



Rodrigo Esteves de Lima Lopes

SOCIOSSEMIÓTICA DA PRODUÇÃO AUDIOVISUAL:

Uma Proposta Metodológica para Análise Multimodal da
Comunicação em Vídeo

Campinas

2012



Universidade Estadual de Campinas
Instituto de Estudos da Linguagem

Rodrigo Esteves de Lima Lopes

SOCIOSSEMIÓTICA DA PRODUÇÃO AUDIOVISUAL:

Uma Proposta Metodológica para Análise Multimodal da
Comunicação em Vídeo

Orientadora: Prof.^a Dr.^a Denise Bértoli Braga

Tese apresentada ao Instituto de Estudos da Linguagem
da Universidade Estadual de Campinas para obtenção
do Título de Doutor em Linguística Aplicada na área de
Linguagem e Tecnologias

Campinas

2012

FICHA CATALOGRÁFICA ELABORADA POR
TERESINHA DE JESUS JACINTHO – CRB8/6879 - BIBLIOTECA DO INSTITUTO DE
ESTUDOS DA LINGUAGEM - UNICAMP

L628s Lima-Lopes, Rodrigo Esteves de, 1973-
Sociossemiótica da Produção Audiovisual: Uma
Proposta Metodológica para Análise Multimodal da
Comunicação em Vídeo / Rodrigo Esteves de Lima Lopes.
-- Campinas, SP : [s.n.], 2012.

Orientador : Denise Bértoli Braga.
Tese (doutorado) - Universidade Estadual de
Campinas, Instituto de Estudos da Linguagem.

1. Multimodalidade. 2. Gramática sistêmica. 3.
Comunicação audiovisual. 4. Pesquisa qualitativa. 5.
Semiótica. I. Braga, Denise Bértoli, 1953-. II. Universidade
Estadual de Campinas. Instituto de Estudos da Linguagem.
III. Título.

Informações para Biblioteca Digital

Título em inglês: Socio-Semiotics of Audiovisual Production: A
Methodological Proposal for the Analysis of Multimodal Video Communication.

Palavras-chave em inglês:

Multimodality

Systemic Grammar

Audiovisual communication

Qualitative research

Semiotics

Área de concentração: Linguagem e Tecnologias.

Titulação: Doutor em Linguística Aplicada.

Banca examinadora:

Denise Bértoli Braga [Orientador]

Rosinda de Castro Guerra Ramos

Edson do Prado Pfutzenreuter

Viviane Maria Heberle

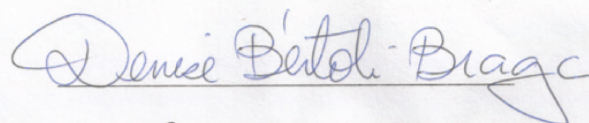
Orlando Vian Júnior

Data da defesa: 09-11-2012.

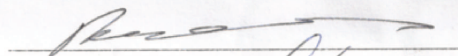
Programa de Pós-Graduação: Linguística Aplicada.

BANCA EXAMINADORA:

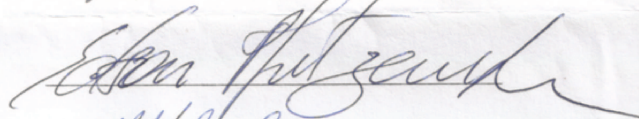
Denise Bértoli Braga



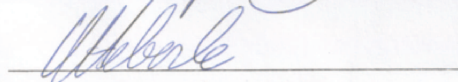
Rosinda de Castro Guerra Ramos



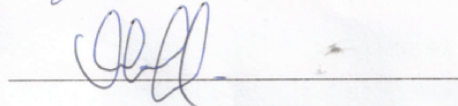
Edson do Prado Pfutzenreuter



Viviane Maria Heberle



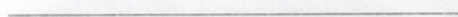
Orlando Vian Júnior



Leila Barbara



Roxane Helena Rodrigues Rojo



Adail Sebastião Rodrigues Júnior



IEL/UNICAMP
2012

Aos meus pais:

Derly Esteves dos Santos Lopes

José Londrino de Lima Lopes (*in memoriam*)

Por me ensinarem a ser quem sou

Às minhas professoras:

Beatriz Berrini

Leila Barbara

Rosinda de Castro Guerra Ramos

Denise Bértoli Braga

Por me ensinarem muito mais do que fazer ciência

Agradecimentos

- À CAPES, pelo financiamento desta pesquisa.
- À minha orientadora, Prof.^a Dr.^a Denise Bértoli Braga, por entender as necessidades deste trabalho e nele acreditar, pelo seu carinho e dedicação durante o processo de orientação.
- À Prof.^a Dr.^a Rosinda de Castro Guerra Ramos, pela dedicação a esta pesquisa, pelo carinho materno, por ensinar-me muito sobre o mundo e sobre mim, pela generosidade, por ser amiga e companheira e pelas sugestões importantes durante os exames de qualificação e de defesa pública.
- Ao Prof. Dr. Marcelo El Khouri Buzato, pelas sugestões importantes durante o exame de qualificação.
- Ao Prof. Dr. Edson do Prado Pfitzenreuter, pelas sugestões importantes durante o exame de qualificação e de defesa pública, pelas profícuas discussões e intercâmbio estabelecidos.
- À Prof.^a Dr.^a Roxane Rojo, por suas contribuições no exame de qualificação fora de área.
- Ao Prof. Dr. Orlando Vian Jr. e à Prof.^a Dr.^a Viviane Heberle por aceitarem participar da banca de defesa pública, pelas sugestões importantes.
- Ao Prof. Dr. José Luiz Meurer (*in memoriam*), pelas profícuas discussões sobre coesão visual e Gramática Sistêmico-Funcional, pelo apoio teórico e pessoal sempre presentes.
- À Prof.^a Dr.^a Leila Barbara, por me iniciar nos caminhos da GSF, pelo apoio acadêmico e pessoal que se traduzem em carinho e amizade.
- To Professor Peter Fries, for his generosity and important contributions to the discussion on visual cohesion and visual Theme/Rheme.
- To Professor Geoff Thompson, for the interesting discussion on multimodal and visual constituency.
- Aos colegas Marco Vale, Cláudio Arantes e José Eugenio, a quem agradeço por acreditarem em meu trabalho.
- A Carol Bottosso, Ana Flora Schlindwein, Heitor Gribl, Viviane Curto e Sônia Gazeta, pelo apoio e ajuda, pelo incentivo e por abrirem caminhos. Pela paciência em nossas sessões de "terapia acadêmica".
- À Shirlei, pela paciência e carinho, além das belas ilustrações.
- Aos amigos Samuel Gomes da Silva, Hélio Brito Souza, Alexandre Maduenho, Paulo Rowlands, Zeca Rodrigues, Cesar Sampaio, Ricardo Bergmann, Cida Lamas, Marcus Vinícius Pedrozo, Emerson Gomes Galhardo, Lucas Barão e Gláucia Carvalho pela amizade e ajuda incondicionais durante tantos anos. Por saberem compreender minhas necessidades e idiosincrasias. Por serem os amigos que sempre admirei por toda vida.
- À Isabel Pacheco Bernini, pela amizade, pelo socorro de última hora e pela ajuda imprescindível.
- À Carol Ventura, pela amizade e paciência quase infinitas, pois qualquer processo de ensino-aprendizagem nasce da interação com nosso par mais competente.

As lições da infância
Desaprendidas na idade madura.
Já não quero palavras
Nem delas careço.
Tenho todos os elementos
Ao alcance do braço.

Carlos Drummond de Andrade

Tenho exercido vários ofícios, esqueci
todos, e assim posso mover-me sem
nenhum constrangimento.

Graciliano Ramos

Ora, meu espírito apresenta uma
peculiaridade, que decerto é uma
inaptidão: a dificuldade de ser por mim
fixado duas vezes no mesmo objeto.

Claude Levi-Strauss

Resumo

Nesta tese tenho por objetivo propor uma metodologia de análise para comunicação multimodal em produções audiovisuais. Sua base teórica está na análise multimodal (KRESS, 1993, 2003, 2005, 2006, 2010; KRESS; VAN LEEUWEN, 1996, 2001, 2002; Lemke, 1998, 2002, 2004, 2006a, b, c) e parte do pressuposto que qualquer ato comunicacional ocorre graças a uma série de convenções de constituição atreladas à situação de comunicação. Essas convenções constituem os diversos modos de linguagem, os quais em interação permitem que os participantes de uma interação construam seu significado. Por conseguinte, acredito que todos os modos de linguagem presentes no corpus desta tese — o que incluiria, por exemplo, a linguagem visual, gestual ou qualquer outro modo de linguagem — têm seus significados expressos por meio de escolhas interrelacionadas e de importância equivalente. O corpus desta pesquisa é composto por 12 vídeos da campanha "Get a Mac" (Compre um Mac), exibida na televisão dos EUA e disponibilizada em sítios de internet. Tal campanha teve desdobramentos no Reino Unido e no Japão, locais onde foram produzidos alguns vídeos publicitários inéditos e com características próprias. Início essa tese com uma reflexão sobre os conceitos epistêmico-teóricos que dão origem à abordagem multimodal, para, posteriormente, discuti-la de forma crítica. Aqui pretendo estabelecer alguns pressupostos que estão na base de interpretação da discussão multimodal, e que serão constantemente retomados no processo de discussão teórica, análise e discussão dos dados. Segue-se então uma discussão sobre aplicação da Metafunção Organizacional (ou Textual) em peças audiovisuais, de forma a complementar a literatura já existente com discussões a respeito de como elementos próprios da linguagem vídeo-fílmica, além de modos de linguagem corporificados, como gestos e olhares, também poderiam ser considerados nessa equação. Passo então para uma discussão metodológica que busca não apenas apresentar o corpus de estudo deste trabalho, discutir quais foram os critérios de coleta, análise e transcrição de dados, trazer algumas considerações importantes sobre a linguagem vídeo-fílmica, bem como propor um modelo de análise para esses textos, objeto de estudo desta tese. Parto do princípio de instanciação como elemento norteador do trabalho do analista, procurando fazer um levantamento sistemático das escolhas em cada um dos modos de linguagem e discutindo como cada um deles contribui para o amálgama de signos que é o significado; observa-se, então, que a instanciação do significado faz com que as diferentes escolhas sejam condicionadas. Dessa forma, aplicarei esse modelo no corpus de pesquisa. Exploro como o conceito de constituição pode ser utilizado na análise de textos multimodais, desenvolvendo uma reflexão sobre a forma como os diferentes constituintes dos diversos modos de linguagem podem ser analisados em conjunto nos vídeos publicitários que compõem o corpus desta pesquisa. Utilizo os resultados desta parte inicial de forma a ampliar a discussão tendo em vista diversas perspectivas, o nível do conteúdo e o discurso, o design como encontro entre o conteúdo e a produção e suas escolhas nos diferentes contextos de produção. Por fim, estabeleço uma reflexão ligada à construção do processo de coesão e fluxo de informações em textos vídeo-fílmicos, com o objetivo de entender como se constrói um sistema de referências internas que permita conectar aquilo que já foi exibido com aquilo que está aparecendo naquele exato momento na tela.

