de se esperar que a identificação de um dos lados como “amigo” e do outro como “inimigo” seja uma atitude ideológica. Assim, procuramos associar as metáforas de *container* e força daquele inventário aos dois lados em conflito – manifestantes e policiais – nos vários relatos jornalísticos recolhidos, na busca de regularidade na sua utilização. O resultado é mostrado na Tabela 38.

Manifestantes	Polícia	Fonte
Manifestantes rompem cerco perto do local da reunião da OMC		
Um grupo de manifestantes conseguiu romper	o cerco policial em torno do centro de convenções de Hong Kong	Terra
Os ativistas se uniram para atacar	as barreiras policiais que mantêm vazias as ruas	Terra
	protegidos com escudos	Terra
	A proteção dos acessos ao edifício está sendo reforçada com vários agentes	Terra
	Mais de 900 manifestantes foram detidos	UOL
Manifestantes entram em choque com a polícia		Globo
para invadir reunião		Globo
a invasão do centro de convenções por manifestantes		Globo
	A polícia de Hong Kong deteve mais de 900 pessoas	Globo
	Os detidos, mantidos sob custódia policial num perímetro guardado	Globo

Tabela 38: Ocorrência de metáforas envolvendo os esquemas de *container* e força.

É importante ressaltar que a justificativa da luta desempenha hoje um importante papel na ideologia da guerra. Mesmo Hitler procurou justificar a invasão da então Tchecoslováquia como um ato de auto-defesa. Afinal, nossos poetas tendem a não mais glorificar os “saqueadores de cidades”: uma guerra “justa” é uma “guerra defensiva”, uma guerra travada em defesa do “nosso perímetro” ou um “revide” à agressão. Avaliando a Tabela 38 sob essa luz, notamos que a distribuição das metáforas conceituais *não é aleatória!* Os “caras bons” dentro do perímetro devem ser protegidos dos “caras maus” que querem invadi-lo pelos heróicos soldados da pátria. Protesto e revide... A simples realização de uma manifestação política ou social é encarada como uma declaração de guerra e é o bastante para a mobilização

das forças patrióticas em defesa do solo sagrado. A identificação com um dos lados em conflito é simples e direta, desprovida de subterfúgios. Os relatos, portanto, estão *comprometidos* com um dos lados em conflito e isso se manifesta também na distribuição das metáforas conceituais encontradas nos relatos selecionados.

4.2. A construção da identidade feminina nos portais

Na seção anterior discutimos implicitamente a função de compromisso ideológico do discurso, sua capacidade de *construir* o mundo que representa e argumentar em favor de *uma* representação particular desse mundo possível. Um outro estudo de caso que deverá nos auxiliar na caracterização discursiva desse gênero digital emergente é a função identitária do discurso, sua capacidade de construir uma representação do Sujeito na interação. Por comodidade, o próximo estudo de caso abordará os canais dedicados à mulher dos principais portais de interesse geral.

4.2.1. Eu, mulher jovem, magra e sensual

Uma simples contagem de *sites* na Internet cujo projeto comercial implique na afluência de público indica que uma das funções primárias da cor é marcar sua diferença em face dos *sites* concorrentes. A cor desempenha aqui um papel de colaborar na “formação” da identidade do próprio *site*. Em outras palavras, o modo cor é um recurso a serviço do discurso identitário. Todos os principais portais de interesse geral se comportam do mesmo modo, e esse comportamento se estende aos seus canais. Igualmente relevante na constituição dessa “identidade comercial” costuma ser a própria diagramação dos *sites*, cada um procurando imprimir suas próprias características sem, contudo, fugir ao padrão global. Isso visa tanto assegurar o reconhecimento do gênero (digital) quanto da própria instância pelo usuário visitante. No entanto, é notável nos portais que cor e imagem não sejam escolhidas

arbitrariamente, visando tão somente fixar uma identidade de *site*, mas sirvam igualmente ao compartilhamento desta identidade com seu público alvo.

Consideremos o canal Mulher, do Terra, cuja versão de 21 de novembro de 2006 é mostrada na Ilustração . A combinação de cores predominante na página é um intervalo entre o rosa e o vermelho, mais próximo deste no espectro que daquele. Pensemos na associação do vermelho à sensualidade e poderemos interpretar esse modo no *site* como significando feminilidade saturada de sensualidade. Mas não inteiramente, já que a cor não alcança o vermelho: trata-se de sexualidade contida, discreta. Esse discurso da sexualidade contida se manifesta igualmente através de outros modos semióticos: de saída na imagem tema que caracteriza o logotipo do site – a rosa, clássico símbolo do amor e da paixão, sempre, porém, aplicado ao universo feminino.

Ilustração 25: O canal Mulher do portal Terra (acesso em 21 de novembro de 2006).

Em artigo sobre o potencial semiótico dos brinquedos, Caldas-Coulthard e Van Leeuwen assinalam que:

A cor pode ser um importante significante de gênero. Pense na tradicional oposição entre as roupas de bebê azuis e cor de rosa, ou os marrons e cinzas que dominam as roupas dos homens e os tons mais claros que estão associados às roupas femininas (CALDAS-COULTHARD & VAN LEEUWEN, 2004).

Ao construir sua identidade, o canal assume um conjunto de marcas visuais de gênero associadas ao seu próprio público alvo: ele projeta em si sua imagem do outro. Ao fazê-lo, porém, ele constrói no outro sua própria imagem: com esse recurso, o canal constrói tanto sua própria identidade quanto a identidade do seu interlocutor, porque “quem tem o poder de representar tem o poder de definir e determinar a identidade” (SILVA, 2000: 91).

The image shows a screenshot of the UOL website's 'Estilo' (Style) section. At the top, there is a banner for Visa with the text 'Só falta isso para você conhecer uma das 7 Maravilhas. Passe o mouse.' Below the banner is a navigation bar with links like 'UOL 10 ANOS', 'ASSINE', 'BATE-PAPO', 'BUSCA', 'CENTRAL DO ASSINANTE', 'E-MAIL', 'SHOPPING UOL', and 'ÍNDICE PRINCIPAL'. The main content area features a large photo of Tom Cruise and Katie Holmes in wedding attire, with the headline 'Armani vestiu os noivos e toda a família de Tom Cruise para o casamento'. To the left is a vertical menu with categories like 'Beleza e Boa Forma', 'Blogs', 'Casa', 'Etiqueta', 'Família', 'Gastronomia', 'Horóscopo', 'Moda', 'Revistas', 'Saúde', and 'Últimas Notícias'. To the right are several promotional boxes for 'Shopping UOL' and 'UOL Links Patrocinados'. At the bottom, there are three small boxes with titles 'NOVA REGRA', 'FAMÍLIA', and 'DANUZA LEÃO', each with a short text snippet. The date '21/11/2006' is visible at the bottom left.

Ilustração 26: O canal dedicado à mulher do UOL (acesso em 21 de novembro de 2006).

Mas queremos fixar algo além dessa constatação: como o discurso só existe no diálogo, na interlocução, esse poder de representar não se resume a *se* representar, a definir sua própria

identidade, mas se estende à representação do outro. No caso do canal, esse diálogo é indireto e temporalmente diferido: a voz que predomina é a do locutor, que orienta seu discurso ao que é sua própria representação do outro. O outro fala indiretamente; sua relação de poder é assimétrica. Portanto, a locução do canal determina e define não apenas sua própria identidade, mas a identidade do seu interlocutor.

Esses artificios de representação de gênero não são exclusividade do canal Mulher do Terra. Vejamos, por exemplo, o canal Estilo do UOL, igualmente dedicado ao público feminino, cuja versão de 21 de novembro de 2006 pode ser vista na Ilustração . Aqui a opção não foi, evidentemente, pelos tons de rosa, no mínimo para assegurar diferenciação com os canais competidores. No entanto, os tons claros citados por Caldas-Coulthard e Van Leeuwen predominam, associados ao tema floral.

Todos os canais dedicados à mulher dos portais mais prototípicos utilizam essa estratégia identitária. A revista Manuela, o *site* assumido pelo portal Pop (<http://www.pop.com.br/barra.php?url=http://www.revistamanuela.com.br>), tem seu visual inteiramente construído com tons de rosa. Um outro canal interessante é o do iG (<http://delas.ig.com.br>), cujo logotipo é baseado no rosa, vermelho e laranja, insistindo na associação da feminilidade à paixão e às “cores do amor” (ver Ilustração).

Isso nos remete à reflexão sobre a imagem (visual) da Mulher (como Objeto de representação) projetada pelos canais através da fotografia (e ilustrações, no caso estrito do

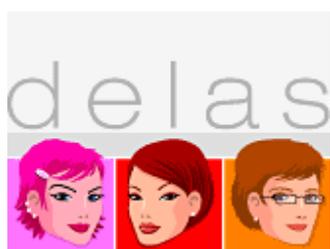


Ilustração 27: O canal do iG

iG). Em todos os casos observados, as mulheres são jovens, magras e esbeltas. Aliás, todas são “modelos”, profissionais da moda ou artistas do cinema e televisão. Na imensa maioria dos casos, as fotografias são “encenadas”, sendo raramente encontradas mulheres “comuns” (e “verdadeiras”) nos portais. Aqui, a ênfase fotográfica é “ficcional”, muito raramente utilizando a fotografia jornalística, baseada em

“fatos” e acontecimentos, a contraparte visual do relato jornalístico. Uma única exceção a esse padrão de beleza foi encontrada: na Ilustração , encontramos a foto de uma mulher idosa. No entanto, essa foto aparece no contexto de matéria sobre o envelhecimento! Isso significa implicitamente que a foto não se destina ao consumo de mulheres idosas, mas daquelas mais jovens, preocupadas com o envelhecimento (futuro). Assinale-se que essa caracterização do universo feminino nos portais não é simplesmente uma função da faixa etária do público usuário; afinal, 10% dos usuários da Internet tem mais de 45 anos (fonte: CGI.BR, 2006). Trata-se de uma *projeção*, uma construção identitária de natureza argumentativa: ao construir

uma imagem do outro voltada para seus desejos e aspirações, o discurso dos portais visa seduzir, do mesmo modo que os recursos análogos utilizados na publicidade.

A síntese (ou talvez o emblema) da beleza feminina nos é fornecida no Terra (ver Ilustração): contra um pano de fundo “tórrido”, baseado nos tons vermelhos da paixão, a Beleza é jovem, quase adolescente; magra, quase anoréxica; sensual, com o reforço artificial da lingerie, emblema da sensualidade em nosso tempo; e narcísica – ela é flagrada diante do espelho, realçando (mas não “produzindo”) sua beleza com cosméticos. Note-se que esta é a foto de ilustração da seção Beleza do *site* (ainda que acessível da *home page* do Terra, este *site* é igualmente acessível a partir do canal Mulher, mas não de nenhum canal “masculino”), que também dispõe de páginas dedicadas ao homem. No entanto, a figura masculina de ilustração (ver <http://beleza.terra.com.br>) exhibe sua beleza nua para os outros e não para si, diante do espelho. Ainda que magra, sua beleza é baseada em músculos, mais próxima do atleta que do “homem de negócios bem sucedido”, possivelmente compartilhando seu ideal de beleza com o imaginário feminino. Trata-se da beleza do universo do consumo: efêmera, que não pode amadurecer e deve ser consumida rapidamente, que requer os “melhores produtos disponíveis no mercado” e que, portanto, precisa consumir para ser consumida; adorno e não substância; contingência dos produtos do consumo.



Ilustração 28: La Beauté, segundo o Terra.

Assinalamos incidentalmente algumas das marcas de diferença da identidade feminina com a projetada para a figura masculina. Mas, sobretudo, indicamos implicitamente as diferenças dessa projeção com aquela imagem da mulher que *não* se quer marcar: ao delimitar uma mulher jovem, magra e sensual, construída em torno do padrão de beleza projetado pelo universo da moda feminina, o que os portais excluem são as mulheres não-jovens, não-magras e não-sensuais. Como esses atributos provavelmente estão relacionados de algum modo a parcela significativa (para dizer o mínimo) das usuárias do canal e seu propósito é abranger a todos, essa exclusão típica da construção da identidade não pode ser absoluta. Assim, essa

identidade deve funcionar como um referencial imaginário, uma meta a ser alcançada (pelo menos no imaginário), uma projeção de desejos e aspirações.

Ora, assumimos que “as identidades são fabricadas por meio da marcação da diferença” e que “essa marcação da diferença ocorre tanto por meio de sistemas simbólicos de representação quanto por meio de formas de exclusão social” (WOODWARD, 2000: 30). Ora, se isso for tudo o que podemos dizer sobre a construção da identidade, é forçoso considerar que aquelas mulheres com um mínimo de autoconhecimento que não sejam capazes de se reconhecer na categoria de Mulher projetada pelo canal estarão imediatamente excluídas – e não serão usuárias do portal, condenando ao fracasso sua estratégia retórica. O fato é que isso *não* é tudo o que podemos dizer sobre a construção da identidade: devemos considerar que ela, ao funcionar como uma projeção de desejos e aspirações, permite que se incluam sob suas marcas todos os que, de algum modo, compartilham esses desejos e aspirações.

4.2.2. Dize-me o que consumes e...

A identidade visual, portanto (pelo menos no escopo dos canais dedicados à Mulher), marca antes de tudo a diferença com o que o Sujeito que nela se reconhece *não deseja* para si, funcionando como um marco imaginário postado num futuro alcançável. Nesse contexto, o canal funciona como um mecanismo de inclusão, assumindo fornecer os *meios* para a usuária alcançar aquela identidade projetada. Isso indica o caminho que devemos tomar no prosseguimento da nossa análise. Trata-se de verificar que meios de inclusão são esses.