Palavras-chave: Multimodalidade; Gramática sistêmica; Comunicação audiovisual; Pesquisa qualitativa; Semiótica.

Abstract

In this thesis I aim at proposing a method of analysis for multimodal communication in audio-visual productions. The theoretical framework is based on multimodal analysis (KRESS, 1993, 2003, 2005, 2006, 2010; KRESS; VAN LEEUWEN, 1996, 2001, 2002; Lemke, 1998, 2002, 2004, 2006a, b, c), which assumes that any interaction takes place due to a series of conventions tied up to the context of communication. These conventions are the different modes of language, which allow participants of an interaction to build their meaning. Therefore, I consider that all modes of language present in the corpus analysed here — which could include, for example, the visual language, gestures or any other mode of language — have their meanings expressed through choices interrelated and of equal importance. The corpus of this research consists of 12 videos from the campaign "Get a Mac", shown in the US television and made available by Internet websites. As we shall see, this campaign was also launched in the UK and Japan, where some new videos were produced. In order to discuss my theoretical framework critically, this thesis begins with a discussion of the epistemic-theoretical concepts that give rise to a multimodal approach. Here I intend to establish some assumptions that will underlie my multimodal interpretation of the corpus, and these assumptions are going to be constantly put on back again in the theoretical discussion and data analysis. This will be followed by a discussion on realization of the Organizational (Textual) Metafunction in audio-visual pieces. This is meant to complement the existing literature with reflections on how specific elements of audio-visual language, as well as embodied modes of language such as gestures and looks could also be considered in this equation. Then, I begin a methodological discussion that seeks not only to present the corpus of the present study, my criteria for data collection, analysis and transcription, but also to propose a model to analyse these texts. I also explore some important concepts regarding audio-visual language. As a result, I assume the instantiation as a guide to the analysis and try to make a systematic survey of the choices in each of the modes of language in the videos. It was essential to understand how each mode of language contributes to the amalgamation of signs that results in meaning, which regulates the different choices. After this, I apply this model in the research corpus in order to explore how the concept of constitution can be used in the analysis of video-filmic texts and discuss how the different constituents of the modes of language can be analysed together. I use the results of this initial part to broaden the discussion on some perspectives, the level of content and discourse, the design as an encounter between content and expression and the choices in different contexts of production. Finally, I try to reflect upon cohesion and information flow in audio-visual texts. My goal in this moment was to understand how the instantiation of video texts creates a system of internal references that allows the viewer to connect what has already been shown with what is being shown at that very moment on the screen.

Keywords: Multimodality; Systemic grammar; Audio-visual communication; Qualitative research; Semiotics.

Lista de Quadros

1	Composição do corpus principal	112
2	Relação entre os elementos atitudinais e composicionais em <i>Networking</i> (V-04)	152
3	Relação entre os elementos atitudinais e composicionais em <i>Boxer</i> (V-09) . .	154
4	Relação entre os elementos atitudinais e composicionais em <i>Security</i> (V-06) .	157
5	Quebras verbais na apresentação	186
6	Comparação em Introdução Problema-1	190
7	Comparação em Introdução Problema-2	193
8	Vídeos e a estratégia estrutural no fluxo de informações	225

Lista de Figuras

1	Níveis ou <i>Strata</i> na GSF	31
2	<i>Strata</i> e sua tradução para o Corpus	32
3	O gênero em relação ao registro e à Língua	37
4	Fluxo de informação em fotografias esportivas	92
5	Propaganda de notebook Sony	93
6	João da Mata no pódio da São Silvestre de 1983	94
7	Sequência de animação em Flash — anunciando produto Sony	95
8	Sequência de <i>O lenhador</i>	98
9	Frame isolado de <i>O lenhador</i>	100
10	Sequência de <i>Life on Mars</i>	102
11	Sequência de <i>Ônibus 174</i>	103
12	Sequência de <i>Disque M para Matar</i>	105
13	Sequência de comercial em <i>Flashback</i>	106
14	Frame ilustrativo do vídeo <i>Restart</i> (v4)	116
15	Planos em vídeo, cinema e fotografia	121
16	Movimentos de câmera	122
17	Janela principal do software ELAN 3.9	131

18	Sistemas de significado e linguagem videográfica	145
19	Elementos atitudinal e composicional em <i>Networking</i> (V-04)	150
20	Componentes atitudinal e composicional em <i>Boxer</i> (V-09)	153
21	Frame em <i>Security</i> (V-06)	155
22	Comparação roupas-Mac x PC em alguns vídeos do corpus	165
23	Constituição da metáfora de domínio computacional	169
24	Metáforas em dois vídeos do corpus	170
25	Solidariedade em <i>Restarting</i> (V-03)	176
26	Solidariedade em <i>Viruses</i> (V-01)	177
27	Planos e composição em <i>Better Results</i> (V-05)	179
28	Intertextualidade entre "Get a Mac" e o cinema	180
29	Intertextualidade audiovisual em duas campanhas Apple	182
30	Vídeo da campanha <i>Switch</i>	183
31	Divisão temporal nos vídeos	184
32	Exemplos de quebra de padrões na apresentação	187
33	Vídeos apenas com quebra do padrão visual	188
34	Composições visuais em "Introdução do Problema"	189
35	Elementos externos em "Introdução do Problema"	191
36	Cortes e olhar em <i>Vista Choose</i> (V-07)	194
37	Frames de Solução do Problema em <i>Out of the box</i> (V-02)	195
38	Solução do problema em <i>Out of the box</i> (V-02) e <i>Meant for Work</i> (V-10) . .	197
39	Comparação entre os quadros: EUA, Japão e Reino Unido	200
40	Office vs. Home — campanha Reino Unido	201
41	Steps — campanha japonesa	203
42	Comparação entre 3 vídeos — Japão/Reino Unido/EUA	208
43	Repetição em <i>key-frames</i> em <i>Surgery</i> (V-08)	215

44	Repetição em <i>key-frames</i> em <i>Vista Choose</i> (V-07)	220
45	Planos e preenchimento	222
46	Planos para preenchimento em <i>Surgery</i> (V-08)	223
47	Componentes estruturais em <i>Meant for Work</i> (V-10)	226
48	Componentes estruturais em <i>Viruses</i> (V-01)	228
49	Baixo uso de componentes estruturais em <i>Fat HD</i> (V-11)	229
50	Baixo uso de componentes estruturais em <i>Breakthrough</i> (V-12)	230
51	Uso de constituintes estruturais exclusivos em <i>Networking</i> (V-04)	231
52	Gestos e o fluxo de informação em <i>Out of the box</i> (V-02)	234
53	Olhares, gestos e sua relação com o fluxo de informação	236
54	Fluxo de informações em <i>Security</i> (V-06)	240

Sumário

Introdução	1
Motivações: o encontro entre o pessoal e as transformações sociais	1
O vídeo como objeto de estudo	9
"Get a Mac"questões gerais sobre a campanha	14
A perspectiva de linguagem adotada e estrutura do trabalho	18
1 Teoria da multimodalidade	23
Introdução	23
1.1 Multimodalidade e suas origens na Sociossemiótica Hallidiana	25
1.1.1 GSF: os conceitos de partida	25
1.1.2 Significado e forma: uma gramática em níveis	27
1.1.3 Função, contexto e metafunção	34
1.2 Transformações no contexto social e linguístico: em busca de uma teoria So-	
ciossemiótica Multimodal	41
1.2.1 Por uma teoria abrangente da comunicação	41
1.2.1.1 A imagem e a tecnologia	43
1.2.1.2 Produção de significado	49
1.2.2 Refletindo sobre a articulação de significados em textos multimodais .	56
1.2.2.1 Discurso	57

1.2.2.2	Design	60
1.2.2.3	Produção	67
1.2.2.4	Distribuição	70
1.2.3	Analisando diversos modos simultaneamente	73
1.3	Resultados parciais	77
2	Metafunção organizacional e multimodalidade	85
	Introdução	85
2.1	Conceitos iniciais	86
2.2	Metafunção organizacional e a linguagem visual	92
2.2.1	A imagem estática	92
2.2.2	A linguagem videográfica	96
2.3	Resultados parciais	107
3	Definição do corpus, temática da campanha e elementos de linguagem videográfica	109
	Introdução	109
3.1	Descrição do corpus	110
3.1.1	Composição do Corpus e procedimentos de coleta	110
3.1.2	Temática da campanha e dos vídeos do corpus principal	114
3.2	Noções básicas de linguagem vídeo-fílmica	118
4	Reflexões sobre a transcrição multimodal	127
	Introdução	127
4.1	Estabelecendo procedimentos de transcrição	129
4.2	Constituição: elementos criadores de significado	138
4.3	Construção do significado: operando vários modos de linguagem	149
4.4	Algoritmo final de análise	158
4.5	Resultados parciais	160

5 O vídeo enquanto sistema de construção de significados: explorando modos de linguagem	163
Introdução	163
5.1 Constituição do discurso: tradução e metáfora	164
5.2 Organização e padrões em "Get a Mac"	179
5.3 Escolhas e manutenção do significado em diferentes ambientes de produção e veiculação	199
5.4 Resultados parciais	209
6 Coesão e fluxo de informações no vídeo	213
Introdução	213
6.1 Repetição e preenchimento de vazios	214
6.2 Fluxo de informação e construção do significado	224
6.2.1 Componentes estruturais	224
6.2.2 Componentes atitudinais	234
6.3 Resultados parciais	242
7 Considerações Finais	245
Referências Bibliográficas	255
A. Exemplo de transcrição manual	265

Introdução

The space of semiotic meanings is
multiplicative: it expands
combinatorially in possibilities.

Theo A. van Leeuwen

Motivações: o encontro entre o pessoal e as transformações sociais

Esta tese analisa um conjunto de vídeos promocionais da Apple Computers, Inc. Esses vídeos, distribuídos pela internet e exibidos nas lojas Apple e canais de TV dos EUA, fazem parte de uma campanha intitulada “Get a Mac”, realizada entre 2007 e 2009. Aqui utilizarei como base teórica a abordagem multimodal (KRESS e VAN LEEUWEN, 1996; LEMKE, 1998; KRESS; VAN LEEUWEN, 2001; LEMKE, 2002; NORRIS, 2002; BALDRY, 2004; LEMKE, 2004; NORRIS, 2004; LEMKE, 2006c; para citar apenas alguns), de forma a compreender como os diferentes modos de linguagem interagem na construção desses textos vídeo-fílmicos. Meu objetivo principal é fornecer insumo para uma discussão teórico-metodológica que auxilie no desenvolvimento de uma metodologia de interpretação de textos multimodais em vídeo, tentando compreender como a construção dos significados perpassa

os diferentes modos de linguagem no que tange tanto à seleção dos elementos em sua gramática, como à combinatória na instanciamento do significado. Se, por um lado, estou consciente de que um estudo dessa natureza não pode pressupor um sistema de análise definitivo ou fechado, qualquer tentativa de sistematização deve se pautar pela consciência de sua capacidade apenas superficial de explicação dos fenômenos; por outro devo observar que isso é uma consequência do próprio fenômeno estudado: dinâmico, o vídeo, como veremos a seguir, tem ganho um grande destaque na cultura popular contemporânea, chegando a ocupar um lugar muito próximo da fotografia.

Essa reflexão inicial acaba por ter dois desdobramentos; o primeiro estaria relacionado a uma discussão sobre o poder explicativo da análise multimodal, e o segundo à criação de um sistema de análise que tome o significado como elemento central da construção da linguagem. Ao se pensar na natureza dos textos que são analisados nesta tese, poder-se-á advogar que as escolhas dentro dos modos de linguagem que os constituem estão longe de ser aleatórias, elas congregam elementos com valor cultural na sociedade ocidental — como olhar, dar as mãos, aproximar-se, trocar risadas — e de valor técnico-informacional — como movimentos de câmera, enquadramentos, escolhas de edição — de forma a vesti-los de um significado específico, funcionalmente determinado pelo objetivo comunicativo de cada um dos vídeos analisados.

O corpus desta pesquisa (ver capítulo 3 na página 109) é composto por 12 vídeos da campanha "Get a Mac"(Compre um Mac). Como veremos, essa campanha teve desdobramentos no Reino Unido e no Japão, locais onde foram produzidos alguns vídeos publicitários inéditos. Alguns desses vídeos também serão agregados a esta tese para análise contrastiva durante o capítulo 5 na seção 1.2.2.3, de forma a constatar se há a manutenção da identidade da campanha nos diferentes contextos de produção.

Para nortear o leitor a respeito de meu interesse pelo tema tratado nesta tese, cabe explicitar meu percurso histórico contextualizando-o de forma a enriquecer as discussões teóricas que busco realizar. Meu primeiro contato com a linguagem audiovisual se deu ainda

na infância, vivida na extinta Cinelândia, em São Paulo. Essa área que compreende a avenida São João, parte da Duque de Caxias, largo do Paissandu, avenida Ipiranga e parte da rua Conselheiro Crispiniano, abrigou um grande número de cinemas na capital entre as décadas de 50 (seu auge) e os anos 80 (sua decadência). A vivência nesse espaço de exposição privilegiada fez-me criar um verdadeiro fascínio pelo cinema, criando um estreito laço entre mim e essa arte. Boa parte dessa exposição foi o que me levou, em 1988, a ingressar como profissional na Rádio e Televisão Cultura de São Paulo, Centro Paulista de Rádio e TV Educativas, local onde tive a oportunidade de trabalhar no departamento de operações até 1998, ano em que me inscrevi no programa de mestrado em Linguística Aplicada da PUCSP. Foi essa experiência profissional e pessoal que me permitiu atuar entre 2002 e 2008 como professor de Linguagem de Vídeo no curso de Comunicação e Multimeios da PUCSP e hoje trabalhar como professor de Novas Tecnologias em Rádio e Televisão e Teoria da Comunicação, além de ser pesquisador, na Faculdade Cásper Líbero.

Se no âmbito pessoal a relação com a linguagem audiovisual (ou vídeo-fílmica, aqui utilizo ambos como sinônimos) é estética, profissionalmente ela tem, pelo menos, duas cores distintas. Como técnico na TV Cultura, atuei em diversas áreas relacionadas à captura e edição de áudio e vídeo, com especial ênfase na gravação de reportagens externas. Já em minha atuação acadêmica, minha relação com esta linguagem está principalmente em uma reflexão crítica sobre sua produção, tentando estabelecer mecanismos de compreensão em seu nível estético, ideológico e estrutural, preocupando-me sobremaneira em levar a discussão sobre essa linguagem além da simples técnica. Surge daí meu interesse pela ampliação dos estudos de análise do discurso¹ em outros modos de linguagem além do verbal.²

Nossa sociedade vem sofrendo profundas transformações, advindas de metamorfoses dentro do sistema econômico em que estamos inseridos (GEE; HULL, 1996a; b), sendo que elas

¹Para uma definição do termo discurso e de seu papel nessa abordagem, ver seção 1.2.2.1, na página 57, e na seção 5.1, na página 164.

²Neste trabalho, utilizo a expressão linguagem verbal para me referir ao componente linguístico de uma interação, tanto em seu modo escrito como falado. Esse termo está em estado de contraponto ao termo linguagem não verbal, aqui utilizado para compreender todos os demais modos de linguagem.

trazem diferentes consequências para o mundo do trabalho, gerando um indivíduo que se forma e se identifica em uma cultura digital de rede (SANTAELLA, 2004; 2005). Se, por um lado, durante o sistema capitalista tradicional, o indivíduo era responsável por uma parte do trabalho, estando, portanto, alienado em relação à totalidade dos processos que resultariam na produção de um bem, por outro seu desligamento da função se dava no momento em que seu período de trabalho ocorria. Todavia, no modelo atual, estamos integrados à constante presença digital, trabalhando, nos relacionando e interagindo em tempo real, sendo que nossa presença e nossa realidade são expandidas em uma série de espaços virtuais, em um contexto no qual a esfera pública e privada começam a se misturar. Como um resultado dessa realidade, todas nossas atividades são expandidas, criam-se extensões corpóreas que garantem nossa entrega à nova realidade digital; nosso nível de comprometimento e nossa capacidade como indivíduos passam a ser medidos pela nossa presença na arena virtual, pela forma como gerenciamos informações.

Um exemplo disso são os telefones celulares que estão cada vez mais presentes em nosso cotidiano. Dados da ANATEL³ mostram que a densidade do parque de telefones celulares no Brasil chegou a 93,80 aparelhos para cada 100 habitantes, com um total de 5,1% desses telefones acessando as redes de dados 3G. No âmbito da telefonia fixa, cresce a utilização de serviços de VoIP (voz sobre IP), que oferecem ligações locais e internacionais por meio da internet a partir de preços bastante competitivos.⁴ Cada vez mais milhares de usuários usam recursos digitais para criação de seu espaço social, vemos a proliferação de redes sociais (como Facebook e MySpace), sistemas de mensagens instantâneas (como MSN, Gtalk e Skype); bancos, concessionárias públicas, universidades e todos os segmentos da sociedade virtualizam, cada vez mais, sua relação com o público — o varejo de imóveis e carros tem na internet e sites de anúncio um grande aliado comercial. O uso crescente de correio eletrônico

³Base do serviço móvel chega a 185 milhões no Brasil. Retirado de <http://oglobo.globo.com/pais/>. Acessado em: 16 de julho de 2010.