O gênero não digital mais próximo desse tipo de canal são as revistas (femininas). Tal como sua contraparte não-virtual, os canais se organizam em torno de seções temáticas, que procuram dar conta dos interesses específicos da audiência pretendida. Ora, a análise sumária dessas seções deverá iluminar o que dissemos no parágrafo anterior. Consideremos a Tabela 39, que lista as seções temáticas encontradas nos três *sites*. Que interesses essas seções projetam? As seções dedicadas à Beleza e à Moda são uma unanimidade. Isso pode ser relacionado àquela função que atribuíamos a esse tipo de *sites*: é possível assumir que os canais procuram, de algum modo, fornecer os meios das usuárias alcançarem aquela identidade visual projetada de que falávamos. Portanto, descobrimos uma outra função para

essa identidade: ela funciona igualmente como um *atrator retórico*, um “pote de ouro” que convida ao *surfing* pelas páginas do canal.

Seção	Terra	UOL	iG
Últimas notícias	♥	♥	
Beleza	♥	♥	♥
Casa	♥	♥	
Culinária, gastronomia, gourmet	♥	♥	♥
Mães e filhos, família	♥	♥	
Moda	♥	♥	♥
Noivas	♥		
Relacionamento	♥		♥
Saúde	♥	♥	
Sexo	♥		
Vida profissional	♥		
Bem-estar da mulher	♥		
Blogs		♥	
Etiqueta		♥	
Horóscopo, astral		♥	♥
Revistas		♥	

Tabela 39: As seções temáticas dos canais dedicados à mulher

Acrescentemos a esses temas relacionados aos padrões de beleza feminina também o canal dedicado à Saúde, pelo menos em parte. Na seção do Terra, 32% das matérias listadas na primeira página em 26 de novembro de 2006 estavam relacionadas à beleza física, como alimentação e ginástica estética. Além disso, 46% da publicidade fornecida ali era de produtos relacionados à beleza. No UOL, esses números caem para 30% e 38%, respectivamente, mas ainda continuam significativos.

Extremamente relevante para a composição dos interesses do público alvo dos canais são os temas relacionados às atividades domésticas: casa, culinária, maternidade e vida familiar. A “rainha do lar” finalmente ressurgiu das cinzas na composição da identidade feminina nos portais. Dois outros temas incidentes apenas no Terra parecem estar à serviço desse “atributo” identitário: a seção dedicada às Noivas e à Vida profissional da Mulher. Este último é de especial interesse: 31% das matérias compiladas em 21 de novembro de 2006 buscavam conciliar a “rainha do lar” à “mulher trabalhadora”, com matérias como “ser mãe também é profissão”, “concilie os filhos e o trabalho”, “levar seu filho ao médico não pode causar demissão”, etc.

Uma última observação interessante refere-se à seção Horóscopo, disponível no UOL e no iG. Trata-se de canais com vida própria, acessíveis a partir da *home page* dos portais. No entanto, sua apresentação também como seção dos canais dedicados à mulher (mas *não* de canais dedicados a outros gêneros) é, para dizer o mínimo, “suspeita”. No UOL, o temário desse canal é cuidadosamente independente do gênero; o fornecimento de serviços de “horóscopo amoroso” chega mesmo a conceder primazia aos homens: ao indicar a seleção do parceiro amoroso, a legenda indica “signo dela(e)”, como se principalmente homens consultassem o serviço (ver Ilustração). No entanto, a composição da publicidade é a mesma do canal feminino: 66% da publicidade está orientada ao público feminino. É como se os anunciantes implicitamente confessassem que conjunção dos astros é “coisa de mulher”, provavelmente deixando para os homens a cotação das bolsas...

Em alguns momentos, exemplificamos as características da construção da identidade feminina chamando a atenção para a composição da publicidade nas páginas. Trata-se de aprofundar um pouco mais o problema. Assumimos que, do ponto de vista do problema que nos interessa aqui – a construção da identidade feminina, a publicidade não é apenas indício dessa identidade tal como construída no *site*: a nosso ver, o discurso da publicidade desempenha um papel ativo nessa construção, na medida em que ele também tem o poder de representar a identidade. Podemos dizer a propósito desse discurso, pelo menos no escopo dos canais dedicados à mulher, o mesmo que dissemos a propósito do discurso do próprio canal.

Não vamos aqui insistir nesse ponto. Basta-nos assinalar que a composição publicitária



A screenshot of a web interface for a horoscope service. The title is "Atração Astral" in orange. Below it, the text reads "O que os astros dizem sobre seu relacionamento". There are two dropdown menus: "SEU SIGNO" and "SIGNO DELA(E)". At the bottom, there is an orange button labeled "CONSULTAR".

Ilustração 29: Horóscopo no UOL.

dos canais analisados em 26 de novembro de 2006, sintetizada na Tabela 40, está bem próxima da sua composição temática. Como se pode notar, 35% dos anúncios veiculados está relacionados à beleza feminina, enquanto que 25% deles se relacionam ao seu papel de “rainha do lar”. Parece-nos relevante aqui a existência de duas vozes discursivas representando (e, portanto, definindo) a identidade feminina: a voz do canal e a voz dos anunciantes. Essas duas vozes atuam em conjunto, colaborando segundo um único plano na constituição de uma imagem da mulher que o *site* quer alcançar.

A única voz calada é da própria usuária, condenada a se representar no discurso do outro e não no próprio...

Categoria do produto	Terra	UOL	iG
Cosméticos, produtos e serviços de beleza em geral	2	2	–
Bebidas leves	1	–	2
Enxoval doméstico	1	1	1
Escola (1º e 2º graus)	–	1	–
Vestuário feminino	–	3	–
Empregos	–	2	1
Presentes, brindes	–	1	–
Supermercados	–	–	1

Tabela 40: Categorias de produtos anunciados nos portais da amostra.

4.2.3. Relacionamentos – o jogo da sedução

Há ainda um último aspecto do discurso identitário sobre a mulher que é preciso abordar. Na Tabela 39, podemos notar que o tema Relacionamento é relevante na visão do Terra e do iG. Trata-se aqui da versão portal da temática do “sexo para as mulheres”. Note-se que o Terra indica seu próprio canal de sexo como seção do canal Mulher; no entanto, se avaliarmos o conteúdo daquele canal, notaremos que não há indicação alguma de que ele assuma de algum modo conteúdo especificamente “feminino”. Ao contrário, nele nada há de distinto em relação aos canais de sexo “para homens” dos demais portais. Desse modo, parece-nos adequado considerar, mesmo que por hipótese, que a seção Relacionamento cumpre esse papel no canal.

De um modo geral, no canal o sexo é cuidadosamente abordado de modo a não relacioná-lo à pura sensualidade, como nos *sites* “masculinos”. Um exemplo típico são as notícias listadas logo na primeira página da seção: dentre as 40 chamadas disponíveis, embora várias digam respeito diretamente à sexualidade, a palavra “sexo” não aparece uma única vez entre as 317 utilizadas. Do ponto de vista vocabular, o “relacionamento” está relacionado a expressões como: flerte, affair, sedução, amor, namoro e “pintar um clima”.