⁴Importante ressaltar que muitas dessas empresas de VoIP também oferecem serviços de vídeo-chamada e chat já incluídas dentro dos programas por ela distribuídos. Exemplos desse tipo de serviço são o Skype <http://skype.com> e o Google Voice <http://www.google.com/googlevoice/about.html>

já substitui boa parte das correspondências em papel, de acordo com a última reunião da União Postal Universal, com sede em Berna, de 2006 para cá apenas 52% das atividades do correio brasileiro estão relacionadas ao envio de cartas. O uso do correio convencional possui tendência de aumento apenas em alguns países da América Latina e do antigo bloco soviético, localidades já conhecidas pela dificuldade de acesso à tecnologia.⁵ Essa realidade tecnológica tem transformado a maneira como interagimos, como nos comunicamos e tem mostrado que a dinâmica da interação social em muito transcendeu os modos e suportes comunicacionais mais tradicionais. Cria-se uma organização social que prima pelo potencial existencial das diferentes informações e interações realizadas em rede na qual a presença no planisfério virtual é tão importante, senão mais, do que a física (LÉVY, 1996). Como um resultado prático dessa nova realidade, nas últimas duas décadas, as fronteiras entre os diferentes suportes comunicacionais têm se transformado em algo tênue, o que se dá, principalmente, devido ao grande fluxo de informações e ao surgimento de novos meios para sua propagação. Boa parte dessa nova realidade deve-se ao extensivo uso de computadores e mídias locativas em nosso cotidiano, com especial ênfase para a sua integração em rede. Isso gera uma nova realidade, que tanto influencia o mundo do trabalho (GEE; HULL, 1996a; b), do ensino (BAWDEN, 2008) e dos meios de comunicação (NEGROPONTE, 2002; STRAUBHAAR; LAROSE, 2004; CANEVACCI, 2009), gerando novos paradigmas de interação e produção textual. Neste trabalho, uso o termo texto de uma forma um pouco mais ampla que a maioria dos estudos em Linguística Aplicada, considerando como tal qualquer ato comunicativo, seja ele sonoro, imagético, performático, atitudinal e de vestuário. Ou seja, não apenas manifestações relacionadas à linguagem verbal, mas também construções significativas que incluam outros modos de linguagem, o que, por vezes, poderia prescindir da fala ou da escrita. Considero, portanto, como texto uma unidade comunicacional que constrói o seu significado através de múltiplos modos de linguagem, sempre em interação com seus leitores. Tal unidade se apoia em seu componente sociocomunicativo — algo que parece nortear todas as produções —,

⁵Dados obtidos de Swiss Info. <http://swissinfo.ch> em 20 maio de 2010.

semântico — uma vez que a criação de significados, por parte do produtor e do leitor, é um pressuposto de qualquer ato comunicacional — e formal — uma vez que todo modo de linguagem possui alguma forma de organização da mensagem.

De fato, essa realidade serviu para revelar apenas uma questão antes pouco observada pelos linguistas: de que todo texto possui diversas instâncias envolvidas na construção de seu significado, sendo a combinação funcional entre diferentes sistemas sociossemióticos o que garante a comunicação humana. Esse fato é importante por ampliar o escopo da Linguística Aplicada (doravante LA). Como área essencialmente interdisciplinar, a LA, em minha opinião, estaria apta a expandir seu escopo de análise e considerar outros modos de linguagem, os quais estenderiam, em muito, a compreensão dos fenômenos comunicacionais. Em uma época que os meios de propagação da informação são tão importantes quanto a informação em si (MCLUHAN, 1994), seria ingênuo pensar que a linguagem possa ser estudada de forma dissociada de outros modos e das mídias em que é utilizada. Por conseguinte, talvez seria ingênuo pensar que tal transformação no construto social e econômico não atingiria as práticas languageiras, alterando padrões de produção, normas sociais de comunicação e a construção de cadeias significativas e simbólicas. Se seguirmos por essa seara, teremos uma arena povoada de possibilidades de significados, todos eles em infinita probabilidade de combinação. Isso nos levaria a pensar que a análise de qualquer semiose em isolamento seria, pelo menos, limitadora, graças à exclusão de significados implicada.

Em primeiro lugar, é importante observar que a LA é uma ciência com especial aptidão para estudar esses fenômenos, o que se dá, principalmente, por sua capacidade de agregar preceitos teóricos advindos de diversas ciências. Boa parte dessa flexibilidade se origina em sua preocupação com um objeto real e contextualizado: a linguagem verbal em suas manifestações em ambientes sociais. Razão que pode explicar o fato da LA ter desenvolvido diferentes abordagens, todas elas de caráter multidisciplinar. No ensino-aprendizagem de línguas, subárea de estudo que se confunde com o próprio nascimento da LA, perceber-se-ão influências teóricas vindas da pedagogia, psicologia, filosofia e linguística teórica (PASCHOAL; CELANI,

1992). Outro exemplo seria a Linguística do Corpus, que se alimenta de conceitos oriundos da computação e dos estudos funcionalistas de linguagem, além de ser largamente utilizada em conjunto com estudos relacionados à tradução, produção de material didático, em estudos descritivos de registros e gêneros, inteligência artificial, multidimensionalidade, além de muitos outros campos (BIBER, 1988; 1995; MCENERY; WILSON, 1996; KENNEDY, 1998; BIBER, CONRAD et al., 1999; BERBER SARDINHA, 2000; 2007). De fato, as possibilidades seriam quase infinitas: praticamente todas as subáreas da LA alimentam-se de insumos teóricos e resultados de pesquisa empíricas provenientes de diversas áreas do conhecimento. Esse dinamismo e multidisciplinaridade inatos permitiriam à LA ampliar seus horizontes em termos dos sistemas semióticos que podem ser tomados como seu objeto de estudo.

Por ser uma disciplina que traz em seu bojo uma preocupação social inerente, muito dos estudos sobre tecnologia na LA tem tido seu foco nas discussões sobre letramento digital, as quais se caracterizariam pela análise do impacto que a tecnologia de comunicação em rede teve em nossas práticas cotidianas (BRAGA, 2007). Segundo Braga (2007, p. 185), os estudos de letramento digital foram inicialmente motivados pelo leque de novas possibilidades que os computadores possibilitam para EAD, com o desenvolvimento de plataformas específicas — exemplos gratuitos seriam o Teleduc e o Moodle além de uma enorme variedade de coursewares proprietários. Além dessas perspectivas, a comunicação digital também foi estudada a partir da perspectiva de sua esfera democratizante (LÉVY, 1998) e do processo de globalização e reestruturação do mercado. Boa parte dessa literatura está centrada nos efeitos sociais, inclusivos ou opressores que a tecnologia traz, poucos são os trabalhos que procuram discutir como produtos textuais efetivos combinam os diversos modos de linguagem no processo de significação, objetivo a que este estudo se presta.

Como veremos no capítulo 1 na página 23, a análise multimodal (KRESS; VAN LEEUWEN, 1996; LEMKE, 1998; KRESS e VAN LEEUWEN, 2001; LEMKE, 2002; NORRIS, 2002; BALDRY, 2004; LEMKE, 2004; 2006c NORRIS, 2004; 2006; para citar apenas alguns) toma um grande impulso por considerar, inicialmente, esse elemento tecnológico como um

fator-chave em um processo de emancipação da imagem em relação à linguagem verbal em suas variantes escrita e falada. Apesar de haver diversos estudos que buscam discutir elementos metodológicos nessa abordagem, poucos são os que efetivamente refletem como os diferentes modos de linguagem se combinam no processo de formação de significado. Boa parte deles, até o presente momento, como é o caso de Baldry; Thibault (2006); Norris (2002); Iedema, (2001), tem seu foco ou no processo de transcrição e anotação ou na definição de unidades de análise. Minha pretensão é ir um pouco além, quero entender como o processo de construção do significado se dá. Apesar da literatura multimodal afirmar que todos os modos em um texto formam um único amálgama responsável pelo significado, não há muitas análises efetivas de como esse processo efetivamente ocorre, e poucas reflexões sobre parâmetros teóricos-metodológicos que orientem análises de produções fílmicas na LA.

De fato, esta tese procura entender como acontece a interação entre os diversos modos de linguagem nos vídeos estudados, de forma a realizar o levantamento minucioso da contribuição que cada um deles pode trazer para a construção de significado. Antes de tudo, vejo essa questão não apenas como teórica, mas também metodológica; existem poucas formas de transcrição⁶ e levantamento de dados em LA que deem conta dessa realidade de análise. Além disso, boa parte das tentativas de análise audiovisual — durante esta tese o leitor terá a oportunidade de ter contato com diversas delas — parte da transcrição da linguagem verbal como elemento determinante para discussão de seu conteúdo; contraexemplos seriam trabalhos como o de Iedema (2001) que analisa a forma como o processo de edição/montagem do filme pode influenciar na interpretação das personagens.

Essa realidade traz uma questão importante à baila, qual é o limite de explicação que as categorias de análise trazidas pela Gramática Sistêmico-Funcional e, supostamente, herdadas pela análise multimodal possuem no contexto dos vídeos estudados nesta tese? De fato, como se observará no decorrer do trabalho, essas categorias necessitam de apoio em uma literatura

⁶Utilizo o termo transcrição de uma forma mais ampla do que seu significado tradicional, que seria a reprodução gráfica dos sons e particularidades da fala. Nesta pesquisa, o termo transcrição está sendo utilizado como anotação textual de elementos significativos, os quais serão posteriormente a base da análise multimodal.

especializada na área técnica de vídeo e comunicação. Isso advém do fato da necessidade de se compreender o fenômeno do vídeo a partir de sua estrutura própria, que conta com diversos modos de linguagem, cada um com uma sintaxe e vocabulário próprios. A ideia central é entender como escolhas com parâmetros tão diferentes podem constituir um significado único, que depende de todas elas.

O vídeo como objeto de estudo

Uma parte significativa dos estudos em LA relacionados à área de tecnologia está centrada, principalmente, nas transformações que ela propicia; buscando-se, mais das vezes, compreendê-la como uma prática social. Há, por conseguinte, um rico conjunto de discussões relacionadas ao processo de experiencição dessas diversas tecnologias, com especial ênfase para a internet, centrando a discussão nos impactos relacionados à exposição a tais tecnologias, além do seu aprendizado e dos ganhos sociais que ele pode trazer.

Entretanto, é importante ressaltar que a internet não é a única experiência tecnológica presente em nosso mundo cotidiano. Como acabei de colocar, nos últimos 30 anos, uma grande transformação tem se operado em nossa arena tecnológica, videogames, minigames, uma infinidade de mídias locativas, além da vulgarização do vídeo e das câmeras digitais. Dentre o conjunto de possibilidades que se estabelece, escolhi discutir o vídeo, pela sua importância cada vez maior em nosso cotidiano. Antes de meados dos anos 80, câmeras de vídeo e videocassetes eram grandes, caros e de difícil acesso, seu processo de miniaturização facilitou sua utilização por pessoas não ligadas aos grandes *broadcasters*.⁷ Como coloca Machado (2007a), a vulgarização gradativa dessa tecnologia gera um processo de fetichização de seu uso, principalmente por parte de artistas, que passam a usar o vídeo cada vez mais em seus experimentos. Ao mesmo tempo o constante crescimento do parque televisivo faz com que o vídeo se torne uma das mídias mais importantes do século XX (ELWES, 2005),

⁷Termo em inglês adotado em língua portuguesa para classificar empresas especializadas em transmissão midiática de informação, como emissoras de TV e rádio.

da vulgarização das TVs e câmeras de gravação, ao videocassete e DVDs, o vídeo tem estado cada vez mais ligado ao nosso dia a dia. Além disso, com a invenção do VHS e tecnologias a ela posteriores, equipamentos de gravação têm se tornado cada vez mais acessíveis ao usuário doméstico, que passa a registrar seu cotidiano.

Esse processo tem início graças a alguns fenômenos sociais que devem ser discutidos. Como já colocado em Lima-Lopes (2011), a virada entre os anos 80 e 90 é um ponto-chave para a discussão do papel do vídeo na sociedade, uma vez que se instaura o início de um processo de vulgarização do vídeo, que culminará com o surgimento da bitola digital e sua íntima relação com a internet, que passará a ser o meio de divulgação preferencial do vídeo.

Isso parece se iniciar em duas frentes: uma relacionada à disponibilização de materiais para gravação e consumo doméstico de filmes e programas de televisão e outro relacionado à disponibilização de *camcoders*⁸ para o consumidor comum. Na primeira, há comércio de aparelhos de videocassete domésticos, que realizam uma dupla função: uma relacionada à possibilidade de assistir a produtos audiovisuais comercializados por distribuidores especializados, e outra ligada à gravação de programas transmitidos pelas emissoras. No caso da primeira, temos a desvinculação entre o aparelho televisivo e a emissora: é o momento em que a TV vira suporte para exibição de algo que não está sendo enviado para a antena, e que nem mesmo é parte da grade de programação. No caso específico do VHS — formato criado pela Panasonic que se torna o padrão de mercado para os aparelhos domésticos —, essa possibilidade de consumo se limita, quase que exclusivamente, a lançamentos cinematográficos. Na segunda, há, pela primeira vez, a possibilidade de gravação de um programa transmitido pela emissora de forma a assisti-lo fora do horário de transmissão previsto em grade. O que se tem nessa época é uma nova interação com a programação; a televisão tem sua função monitor descoberta pela grande massa. Nessa nova função, constrói-se um embrião do futuro internauta que será capaz de construir sua própria programação e ir além daquilo que lhe é oferecido.

⁸Camcoders são equipamentos que gravam vídeo e áudio em um mesma mídia, ela é câmera e gravador em uma só unidade. Ela se contrasta com câmeras que se conectam a unidades de gravação externas.

Já a disponibilização de *camcoders* tem uma história relacionada ao processo de miniaturização dos equipamentos de vídeo: o nascimento de bitolas, como Hi-8 e Mini DV, e mesmo gravação direta em DVDs convencionais, facilita um processo de apropriação tecnológica anteriormente possível apenas aos profissionais da área de comunicação. Mesmo sem recursos mínimos de edição, o barateamento e a simplificação gradual desses equipamentos já permitem o registro do cotidiano por indivíduos comuns. Este material ainda não é divulgado além do círculo familiar ou de amizades mais próximas, o seu suporte de exibição é o próprio aparelho de televisão e os diversos players, por meio de adaptadores, fitas ou de cabos modelo RCA para áudio e vídeo, que aos poucos se tornam padrão nos equipamentos domésticos (LIMA-LOPES, 2011).

A vulgarização tecnológica do vídeo e sua inclusão como uma tecnologia do cotidiano teve um especial crescimento com o surgimento das bitolas 100% digitais para captura de vídeos e *stills* fotográficos, aliadas à proliferação dos computadores. Inicialmente pensadas para modificar o sistema de distribuição de imagens, as câmeras digitais — que acabaram por substituir a película fotográfica, último remanescente do celuloide fotossensível em nossa cultura tecnológica —, a partir do início do século XXI, passam a se tornar o padrão de mercado, quebrando algumas barreiras importantes. A primeira está relacionada ao seu valor de aquisição; mais baratas que as *camcoders* de vídeo digital disponíveis no mercado, gradativamente vão tendo sua capacidade técnica de gravação de vídeo melhorada. Para o consumidor popular, ela causa outro impacto importante, que é a não aquisição dos equipamentos semi-profissionais. Essa teia de inclusão e vulgarização do vídeo não poderia estar completa sem a criação das câmeras de vídeo para celulares, cuja constante evolução permitiu, mesmo com as diversas limitações técnicas impostas, um registro foto e videográfico bastante peculiar do cotidiano (LIMA-LOPES, 2011, p. 53-54).

Como resultado prático dessa realidade, vivida em especial nos últimos dez anos, o vídeo tem sido um meio cada vez mais acessível, permitindo que indivíduos não profissionais realizem registros e os distribuam livremente na internet. Nesse contexto, a metodologia

aqui desenvolvida pode ser importante tanto em uma perspectiva que vise ao letramento para captação, produção e interpretação desses textos, algo que efetivamente não acontece em nossos ambientes escolares, quanto em uma perspectiva analítica e de compreensão de linguagem, dado que são poucos os estudos em LA que se debruçam sobre tal objeto. No que tange à produção, esse método pode ser relevante por ajudar na exploração e conscientização do significado de elementos de linguagem que podem ser utilizados em produções de caráter privado ou público. Digo isso porque cada vez mais os canais de divulgação online de vídeo, como *YouTube* e *Vimeo*,⁹ estão sendo utilizados para divulgação de vídeos que discutem questões sociais do cotidiano de seus usuários. Outra implicação importante desta tese é para o processo de recepção desses vídeos. De certa forma, a consciência de como as escolhas nos modos de linguagem se tornam significativas pode ser um importante elemento para o desenvolvimento de criticidade a respeito daquilo que se assiste, tanto nesses canais como nos diferentes meios de massa (como televisão e cinema). A ideia central nesse ponto seria ajudar a desenvolver ferramentas sofisticadas de análise e percepção discursiva, ajudando a melhor discutir os processos ideológicos presentes nas produções audiovisuais.

No caso específico das peças audiovisuais, Martinec e Salway (2005) colocam que os novos desenvolvimentos em mídias eletrônicas criam um ambiente propício para uma grande emancipação da imagem na cultura ocidental, possibilitando a manipulação e livre distribuição de textos que agregam em si modos de linguagem além do verbal. Se observarmos um exemplo isolado dessa nova tecnologia, o formato de livre distribuição conhecido como *.pdf, veremos que *ebooks* já são capazes de suportar vídeo e áudio agregados em textos com formatação próxima a livros convencionais. Nesse sentido, entender essa transformação e modificação de linguagem propiciadas por tal realidade traria resultados mais do que gratos à LA. Isso porque estudar as relações entre os diferentes modos de linguagem pode contribuir para melhor utilização dessas mídias em situações de ensino-aprendizagem, além de nos ajudar a compreender como o seu uso afeta as relações sociais e humanas. Um dos resultados práticos

⁹Ver <http://vimeo.com/>

dessas modificações é a valorização da linguagem visual; como colocam Kress (2005); Kress e van Leeuwen (1996); Iedema (2001; 2003), Baldry e Thibault (2006), os meios digitais têm propiciado grande ênfase na imagem como um dos meios preferenciais de comunicação, levando a uma desvalorização da linguagem verbal, em especial sua modalidade escrita. Para esses autores, trata-se do estabelecimento de um novo paradigma, no qual não existe mais um código preferencial para comunicação, todos os modos de linguagem são igualmente importantes. Isso cria necessidade de discussão mais profunda sobre o vídeo, uma vez que esse meio de expressão está sendo utilizado nas mais diversas ambiências sociais e políticas sem que haja uma reflexão sobre suas potencialidades sociais. É importante pensar que a escola precisa desenvolver parâmetros para incorporação dessas novas práticas sociais, o que passa necessariamente pela compreensão de seus elementos estruturais básicos.