O “relacionamento” é a contraparte pública do sexo: este é aqui contido, torna-se matéria íntima e privada, inadequada ao tratamento “público” de um *site*; na sua versão

“pública, ele é relacionado aos sentimentos, não aos instintos. Quando é tratado diretamente, é matéria de clínica médica. Por exemplo, na seção Saúde da Mulher é apresentada toda uma nova seção dedicada a “Sexo e fertilidade”; além disso, matéria abordando a masturbação publicada no iG o faz em tom de conselho médico (ver http://delas.ig.com.br/colunas/fale_de_sexo/, em 26 de novembro de 2006).

No entanto, o clássico imaginário do amor no universo feminino torna-se apenas isto – um clássico, um filme antigo (e *cult*) a ser visto somente em reprises, que deve ser apreciado, mas não necessariamente praticado. É curioso que das seis matérias tematizando diretamente o “amor” no Terra em 26 de novembro de 2006, em cinco delas o amor é definido como uma patologia na melhor das hipóteses equivalente ao alcoolismo. Como se pode notar pela Tabela 41, na única matéria em que o amor não é encarado como uma patologia ele, no entanto, é implicitamente tratado como um valor já gasto, que requer um enfoque original para ser assumido pelo universo feminino de valores.

Como tratar o amor patológico
Descubra se você está passando do limite do amor saudável
Amor patológico: você já viveu um?
Dez formas originais de declarar seu amor
Os 12 passos do MADA (Mulheres que Amam Demais Anônimas)
Mulheres que Amam Demais Anônimas (MADA)

Tabela 41: O amor como patologia, segundo o Terra.

As palavras chave no canal são “namoro” (seis matérias publicadas) e sedução (duas matérias); são elas que melhor representam o conteúdo da seção, provavelmente o melhor sentido para seu tema – o Relacionamento. É possível relacionar essa “preferência” de enfoque do tema ao próprio universo do consumo, assinalado mais atrás, eminentemente efêmero e fugaz: os relacionamentos “saudáveis” (ou pelo menos os preferíveis) não são longos, nem excessivos; não são bens duráveis (duas matérias tematizam os homens “grudentos”, que não “largam o pé” das mulheres), que devem ser consumidos imediatamente.

Essa efemeridade do relacionamento pode ser também inferida do papel desempenhado pelo casamento nesse universo de interesses. Consideremos a Tabela 42, que lista chamadas para matérias que tratam do casamento do ponto de vista da mulher projetada pelo Terra. Como se pode notar, o casamento é mais um problema que uma extensão dos relacionamentos entre parceiros: não há matérias que tematizem positivamente o casamento. Ele é uma

instituição sem prazer (matéria 1), sujeito à violência doméstica (matéria 4), que requer flertes fora dele para se sustentar (matéria 5), instituição cuja manutenção “com paixão” requer a divulgação de segredos até hoje não revelados (matéria 6). É curiosa também a escolha de foco da matéria 3: em lugar de assinalar que 83% das pessoas estão satisfeitas com o casamento, um foco positivo possível na pesquisa, ele prefere enfatizar o fato de que 17% delas não estão. Pode-se interpretar, por fim, que o casamento é freqüentemente escolha prematura (matéria 2) e que merece uma festa quando termina (matéria 7). Não resta, portanto, dúvidas de que “relacionamento” e “casamento” não são, no canal do Terra, atividades do mesmo domínio no universo de preocupações da Mulher.

1	Manter o casamento sem prazer vale a pena?
2	Elas dizem sim: muitas garotas ainda escolhem casar cedo
3	17% das pessoas se arrependem do casamento, diz pesquisa.
4	Mulheres mais instruídas sofrem menos violência doméstica, aponta OMS.
5	Flerte fora do casamento ajuda na relação
6	Dez segredos para manter a paixão no casamento
7	Festas de divórcio viram moda nos EUA

Tabela 42: O canal Mulher tematiza o casamento.

4.2.4. A título de síntese

Jovem, magra e sensual – tal o padrão de beleza a que aspira a Mulher dos portais, que constrói a imagem que ela deseja projetar para si, que define seus padrões de comportamento e, finalmente, determina seus interesses. Inevitavelmente, essa Mulher projetada pelos portais deve estar interessada em temas ligados à beleza e à moda, garantes desse padrão, em especial naquela vertente da moda associada à imagem de si – vestuário, cosméticos e serviços estéticos. Porém, um outro conjunto de atributos entra também na construção dessa Mulher: ela é a “rainha do lar”, dedicando parte substancial dos seus interesses a temas ligados à vida doméstica e familiar, em especial a casa e os filhos.

No entanto, há uma curiosa ausência nessa vida familiar: o parceiro. Uma inspeção rápida nas seções ligadas à vida doméstica e familiar do Terra resultou numa única referência à contraparte masculina da vida familiar: a matéria sobre o Feng Shui e seus efeitos na

“energia sexual do casal” (ver [http://mulher.terra.com.br/interna/0,, OI630479-EI4788,00.html](http://mulher.terra.com.br/interna/0,,OI630479-EI4788,00.html), na seção Casa & Cia). A julgar pela seção Mães e filhos (o nome diz tudo), os bebês são enviados pela cegonha ou recolhidos diretamente dos repolhos. Até mesmo a seção Noivas padece da ausência, com uma única matéria onde a presença masculina é obrigatória: “dicas para apimentar a lua de mel” (ver <http://mulher.terra.com.br/noivas/interna/0,,OI524523-EI4883,00.html>). Isso não é privilégio do Terra. O canal Estilo do UOL é mesmo mais radical: os únicos homens presentes em *todo o site* (que não tem uma seção dedicada ao Relacionamento) são *chefs* famosos (na seção de Gastronomia), um famoso jogador de futebol (na seção Moda) e alguns médicos nas seções Saúde e Notícias. Esses dados foram colhidos em inspeção realizada em 27 de novembro de 2006.

É possível que esse deslocamento da figura masculina mesmo da vida familiar e doméstica esteja ligado à depreciação do papel do casamento, assinalada mais atrás, quando discutimos esse outro importante foco de interesse da Mulher dos portais – os relacionamentos. Essa depreciação parece estar associada a uma valorização das relações rápidas e possivelmente efêmeras, enfatizando o namoro e os “affairs”. Isso tudo parece compor uma auto-imagem próxima aos estados narcísicos, compatível com o centro de interesse na beleza. Essa Mulher, mesmo quando se relaciona com os homens, se centra em si mesma. O curioso é que, ainda assim, ela assume para si o papel de “rainha do lar” que lhe reserva a sociedade masculina: ela apenas desloca imaginariamente desse lar os seus súditos homens...

Essa imagem visual e mental que vimos ser construída ao longo dessa análise sumária pode ser corroborada por uma última inspeção, agora às notícias que os canais femininos publicam (em 26 de novembro de 2006). Um sumário dessas notícias pode ser visto na Tabela 43. Note-se que não colhemos dados suficientes por tempo suficiente para utilizá-los na composição de uma tendência. O que é possível interpretar desses poucos dados é apenas a predominância dos temas relacionados àquela imagem da Mulher que inferimos durante a análise.

Conteúdo temático por categoria	Terra	UOL
Vida doméstica e familiar	42,50	10,34
Beleza e moda	5,00	44,83
Relacionamento	37,50	6,90
Outros	15,00	37,93

Tabela 43: Composição das notícias dedicadas à mulher em 21 de novembro de 2006.