Este trabalho parte do pressuposto que a própria linguagem é um processo tecnológico (FLUSSER, 2002) e multimodal. Por conta disso, esta tese realizará uma análise que parte da materialidade multimodal do vídeo como forma de discussão dos principais construtos da teoria e da análise do vídeo. Diferentemente de outros estudos (VAN LEEUWEN, 2001; 2002; BÖHLKE, 2008; IEDEMA, 2001) , este se diferencia por dois motivos principais: a discussão a partir do vocabulário visual próprio e a observação conjunta dos diferentes modos de linguagem. No caso do primeiro, partirei da análise da estrutura constituinte do vídeo, como planos, enquadramentos, seus processos de repetição e sequencialização como insumo para a reflexão sobre linguagem. Nesse sentido, este trabalho procura entender como a instanciamento se dá no espaço semiótico próprio do vídeo; meu ponto de partida será sua constituição técnica.¹⁰ No segundo, não pressuponho superioridade da linguagem verbal em relação aos

¹⁰É relevante ressaltar que os conceitos de linguagem e técnica não são necessariamente separáveis. De maneira geral, o conhecimento de técnicas específicas de produção artística se caracteriza como elemento essencial para produção dessas obras. A competência técnica representa a capacidade de utilizar a sintaxe e vocabulário específico de cada arte. Assim, conhecer tintas, sua textura, pincéis e técnicas que permitam a um pintor, por exemplo, realizar sua obra é algo imprescindível para tal artista. Meu objetivo aqui foi ressaltar que esse vocabulário específico não é necessariamente dividido por analistas da área de LA ou mesmo por indivíduos em comunidades que estão iniciando a utilização de vídeo em seu registro cotidiano. Isso leva à necessidade de construção de propostas metodológicas de análise e letramento que ajudem a superar essa questão, sendo a primeira o foco primeiro deste trabalho.

demais modos de linguagem; meu objetivo é partir da construção do texto audiovisual como um todo, sem pressupor que algum modo de linguagem tenha a prerrogativa de ditar os caminhos do processo de significação. Essa postura leva este estudo a possuir algumas características peculiares: ele parte da materialidade dos modos de linguagem presentes nos vídeos de forma a desenvolver uma proposta metodológica que auxilie no processo de compreensão da construção de seu significado. É importante ressaltar que a abordagem multimodal, em sua essência, pressupõe a interação entre o componente social e o estrutural da linguagem. Busco, por conseguinte, entender como a construção do discurso — ver seção 1.2.2.1, na página 57 e seção 5.1, na página 164, para discussões e definições sobre discurso — manifesta-se nos textos audiovisuais estudados, partindo da produção final do texto audiovisual de forma a discutir como ele reflete questões ideológicas e discursivas.

"Get a Mac": questões gerais sobre a campanha

Nesse contexto a análise dos vídeos da campanha "Get a Mac" se coloca como um ponto de partida para desenvolver, como já dito, um conjunto de procedimentos analíticos e metodológicos para análise de textos multimodais, de forma a refletir criticamente sobre tais vídeos e sobre a abordagem multimodal. Na verdade, busco ir além da constatação da importância do meio digital, que hoje em muito já extrapola o simples uso do computador e está em mídias como o videogame, tanto em console de mesa quanto portáteis, equipamentos audiovisuais (como DVDs, televisores e B-Rays já capazes de acessar a rede). Compreender esse processo é importante na formação de indivíduos não apenas capazes de avaliar essas produções criticamente, mas também de produzi-los de forma competente.¹¹ Isso possibilita-

¹¹Não estou utilizando o termo competente a partir de seu significado na abordagem gerativista. A vulgarização do vídeo tem sido algo relativamente recente em nossa comunidade e, como já discutido em Lima-Lopes (2012), está ligada de forma íntima com a miniaturização, barateamento e convergência midiática propiciados pela microinformática e telefonia celular. Por conta disso, acredito que as ferramentas necessárias para produção de textos vídeo-fílmicos exige um conhecimento técnico específico, ver nota de rodapé 10, na página 13, que pode ser o foco de pesquisas descritivas, como é o caso desta tese, e ações de letramento específicas.

ria a utilização desse meio como elemento de expressão de significados socialmente motivados dentro dos mais diferentes contextos de situação (ou registros).

Não obstante, existem poucas pesquisas que tomam para si tais objetivos, e este trabalho visa a preencher parte de tal lacuna. A opção por vídeos publicitários se dá pelo forte papel que essa mídia possui dentro dos meios de comunicação massiva. Como bem colocam Adorno e Horkheimer (2002), os meios de comunicação de massa têm a questão mercadológica como seu elemento central, sendo isso que propicia o surgimento da publicidade como forma de financiamento para as atividades daquilo que os autores nomeiam como Indústria Cultural. A Apple Inc., empresa que produz as peças publicitárias estudadas nesta tese, tem se destacado nos últimos 20 anos por ser a única empresa de computadores pessoais capaz de competir com a Microsoft Inc. O foco principal da Apple, desde o lançamento da primeira versão do seu computador iMac, em 1994, tem sido a criação de produtos que se destacam pelo design diferenciado e funcionalidade; temas amplamente explorados em todas as suas campanhas.

No caso específico da campanha "Get a Mac", observaremos que as escolhas referentes aos assuntos tratados em cada uma das peças publicitárias em vídeo estão relacionadas a uma tentativa clara de contrapor dois conceitos diferentes de computadores: a plataforma Macintosh, comercializada pela Apple, e a plataforma PC, inicialmente criada pela IBM e, posteriormente, licenciada para diversos fabricantes. Como resultado disso, todos os comerciais comparam essas plataformas, mostrando, invariavelmente, a superioridade dos computadores Apple. Essa é uma contraposição entre duas filosofias comerciais distintas: a IBM e suas licenciadas são caracterizadas por um sistema de produção aberto, diversas empresas fabricam peças compatíveis entre si que podem ser intercambiáveis ou indistintamente utilizáveis na montagem de um computador. Nesse modelo, não são controladas questões relacionadas ao design das peças vendidas por esses fabricantes, apenas as especificidades técnicas e padrões de plataforma. Já os programas, por seu turno, são produzidos por diversas empresas, sendo sua aceitação no mercado ou os acordos comerciais de distribuição sua garantia de sobrevivência. No início do século XXI, a IBM passou a se dedicar apenas

à computação de grande porte, deixando o mercado de computação pessoal exclusivamente nas mãos de empresas licenciadas.

Já a Apple está apoiada em sistema de produção totalmente diferente. Não existem empresas licenciadas para produção dos componentes de seus computadores, ela monitora essa produção sob encomenda e monta suas próprias máquinas, além de comercializá-las em suas próprias lojas ou revendedores autorizados. Apesar de permitir que programadores e empresas independentes criem softwares para sua plataforma, a Apple se responsabiliza pela criação do sistema operacional de suas máquinas, o MacOs. Além da produção de componentes, a Apple controla todas as questões relativas ao design de suas máquinas, tanto no que diz respeito à sua parte técnica (ou interior) quanto à parte estética (ou exterior), com influência direta sobre a concepção visual dos produtos. Dessa forma, os produtos vendidos ao redor do mundo são virtualmente idênticos, apenas com algumas modificações de caráter mandatório; por exemplo no Brasil, Iphones podem ser desbloqueados graças às normas da ANATEL. Já nos países de língua nórdica, eslava e na Alemanha, os teclados são adaptados para o alfabeto local, o volume dos Ipods são menos potentes em países do mercado comum europeu, devido a regulamentações da área da saúde, entre outras modificações.

Um ponto relevante a ser observado é que, a partir do ano de 2006, computadores Apple passaram a utilizar uma arquitetura de processamento baseada em chipsets¹² da marca Intel, famosa por produzir processadores e placas-mãe para computadores da arquitetura PC. De certa forma, essa realidade fez com que algumas diferenças fossem diminuídas, especialmente em termos de desempenho e constituição interna das máquinas. Apesar de utilizar chipsets fabricados exclusivamente para Apple, essa transformação aproximou Macs de PCs; agora com ajustes técnicos, não necessariamente simples, já é possível adaptar softwares de fora a funcionarem em ambos os sistemas. A própria Apple lançou um software chamado *Boot Camp*, que possibilita a instalação do sistema operacional Windows em seus computadores;

¹²Um chipset pode ser definido como um conjunto de circuitos integrados, projetados para funcionar de forma conjunta. São os principais componentes da placa-mãe de um computador, o que inclui o processador.

além disso um exército de hackers, em diversas ocasiões, portou as diversas versões do sistema operacional da Apple, o MacOS X, para rodarem em máquinas não autorizadas.

A consciência da minimização dessas diferenças talvez tenha influenciado na forma como a campanha "Get a Mac" foi produzida. Além de ressaltar as diferenças no sistema de produção, a campanha tenta minimizar as possíveis semelhanças entre os dois sistemas, Mac e PC são representados como sistemas totalmente diferentes, sem proximidade alguma entre suas arquiteturas de microinformática. A consciência das diferenças mercadológicas e das semelhanças em termos técnicos ajudam a compreender muitas das escolhas semióticas presentes nos vídeos que compõem essa campanha, em especial a oposição dicotômica ali colocada. Se seguirmos por esse caminho de interpretação, podemos afirmar que as personagens dos vídeos representam não apenas a oposição comercial entre as duas principais plataformas computacionais, mas dois estilos diferentes de produção e gerenciamento de negócios. Os subtemas tratados em algumas das peças de vídeo — como funcionalidade, design, segurança, facilidade de uso, utilização pessoal vs. corporativa, entre outras —, parecem ser ampliações dessa dicotomia básica. Observa-se, assim, que "Get a Mac" parece ser uma campanha pensada de forma a reafirmar alguns mitos em relação à diferença entre essas plataformas, tentando associar um valor aos computadores Apple que vão além de suas qualidades técnicas.