4.3. O discurso para todos e para o mesmo

Como vimos na seção 3.1. *Satisfação garantida ou seu dinheiro de volta: o domínio temático dos portais*, a diversidade temática é típica dos portais. Mas não qualquer diversidade – apenas aquela que satisfaz a maior variedade e a maior quantidade possível de usuários. Agradar a todos o maior tempo possível – esse é o mote dos portais. Poderíamos imaginar que esse mote se resolvesse num *discurso da diversidade*, orientado a todos porque falando do ponto de vista de cada um. Não foi isso o que pudemos constatar nas seções anteriores, porém. A quantidade não se resolve aqui na variedade: para atingir a todos, os portais não segmentam a audiência, mas a reduzem à sua expressão mais simples, numa operação análoga aos enunciados da publicidade. Em outras palavras, para se dirigir a Todos, os portais se dirigem a um único – o Consumidor, como vimos na seção 4.2. *A construção da identidade feminina nos portais*.

A diversidade temática aparente dos portais é, portanto, enganosa: ela não se estende aos discursos sobre os diversos temas que aborda. Isto ficou bastante evidente na análise dos canais dedicados à Mulher. De todas as mulheres possíveis, disponíveis para assumir o ponto de vista da enunciação, os canais escolheram uma única – inteiramente imaginária, projetada explicitamente como uma meta, um vir-a-ser mais que um ser. Uma Mulher, não uma mulher: uma abstração de consumo, um objeto de desejo narcísico, um conjunto diminuto de atributos passíveis de ser endereçados pelos produtos de consumo. E ao orientar-se a um (ponto de vista), os portais calam o outro... Trata-se, pois, de um mecanismo de controle do discurso do outro.

No presente capítulo, abordamos dois modos diversos desses mecanismos de controle. Na seção 4.1. *A retórica da verdade no relato jornalístico*, tratamos do compromisso ideológico do relato jornalístico, que assume um conjunto de enunciados possíveis em detrimento de outros, igualmente possíveis mas com outro compromisso. Neste relato, nem todos os discursos sobre os eventos narrados podem se expressar: eles são excluídos. Na seção 4.2. *A construção da identidade feminina nos portais*, abordamos esse controle por um outro ângulo, o da construção de uma identidade. Esta operação é sempre feita por exclusão – de outras identidades. No entanto, ela é “inesperada” num gênero que se pretende inclusivo, orientado a todos e não a alguns.

Assim, se as duas últimas abordagens são representativas dos discursos vigentes nos portais, devemos concluir que eles se constituem num espaço de exclusão – dos discursos do outro. Cabe, então, a pergunta: trata-se aqui de mais um atrator retórico? A resposta é sim! Para atender a maior gama de interesses e preferências, os portais precisam se tornar um espaço confortável, orientado ao prazer e à satisfação. Eles, portanto, precisam excluir desse espaço a inquietação, o perigo inerente aos contra-discursos, à palavra problemática e diferente.

Para retomar o modelo proposta na seção *1.5. O hipertexto e as teias do ciberespaço*, os portais se posicionam nos quadrantes do surfista: ele é um espaço de prazer e diversão, alheio aos tormentos da navegação por mares desconhecidos, com suas tempestades e monstros marinhos – mas também com o prêmio das descobertas.

5. FECHANDO PARA BALANÇO

O livro pode valer pelo muito
que nele não deveu caber.
(Guimarães Rosa Tutaméia)

À luz das discussões que travamos nos últimos capítulos, é possível concluir que fomos capazes de responder afirmativamente à questão (iii) levantada na seção *Um problema e muitas hipóteses* da introdução a esta dissertação. Portanto, a concepção com a qual trabalhamos foi suficiente para não concordarmos com a recusa de Marcuschi em categorizar, sob qualquer hipótese, as chamadas *home pages* entre os gêneros do discurso. É certo que, para isso, não recorremos a apenas uma das tradições de pesquisa sobre gêneros, mas procuramos utilizar, onde cabível, contribuições diferentes na formação da nossa própria concepção. Fomos obrigados também a incorporar ao ferramental conceitual normalmente utilizado para abordar os gêneros as novas características discursivas do hipertexto. Isso foi necessário em virtude do papel central desempenhado pelo discurso “hipermodal” nesse novo campo de atividade humana – a Internet. No entanto, isso não representou uma revisão das teorias dos gêneros do discurso, mas simplesmente uma extensão conceitual para além dos domínios estritos da linguagem “verbal”.

Uma questão derivada da anterior foi igualmente respondida, caso nossa argumentação seja considerada consistente: não é necessário utilizarmos categorias vagas como a de *suporte* para darmos conta dessa classe especial de *sites* Internet. Aqui procuramos argumentar que a categoria não é produtiva simplesmente porque não agrega valor de conhecimento ao problema: ela não ajuda a distinguir eventos e artefatos no ciberespaço. Em favor dessa linha de raciocínio, argumentamos que códigos suportarem outros códigos, isto é, servirem também como “veículos” para outros gêneros e categorias discursivas não é acidental, mas típico da noção de *código* dentro do domínio teórico que empregamos. Nesse sentido, se precisarmos dar ênfase a esse papel “portador” (de outros gêneros e artefatos textuais) dos portais de interesse geral entre tantas categorias de *sites*, julgamos mais apropriado o recurso à categoria de “hipergênero” proposta por Bonini e discutida na seção 2.2. *Canais (marcianos) e veículos de comunicação*. A utilização dessa nova categoria na descrição, se necessária, nos permite conceber os portais como parte de uma classe mais geral de gêneros, da qual fazem parte

também jornais, revistas e outras classes de artefatos comumente referidas como “veículos de comunicação”.

Além disso, se tomarmos como canônico o modelo interpretativo do ciberespaço (e dos hipertextos) discutido sumariamente no capítulo 1. *O ciberespaço: interação à velocidade da luz*, podemos concluir também que os portais de interesse geral podem ser interpretados como uma implementação “degenerada” do hipertexto, esse novo *modo de enunciação digital* (XAVIER, 2002). Trata-se de um hipertexto “fraco”, de baixa densidade, produzido por uma estratégia retórica destinada a atrair e manter o usuário na órbita dos portais, o que é o inverso do suposto no hipertexto “forte” (e canônico), que praticamente convida o usuário a mover-se para longe do hiperdocumento corrente, a integrar no texto presente os inúmeros textos virtuais apontados por ele e que constituem, em conjunto, o seu hipertexto. São essas estratégias de construção orientadas a uma retórica atrativa que dissemos significar, metaforicamente, uma “distorção”, uma “curvatura positiva” (para dentro, em direção ao seu próprio centro) no ciberespaço.