Como discutirei na análise dos dados, ver capítulos 4, 5 e 6, os vídeos que compõem o corpus desta tese possuem um uso da câmera e da objetiva bastante simples. Os anúncios em vídeo de "Get a Mac" possuem uma câmera parada e sem movimentos, não há passagens de foco, apenas mudanças de enquadramento, na sua maioria com corte seco. Esse tipo de estrutura composicional centra a ação dos vídeos na fala e nas ações das personagens, que lembram em muito as comédias físicas do início do século XX, em especial as de Buster Keaton. Isso resulta em gestos e expressões exagerados, todos sempre muito presentes e marcados.

Isso influenciou este trabalho de duas formas. Em um primeiro momento, direcionou a análise para alguns elementos que pareceram relevantes no que tange à sintaxe fílmica desses

comerciais: a sequência de cortes. Em todas as peças publicitárias em vídeo de "Get a Mac" se estabelece uma sequência bastante previsível de cortes, algo que parece estabelecer um padrão claro, que ajuda no reconhecimento de como ocorre o fluxo de informações nestes vídeos. Outra influência importante está na escolha dos filmes e vídeos utilizados na discussão teórica, ver capítulos 1 e 2, obras cujo uso dos recursos de linguagem poderiam ser classificados como "cavaleiros de plateia" ou narradores clássicos.

Perspectiva de linguagem adotada nesta pesquisa e estrutura do trabalho

Acredito que todos os modos de linguagem presentes nos vídeos analisados nesta tese — o que incluiria, por exemplo, a linguagem visual, gestual ou qualquer outro modo de linguagem — têm seus significados expressos por meio de escolhas interrelacionadas e de importância equivalente. Isso torna importante ressaltar que, nesta tese, tomo como válidos os pressupostos da linguagem multimodal, em especial de que a língua constitui um sistema semiótico que está em interação com outros sistemas que, de forma conjunta, constituem o significado. Todos esses sistemas são, por conseguinte, de igual importância no processo comunicacional, o que tiraria a centralidade da linguagem verbal como sistema preferencial de constituição de significados (BALDRY; THIBAUT, 2006). Por conta disso, discuto os modos de linguagem presentes nos vídeos analisados nesta tese em pé de igualdade, pressuponho que o significado é constituído de forma simultânea, integrando todos os sistemas semióticos à disposição durante a fase de produção. Isso explica por que, na análise dos dados, não parto da linguagem verbal como ponto de motriz para compreensão dos textos audiovisuais estudados, busco analisar os modos de linguagem de forma simultânea observando qual o papel que cada um possui para a criação do amálgama de significados que tais vídeos representam. No Brasil, já existem diversas pesquisas que caminham pela mesma seara, como seria o caso de Almeida (2006), que trabalha com anúncios de brinquedos, Böhlke (2008), que discute

a questão da imagem feminina em anúncios de medicamentos, Ferreira (2011), que faz um estudo comparativo entre a relação texto vs. imagem em anúncios da revista Times, e Moura (2011), que analisa a realização de produtos audiovisuais realizados por meninos de rua. A exemplo de Böhlke e Moura, tomo o vídeo como objeto de análise em minha tese. No caso da primeira, observamos que a análise das peças prima pela discussão de frames individuais; a exemplos de análises como as de Van Leeuwen (2001), cada frame é congelado e analisado como se fosse uma fotografia estática. Apesar da importância desse tipo de trabalho, que traz grandes insights sobre a constituição semiótica da composição, meu interesse converge com o de Moura (2011), cujo objetivo é discutir a imagem em movimento, considerando o texto videográfico como uma unidade mínima de análise. No caso de Moura, sua preocupação era compreender como os meninos de rua utilizavam tal linguagem como forma de articulação e expressão de sua visão de mundo. No caso desse estudo, meu interesse está voltado para trazer à baila uma discussão que sirva como ponto de partida no desenvolvimento de uma sistemática de estudo da linguagem em movimento, algo que permita desenvolver algum suporte para interpretação semiótica e para o letramento em vídeo.

De forma a conseguir realizar esta pesquisa, a estrutura deste trabalho foi definida em sete capítulos. No capítulo 1, Teoria da multimodalidade, na página 23, reflito sobre os conceitos epistêmico-teóricos que dão origem à abordagem multimodal, para, posteriormente, discuti-la de forma crítica. Aqui pretendo estabelecer alguns pressupostos que estão na base de interpretação da discussão multimodal, e que serão constantemente retomados no processo de discussão teórica, análise e discussão dos dados. Segue-se então o capítulo 2, na página 85, cujo objetivo é refletir sobre aplicação da Metafunção Organizacional (ou Textual) em peças audiovisuais, de forma a complementar a literatura já existente com discussões a respeito de como elementos próprios da linguagem vídeo-filmica, além de modos de linguagem corporificados, como gestos e olhares, também poderiam ser considerados nessa equação.

Já no capítulo 3, Definição do corpus, temática da campanha e elementos de linguagem videográfica, na página 109, procuro apresentar o corpus de estudo deste trabalho, discutir

quais foram os critérios de coleta, análise e transcrição de dados, além de trazer algumas considerações importantes sobre a linguagem vídeo-filmica.

Passo então para o capítulo 4, Reflexões sobre a transcrição multimodal, na página 127, cujo objetivo é propor um modelo de análise para textos multimodais, em especial os vídeo-fílmicos, objeto de estudo desta tese. Neste capítulo, parto do princípio de instanciação como elemento norteador do trabalho do analista, procurando fazer um levantamento sistemático das escolhas em cada um dos modos de linguagem e discutindo como cada um deles contribui para o amálgama de signos que é o significado. Observa-se que a instanciação do significado faz com que as diferentes escolhas sejam condicionadas de forma a oferecer a constituição de diferentes facetas do significado. Entendo que esse modelo pode oferecer aos futuros pesquisadores parâmetros para nortear análises e reflexões sobre a escolha ou construção de parâmetros mais precisos.

No capítulo 5, O vídeo enquanto sistema de construção de significados: explorando modos de linguagem, página 163, busco discutir os vídeos analisados a partir de duas perspectivas. Na primeira, exploro como o conceito de constituição pode ser utilizado na análise de textos multimodais, desenvolvendo uma reflexão sobre a forma como os diferentes constituintes dos diversos modos de linguagem podem ser analisados em conjunto nos vídeos publicitários que compõem o corpus desta pesquisa. Na segunda, utilizo os resultados desta parte inicial de forma a ampliar a discussão tendo em vista quatro perspectivas, o nível do conteúdo e o discurso, o design como encontro entre o conteúdo e a expressão, a produção e suas escolhas nos diferentes contextos de produção e a expressão nos processos de distribuição.

Por sua vez, o capítulo 6, Coesão e fluxo de informações no vídeo, página 213, traz questões relacionadas à construção do processo de coesão e fluxo de informações em textos vídeo-fílmicos. Meu objetivo é entender como se constrói um sistema de referências internas que permita conectar aquilo que já foi exibido com aquilo que está aparecendo naquele exato momento na tela. Uma análise detalhada demonstra que há diversas estratégias de repetição e manutenção de signos em quadro, que contribui para a construção da tecitura textual nos

vídeos de "Get a Mac". Aqui também foi realizada uma minuciosa análise da forma como os componentes estruturais e atitudinais estão relacionados com o processo de sequencialização da informação nos vídeos estudados. Seguem-se as Considerações Finais, na página 245, que têm por objetivo discutir os resultados desta tese além de apresentar possíveis aplicações e as limitações deste trabalho.

Capítulo 1

Teoria da multimodalidade

We need conceptual frameworks to help us cope with the complexity and the novelty of these new multimedia constellations [...]

Jay A. Lemke

Introdução

Aqui reflito sobre os conceitos epistêmico-teóricos que dão origem à abordagem multimodal, para, posteriormente, discuti-la de forma crítica. Como já colocado na introdução deste trabalho, pretendo estabelecer alguns pressupostos que estão na base de interpretação da discussão multimodal, e que serão constantemente retomados no processo de discussão teórica, análise e discussão dos dados. A função deste capítulo é fornecer base de interpretação para os núcleos teórico-analíticos que se seguem: conceitos como modos de linguagem, contexto de cultura, Metafunções, níveis de contexto, escolhas e outros tão comumente citados, e raramente definidos em estudos de multimodalidade, serão aqui discutidos e definidos em sua origem, a teoria hallidiana. Naturalmente, este capítulo também busca discutir as adaptações

necessárias que permitiram a transposição teórica representada pela Multimodalidade, além de refletir criticamente alguns pontos dessa teoria.

A abordagem multimodal tem sua origem a partir dos trabalhos do linguista britânico M. A. K Halliday (HALLIDAY, 1973; 1978; 1985; 1994; HALLIDAY; HASAN, 1989; HALLIDAY; MATTHIESSEN, 2004), e parece ter evoluído em duas frentes complementares. A primeira, tem como principais autores Kress e van Leeuwen (KRESS, 1993; KRESS; VAN LEEUWEN, 1996; 2001; 2002; KRESS, 2003; 2005; 2006; 2010) e Lemke (LEMKE, 1998; 2002; 2004; 2006a; b; c) e atesta que o desenvolvimento tecnológico tem criado novas formas de significação textuais que, de certa forma, causam uma perda de status por parte da escrita. Assim, a palavra grafada, em especial em sua circulação dentro dos diferentes *mass media* — como é o caso do livro, jornal, entre outros —, tem perdido seu poder dentro de nossa sociedade como forma preferencial de divulgação do conhecimento chamado letrado. Esse primeiro grupo tem um papel seminal no desenvolvimento da abordagem por trazer trabalhos de fundo teórico, expandindo a sociossemiótica para outros modos de linguagem que permitiram a autores como Norris (2002; 2004; 2006); e O'Toole (1994, 2004) afirmar — a partir de análises de produtos de mídia, além de interações do cotidiano —, que língua não é o único modo de linguagem utilizado no processo de construção de significado, seja qual for a situação de comunicação na qual estejamos envolvidos. A segunda estaria relacionada a uma transformação que se opera no trabalho de Kress (2010), que passa a argumentar em favor de uma Teoria Multimodal que se constitua como uma ampla forma de compreensão do processo de comunicação humana. Nesse novo contexto, sua obra vai ganhando uma relativa independência da Gramática Sistêmico-Funcional (doravante GSF), em termos de adaptação de suas categorias de análise. Isso, entretanto, sem que haja um distanciamento da base epistêmico-teórica dada pela teoria hallidiana.