Como vimos de uma forma geral, essas estratégias retóricas só são eficazes se imaginarmos um modelo onde o usuário do portal corresponda a um surfista centrípeta, isto é, um usuário que prefira o prazer à aventura, disposto a divertir-se numa única praia e não a aventurar-se em outros continentes incógnitos (e, portanto, perigosos), cujo comportamento navegacional seja orientado a um centro – o próprio portal. Isto significa que, se concordarmos com os argumentos apresentados ao longo da pesquisa, concluiremos que respondemos afirmativamente à questão (i) posta na introdução: os portais de interesse geral afetam o comportamento navegacional dos usuários. Evidentemente, essa questão não poderia ter uma resposta direta sem uma pesquisa de campo junto aos próprios usuários dos portais. Portanto, nossa resposta foi indireta: ela foi inferida de um modelo retórico imaginado para os portais e discutido especialmente na seção 3.3. *O labirinto do surfista*.

Foi possível inferir também que esse modelo, para ser eficaz, implica no *controle do discurso*, controle que se exerce principalmente pelo controle da *palavra do outro*, pela sua exclusão do universo discursivo. Procuramos ressaltar que o espaço discursivo (hipermodal) dos portais precisa exorcizar a palavra do Outro para ser retoricamente eficaz. E é ao Outro que o hipertexto dá voz, quando constitui um espaço de ligações perigosas (os *links*) entre os textos – e é esse Outro que o hipertexto fraco e pouco denso dos portais cala com sua retórica centrípeta. Embora tenhamos enfatizado esse aspecto principalmente ao tratarmos da interação com os dispositivos de busca, na seção 3.2. *O controle da palavra do outro nos*

dispositivos de busca, esse controle é uma decorrência lógica do modelo de hipertexto centrípeto.

Como assinalamos no capítulo 1. *O ciberespaço: interação à velocidade da luz*, a Internet é uma rede sem centro, avessa a controles, construída de tal modo que o espaço social que constitui é um território eminentemente interativo, onde é muito fácil os usuários se tornarem centros transmissores de informação. E se isso é verdade, ele deve potencialmente ser um espaço de manifestação explícita da palavra do Outro. Ora, se os portais se constroem com base numa estratégia retórica centrípeta, que tem como subproduto a exclusão da palavra do Outro, ele deveria também ser um espaço de baixa interatividade. Isso não foi explorado em nossa pesquisa e, certamente, é uma de suas mais importantes lacunas. Para avaliar essa hipótese seria necessário investigar os *chats*, *blogs* e serviços de *e-mail* fornecidos pelos portais, procurando verificar a incidência de possíveis estratégias de redução e controle da interatividade. Trata-se de um possível desenvolvimento, ainda inteiramente dentro dos horizontes da presente pesquisa.

A pergunta de pesquisa (ii) a propósito dos mecanismos utilizados pelos portais para curvar o ciberespaço não podia, evidentemente, ser respondida de modo completo no curso de uma pesquisa modesta: se isso ocorre, é previsível que essa estratégia se infiltre em todos os domínios do universo discursivo dos portais e não apenas naqueles analisados. Por isso, escolhemos apenas três mecanismos – o domínio temático multifacetado, a interação com os dispositivos de busca e a estrutura navegacional dos portais. Como assinalamos, a escolha recaiu sobre aqueles mecanismos mais típicos do hipertexto tal como implementado nos portais. Naturalmente, tais mecanismos poderiam ser analisados sob inúmeros ângulos, porém, julgamos que a investigação do seu uso como atrator retórico pode ser considerada como um saldo relevante da pesquisa, até porque não se trata de um uso óbvio.

Aqui é preciso enfatizar que procuramos chamar atenção para o fato de que tanto a interação com dispositivos computacionais quanto a estrutura de *links* de um hipertexto podem ser encarados como parte do domínio discursivo dos artefatos culturais do ciberespaço. E isso foi feito simplesmente argumentando em favor do seu uso retórico. Isso nos sugere mais uma fronteira interdisciplinar para a Lingüística contemporânea, particularmente aquela que desemboca nas inúmeras vertentes da(s) Análise(s) do Discurso: trata-se das fronteiras com a Ciência da Computação, particularmente em seus ramos sociológicos.

Por fim, é relevante assinalar que a pergunta de pesquisa (iv), a propósito da imagem de sujeito construída a partir do funcionamento discursivo dos portais, foi a mais incompletamente respondida. A opção por aqueles “estudos de caso”, porém, foi suficiente

para assinalar que se trata de uma investigação promissora por sua fecundidade. Um dos desenvolvimentos possivelmente mais interessantes foi sugerido pela hipótese implícita com a qual trabalhamos na seção 4.2. *A construção da identidade feminina nos portais*: a que implica numa relação promíscua entre o discurso publicitário e o discurso dos portais na constituição da imagem do sujeito. Como essa hipótese por si só mereceria toda uma pesquisa, o melhor que pudemos fazer foi indicar o viés como uma hipótese provisória.

Como dissemos, o modelo interpretativo com o qual trabalhamos prevê um ciberespaço sem centro, que privilegia a liberdade de expressão e, como tal, o discurso do outro em suas múltiplas formas, perigosas e inquietantes, problemáticas e aventureiras. No entanto, os portais de interesse geral são o contrário de tudo isso, como ficou evidente ao longo dos últimos capítulos. E essa categoria de *websites* não é um fenômeno temporário e de curto alcance: as pesquisas de audiência e hábitos de navegação no Brasil (e nos países pesquisados pelo NetRatings) o confirmam. Esse modelo interpretativo é, então, romântico e irrealista?

Talvez. No entanto, devemos considerar que a tecnologia apenas fornece o solo para o plantio – são os seres humanos que semeiam o campo. E eles o fazem com base nos seus próprios hábitos de cultivo anteriores, no caso com base nos velhos meios de comunicação de massa e sua cultura centrada no espetáculo. É que o ciberespaço não é um espaço social inteiramente novo, mas um novo campo de atividade comunicativa nascido de novas tecnologias destinadas aos espaços sociais existentes: se elas trazem novas características a esses espaços, novos problemas e desafios – em outras palavras, se impõe mudanças às relações sociais e aos discursos construídos em torno dessas relações, ainda assim elas se revestem das velhas formas e dos antigos problemas. Continuidade e mudança, revolução e reação, velho e novo – eis o ciberespaço. A questão que se põe torna-se, então: de que lado do ciberespaço queremos estar?

Esta dissertação termina aqui – mas não o percurso que conduziu a ela. Não porque tenhamos estabelecido objetivos ambiciosos demais para uma viagem tão curta. Ao contrário, mesmo contrariando uma natural inclinação para a estrada e não para a chegada, acreditamos ter estabelecido – na maior parte dos casos – objetivos simples e fáceis de serem alcançados, expressos por hipóteses de alcance modesto e limitado. É que o território onde procuramos situar nossa pesquisa é vasto e ainda muito pouco explorado. Era assim quando iniciamos o percurso – e muito pouco mudou no ponto em que nos encontramos. Se esta pesquisa (e seus poucos subprodutos na forma de artigos e comunicações em alguns eventos) tiver servido para despertar algum interesse, mesmo que limitado e superficial, nesse novo campo de atividade humana, julgaremos que todo o esforço foi plenamente recompensado.

REFERÊNCIAS

ARAÚJO, J. C.; BIASI-RODRIGUES, B. (Orgs.). *Interação na Internet*. Rio de Janeiro: Lucerna, 2005.

ASHBY, W. R. *An introduction to cybernetics*. London: Chapman & Hall, 1957.

BAKHTIN, M. *Marxismo e filosofia da linguagem: problemas fundamentais do método sociológico na ciência da linguagem*. Tradução do francês por Michel Lahud e Yara F. Vieira. São Paulo: Hucitec, 1979.

_____. *Problemas da poética de Dostoiévski*. Tradução do russo por Paulo Bezerra. Rio de Janeiro: Forense-Universitária, 1981.

_____. *Estética da criação verbal*. Tradução do russo por Paulo Bezerra. 4. ed. São Paulo: Martins Fontes, 2003.

BERLO, D. K. *O processo da comunicação: introdução à teoria e à prática*. Tradução do inglês de Jorge Arnaldo Fortes. São Paulo: Martins Fontes, 1979.

BERNERS-LEE, T. *Information management: a proposal*. World Wide Web Consortium, 1989. Disponível em: <<http://www.w3.org/history/1989/proposal.html>>. Acesso em: 20 maio 2004.

BERNERS-LEE, T.; CAILLIAU, R.; GROFF, J.; POLLERMAN, B. *World Wide Web: the information universe*. CERN. Disponível em: <http://mirrors.webthing.com/view=Betsie/www.w3.org/History/1992/ENRAP/Article_9202.pdf>. Acesso em: 15 abril 2006.

BONINI, A. Veículo de comunicação e gênero textual: noções conflitantes. *DELTA*. vol. 19, nº 1, 2003. Disponível em <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-44502003000100003&lng=en&nrm=iso>. Acesso em: 18 fevereiro 2005.

CALDAS-COULTHARD, C. R.; VAN LEEUWEN, T. Discurso crítico e gênero no mundo infantil: brinquedos e a representação de atores sociais. *Linguagem em (Dis)curso*, Tubarão, v. 4, n. Especial, 2004. Disponível em: <<http://www3.unisul.br/paginas/ensino/pos/linguagem/revista/revista.htm>>. Acesso em: 17 agosto 2004.

CASTELLS, M. *A galáxia da Internet: reflexões sobre a Internet, os negócios e a sociedade*. Tradução do inglês de Maria Luiza X. De A. Borges. Rio de Janeiro: Jorge Zahar, 2003.

_____. *A sociedade em rede*. Tradução do inglês de Roneide Venâncio Majer. 8.ed. São Paulo: Paz e Terra, 2005.

CGI.BR. Pesquisa sobre o uso das tecnologias da informação e da comunicação no Brasil 2005. Comitê Gestor da Internet no Brasil, 2006. CD-ROM.

_____. *Indicadores da Internet no Brasil*. Disponível em: <<http://www.nic.br/indicadores/>>. Acesso em: 28 julho 2006.

COMO fotografar viagens. Rio de Janeiro: Globo, 1988. (Curso Prático de Fotografia Globo).

DELEUZE, G. *Lógica do sentido*. Tradução do francês de Luiz Roberto Salinas Fortes. São Paulo: Perspectiva, 1974.

DELEUZE, G.; GUATTARI, F. *Mil platôs: capitalismo e esquizofrenia*. Tradução do francês de Aurélio Guerra Neto et all. São Paulo: Editora 34, 1996.

DIONISIO, Â. P.; MACHADO, A. R.; BEZERRA, M. A. (Orgs.). *Gêneros textuais & ensino*. 3.ed. Rio de Janeiro: Lucerna, 2005.

ECO, U. *Tratado geral de semiótica*. Tradução do italiano de Antônio de Pádua Danesi e Gilson Cardoso e Souza. São Paulo: Perspectiva, 1980.

FAIRCLOUGH, N. *Discurso e mudança social*. Tradução do inglês de Izabel Magalhães et all. Brasília: Editora Universidade de Brasília, 2001.

FERNANDES, M. R. *Engenhos de busca distribuídos: uma abordagem visando escalabilidade para crawling e indexação*. 2001. Dissertação (Mestrado em Ciência da Computação) – Centro de Computação, Universidade Federal de Pernambuco, Recife.

FOLHA DE SÃO PAULO. *Manual geral da redação*. São Paulo: Folha de São Paulo, 2001.

FOUCAULT, M. *A arqueologia do saber*. Tradução do francês de Luiz Felipe Baeta Neves. 2.ed. Rio de Janeiro: Forense-Universitária, 1986.

_____. *A ordem do discurso*. Aula inaugural no Collège de France, pronunciada em 2 de dezembro de 1970. Tradução do francês de Laura Fraga de Almeida Sampaio. 10.ed. São Paulo: Loyola, 2004.

FREGE, G. *Lógica e filosofia da linguagem*. São Paulo: Cultrix, 1978.

GOOGLE corporate information: company overview. *Google*. Disponível em: <<http://www.google.com/intl/en/corporate/index.html>>. Acesso em: 29 outubro 2006.

HODGE, R.; KRESS, G. *Social semiotics*. Cambridge: Polity Press, 1988.

HOEY, M. *Textual interaction: an introduction to written discourse analysis*. London: Routledge, 2001.

IPROSPECT. *Search engine user behavior study*. 2006. Disponível em: <http://www.iprospect.com/premiumPDFs/WhitePaper_2006_SearchEngineUserBehavior.pdf>. Acesso em: 27 Agosto 2006.

KLIR, G. J. Uncertainty-based information. In: MELO-PINTO, P.; TEODORESCU, H.; FUKUDA, T. *Systematic organization of information in fuzzy systems*. AmsterdaM: IOS Press, 2003.

KRESS, G.; VAN LEEUWEN, T. *Reading images: the grammar of visual design*. London: Routledge, 1996.

_____. *Multimodal discourse: the modes and media of contemporary communication*. London: Arnold, 2001.

- LEMOS, A. *Cibercultura – tecnologia e vida social na cultura*. Porto Alegre: Sulina, 2002.
- LÉVY, P. *O que é o virtual?* Tradução do francês de Paulo Neves. São Paulo: Ed. 34, 1996.
- _____. *Cibercultura*. Tradução do francês de Irineu da Costa. São Paulo: Ed. 34, 1999.
- MAINGUENEAU, D. *Análise de textos de comunicação*. Tradução do francês de Cecília P. de Souza e Silva e Décio Rocha, 2. ed. São Paulo: Cortez, 2002.
- MARCUSCHI, L. A.; XAVIER, A. C. (Orgs.). *Hipertexto e gêneros digitais: novas formas de construção do sentido*. Rio de Janeiro: Lucerna, 2004.
- MCLUHAN, M. *Os meios de comunicação como extensões do homem*. Tradução do inglês de Décio Pignatari. 5. ed. São Paulo: Cultrix, 1979.
- _____. *Visão, som e fúria*. Tradução de César Bloom. In: LIMA, L. C. (Org.). *Teoria da cultura de massa*. Tradução do inglês de César Bloom. 3. ed. Rio de Janeiro: Paz e Terra, 1982.
- MEURER, J.L.; BONINI, A.; MOTTA-ROTH, D. (Orgs.). *Gêneros: teoria, métodos, debates*. São Paulo: Parábola Editorial, 2005.
- MOLES, A. Doutrinas sobre a comunicação de massa. In: LIMA, L. C. (Org.). *Teoria da cultura de massa*. Tradução do francês de Amrah de Araújo Cardoso e Maria Helena de Oliveira Torres. 3. ed. Rio de Janeiro: Paz e Terra, 1982.
- NELSON, T. H. A file structure for the complex, the changing and the indeterminate. *Proceedings of the ACM 20th National Conference*, Cleveland, 1965. p. 84-100.
- O ESTADO DE SÃO PAULO. *Manual de redação e estilo*. Organizado por E. Martins. São Paulo: Câmara Brasileira do Livro, 1990.
- OLIVEIRA, R. P. Semântica. In: MUSSALIM, F.; BENTES, A. C. (Orgs.). *Introdução à lingüística: domínios e fronteiras*. 3. ed. São Paulo: Cortez, 2003. v.2, p. 17-46.
- ORLANDO, R.A.S. *A comunicação on-line e os portais da web: uma abordagem semiótica*. 2001. 284f. Dissertação (Mestrado em Múltiplos Meios) - Instituto de Artes, Universidade Estadual de Campinas, Campinas.
- OSTASIEDWICZ, W. Uncertainty and unsharpness of information. In: MELO-PINTO, P.; TEODORESCU, H.; FUKUDA, T. *Systematic organization of information in fuzzy systems*. AmsterdaM: IOS Press, 2003.
- PAGE, L.; BRIN, S.; MOTWANI, R.; WINOGRAD, T. *The PageRank citation ranking: bringing order to the web*. Stanford University, 1998. Disponível em: <<http://dbpubs.stanford.edu:8090/pub/1999-66>>. Acesso em: 5 outubro 2005.
- POPPER, K. R. *A lógica da pesquisa científica*. Tradução do inglês de Leônidas Hegenberg e Octanny Silveira da Mota. São Paulo: Cultrix, Universidade de São Paulo, 1975.
- RELATÓRIO analisa uso da Internet no Brasil, Estados Unidos e Espanha. *IBOPE*, Press Releases, 14 julho 2005. Disponível em: <<http://www.ibope.com.br/calandraWeb/>

servlet/CalandraRedirect?temp=5&proj=PortalIBOPE&pub=T&db=caldb&comp= Noticias&docid=4325C3D9D3EDC9998325703E00567FC7>. Acesso em: 11 agosto 2005.

RODRIGUES, R. H. Análise de gêneros do discurso na teoria bakhtiniana: algumas questões teóricas e metodológicas. *Linguagem em (Dis)curso*, Tubarão, v. 4, n. 2, 2004. Disponível em: <<http://www3.unisul.br/paginas/ensino/pos/linguagem/0402/08.htm>>. Acesso em: 15 abril 2005.

SAEED, J. I. *Semantics*. London: Blackwell, 1997.

SHANNON, C. E. A mathematical theory of communication. *The Bell System Technical Journal*. v. 27, 1948. Disponível em: <<http://cm.belllabs.com/cm/ms/what/shannonday/paper.html>>. Acesso em: 4 junho 2005.

SILLER, W.; BUCKLEY, J. J. *Fuzzy expert systems and fuzzy reasoning*. New Jersey: John Wiley & Sons, 2005.

SILVA, T. T. (Org.) *Identidade e diferença: a perspectiva dos estudos culturais*. 2. ed. Petrópolis: Vozes, 2000.

SULLIVAN, D. Search Engine Size Wars V erupts. *Search Engine Watch*, 11 Novembro 2004, Search Ratings & Stats. Disponível em: <<http://blog.searchenginewatch.com/blog/041111-084221>>. Acesso em: 28 Outubro 2006.

_____. Nielsen NetRatings search engine ratings. *Search Engine Watch*, 22 Agosto 2006, Search Ratings & Stats. Disponível em: <<http://searchenginewatch.com/showPagehtml?page=2156451>>. Acesso em: 28 Outubro 2006.

SWALES, J. M. *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press, 1990.

_____. Re-thinking genres: another look at discourse community effects. In: RETHINKING GENRE COLLOQUIUM, 1992, Ottawa. Mimeografado.

TAYLOR, J. R. *Linguistic categorization: prototypes in linguistic theory*. Oxford: Clarendon Press, 1991.

TEODORESCU, H. Information, data, and information aggregation in relation to the user model. In: MELO-PINTO, P.; TEODORESCU, H.; FUKUDA, T. *Systematic organization of information in fuzzy systems*. AmsterdaM: IOS Press, 2003.

THE FREE software definition. *GNU project – the free software foundation*. Disponível em: <<http://www.gnu.org/philosophy/free-sw.html>>. Acesso em: 25 julho 2006.

TOP 10 – domínios mais visitados em milhares de usuários únicos. *IBOPE*. Disponível em: <http://www.ibope.com.br/eratings/audiencia/internet/audiencia_internet_nacional_set03_no>. Acesso em: 26 maio 2004.

UOL, o melhor conteúdo da Internet. *UOL*. Disponível em <<http://sobre.uol.com.br>>. Acesso em: 2 fevereiro 2006.

WEAVER, W. Contribuições recentes à teoria matemática de comunicações. In: SHANNON, C. E.; WEAVER, W. *A teoria matemática da comunicação*. Tradução do inglês de Orlando Agueda. São Paulo, Rio de Janeiro: Difel, 1975. p.3-29.

WEB usage data. *Nielsen//NetRatings*. Disponível em: <http://www.netratings.com/news.jsp?section=dat_to>. Acesso em: 28 julho 2006.

WHATIS.COM. *Spam*. Disponível em: <http://searchmobilecomputing.techtarget.com/sDefinition/0,290660,sid40_gci213031,00.html>. Acesso em: 19 Junho 2005.

WIENER, N. *Cibernética e sociedade*. O uso humano dos seres humanos. Tradução do inglês de José Paulo Paes. 9. ed. São Paulo: Cultrix, 1993.

WITTGENSTEIN, L. *Investigações filosóficas*. São Paulo: Cultrix, 1979. (Os Pensadores).

WOODWARD, K. Identidade e diferença: uma introdução teórica e conceitual. In: SILVA, T. T. (Org.). *Identidade e diferença: a perspectiva dos estudos culturais*. 2. ed. Petrópolis: Vozes, 2000. p. 7-72.

XAVIER, A. C. *O hipertexto na sociedade da informação: a constituição do modo de enunciação digital*. 2002. 218f. Tese (Doutorado em Lingüística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas.

FRANÇA, V.O. *Linguagem hipertextual: a parole na fronteira digital*. 2003. Dissertação. (Mestrado em Letras) – Instituto de Letras, Universidade do Estado do Rio de Janeiro, Rio de Janeiro.

ZADEH, L. A. Toward a perception-based theory of probabilistic reasoning. In: MELO-PINTO, P.; TEODORESCU, H.; FUKUDA, T. *Systematic organization of information in fuzzy systems*. Amsterdam: IOS Press, 2003.