No caso deste capítulo, a linha condutora de minha argumentação será o trabalho de Kress, principalmente por sua importância na construção de uma teoria multimodal que busca independência teórica. De forma a cumprir esse objetivo, o capítulo a seguir se divide

em 3 seções. Na primeira, Multimodalidade e suas origens na Sociossemiótica Hallidiana, na página 25, discuto os pontos-chave da GSF que serviram como base para a construção da Sociossemiótica Multimodal de Kress. É relevante observar que essa seção não tem por objetivo ser uma discussão a respeito da totalidade da teoria hallidiana, ali serão trazidos à baila apenas os conceitos que são importantes para a compreensão da transposição teórica utilizada como base teórica para este estudo. Segue-se, então, em Transformações no contexto social e linguístico: em busca de uma teoria Sociossemiótica Multimodal, na página 41, uma reflexão sobre o desenvolvimento da teoria multimodal; tal reflexão tem caráter crítico e visa também a discutir sobre suas possíveis limitações teórico-práticas. Esta seção será dividida em duas outras subseções, em A imagem e a tecnologia, na página 43, discuto os primeiros estágios de construção de uma teoria multimodal; passando então para Produção de significado, na página 49, de forma a discutir como ela evolui da percepção da transformação do meio para uma tentativa de discussão mais abrangente da comunicação.

Encerro o capítulo então com os Resultados parciais, na página 77.

1.1 Multimodalidade e suas origens na Sociossemiótica Hallidiana

1.1.1 GSF: os conceitos de partida

Nesta seção meu objetivo é discutir as características mais relevantes da GSF que a permitiram servir como base teórica para uma teoria multimodal do estudo da linguagem.

Existiriam várias possibilidades de transposição teórica que serviriam ao propósito de ampliar os horizontes da LA para outras semioses que não a verbal, neste trabalho faço uma escolha consciente pelas que descendem dos estudos funcionais de linguagem, em especial nos trabalhos do linguista inglês Michael Halliday (HALLIDAY, 1973; 1978; 1985; 1994; HALLIDAY; MATTHIESSEN, 2004) e sua Gramática Sistemico-Funcional. A GSF

vem se desenvolvendo como um modelo de análise linguística que possui uma forte vocação semântico-funcional, tendo como preocupação principal a forma como a linguagem é estruturada para o uso (EGGINS, 1994, p. 23-24). Nessa abordagem, língua e sociedade se influenciam conjuntamente: ao realizarmos escolhas no nível discursivo, sintático ou lexicogramático,¹³ estamos agindo sobre as pessoas, sobre o mundo e organizando a linguagem para seu uso.

Nesta abordagem a linguagem — já incorporando a transposição desse conceito para textos multimodais — é definida como um sistema de construção de significados baseados nas escolhas que o falante faz, de forma consciente ou não, realizando uma função em uma determinada situação discursiva. Esses sistemas formariam uma rede (*system network*), ou um conjunto de opções interligadas, com o objetivo de produzir significado. Essa noção de sistema privilegia a construção de uma gramática— aqui encarada como um construto que permite a expressão do significado em linguagem — paradigmática: cada escolha é feita em desprezo a outras possíveis dentro do sistema e aceitáveis naquele contexto (HALLIDAY, 1994, p. 15; HALLIDAY; MATTHIESSEN, 2004, p. 22).

Ao nos comunicarmos, nossa unidade mínima de significado é o texto, sendo este entendido como qualquer produção comunicativa que faça sentido para alguém em um dado contexto (HALLIDAY; MATTHIESSEN, 2004, p. 03). Para Halliday, o texto pode ser pensado ora como um espécime ou como um artefato. Se nos debruçarmos sobre o texto a partir da primeira perspectiva, estaremos preocupados em compreender em que medida uma dada manifestação de linguagem nos traz informações a respeito do sistema em que fora produzido. Contudo, se pensarmos em um texto como um elemento em si (ou seja, como um artefato), nosso foco pode se alterar, e nossa preocupação seria entender porque o texto significa o que parece significar, ou mesmo porque seus leitores lhe atribuem este ou aquele significado. Ou seja, para Halliday, se, por um lado, pensarmos um texto como um espécime,

¹³A GSF tradicionalmente não diferencia os níveis lexicais e gramaticais. Isso porque os analistas que seguem essa abordagem acreditam que qualquer escolha no nível lexical implica mudanças na expressão gramatical e vice-versa.

estaremos observando-o como a manifestação de estruturas e da gramática de uma dada língua, ao passo que se, por outro lado, o encarmos como um artefato, ele será o produto único de uma situação comunicativa específica, estaremos buscando compreender como os vários recursos foram empregados para produção do significado. Logo, ambas as abordagens são complementares em termos da compreensão da real função que um texto deve possuir. Isso se dá porque, dentro desse espectro, significado e estrutura estão diretamente interligados, são interdependentes e motivados pelo objetivo social da comunicação (HALLIDAY; MATTHIESSEN, 2004, p. 03-04).

1.1.2 Significado e forma: uma gramática em níveis

As características anteriormente definidas levam essa abordagem a se preocupar com a linguagem em sua totalidade: estrutura e significado passam a ter a mesma importância dentro de um sistema que os vê como indissociáveis, uma vez que, para Halliday não faria sentido separar forma de significado, elas compõem um único elemento (HALLIDAY, 1994, p. 15). Isso faz com que cada vez que um significado seja expresso, a escolha de como ele o será se torna algo não arbitrário (KRESS, 1993; STUBBS, 1996), fazendo com que qualquer escolha no nível estrutural reflita em uma escolha no nível semântico e vice-versa. Essa não-arbitrariedade do signo¹⁴ e das escolhas vai diretamente contra uma já enraizada tradição dentro dos estudos linguísticos (KRESS, 2010; STUBBS, 1996). Ela se justifica, em GSF, graças à relação que a escolha estrutural tem com a representação de um significado,

¹⁴A linguística teórica e a aplicada partem, tradicionalmente, do conceito de signo definido por Saussure em seu "Curso de Linguística Geral" (Saussure, 1959). Para Saussure (1959), o signo se caracterizaria como um conceito dicotômico: o significante seria definível como a imagem acústica, esta responsável por remeter a um significado, um conceito abstrato e comum aos falantes de uma determinada língua. A relação entre o significante e o significado seria de caráter arbitrário e convencional: o único motivo que justificaria o nascimento dessa relação seria um acordo, uma convenção entre os falantes de um determinado idioma. Diferentemente de outras disciplinas — a semiótica piereana é um exemplo de discussão mais aprofundada —, pouco tem se discutido, desde Saussure, a constituição do signo na linguística. A maioria das escolas parece tomar esse conceito saussureano como um axioma dentro da linguística; talvez a GSF seja uma das primeiras teorias linguísticas a se contrapor a tal ideia, introduzindo a questão da não arbitrariedade do signo. Para a teoria hallidiana, esse princípio é importante por ajudar a justificar a relação entre todos os níveis de significado: estrutura, expressão e significado passam a ser indissociáveis, uma vez que realizam uma função social.

dentro do que Halliday (1994) chama de potencial semiótico (ou potencial significados), que poderia ser definido como o conjunto de possibilidades significativas de uma escolha em um dado contexto. Kress (2010), para explicar esse conceito, realiza um paralelo com as artes plásticas: digamos que um escultor, por exemplo, escolha uma determinada matéria-prima para o seu trabalho, ele certamente sabe quais tipos de formas são exequíveis em cada uma das possibilidades que lhe são apresentadas. Ao escolher fibra de vidro ao invés de madeira ou pedra sabão, ele está escolhendo também as possibilidades de realização em termos de forma e significado, uma vez que alguns deles não poderão ser alcançados graças à escolha. Ao mesmo tempo, toda uma série de possibilidade se abre, dado que a fibra de vidro tem características que facilitam a expressão de alguns arranjos, formas e significados; um modelo de funcionamento da resolução de linguagem em seu aspecto matérico.¹⁵

Essa metáfora pode ser diretamente transferida para o uso da língua (no caso da GSF) e da linguagem (no caso da Multimodalidade): ao escolhermos uma determinada forma de expressão, estamos pré-selecionando os significados que podem ser por ela expressos; ao mesmo tempo que a expressão de determinado significado acontece, estamos também pré-determinando quais elementos são passíveis de utilização para sua expressão.

Essa relação direta entre a estrutura e o significado criariam um sistema de escolhas que seria, dentro da teoria de Halliday, semanticamente motivado: não há diferença entre estrutura e o seu significado, ao operarmos em qualquer um desses níveis (ou *strata*) estamos na verdade modificando todos simultaneamente. Em outras palavras, ao escolhermos um significado a ser expressado em uma determinada situação social, estamos ao mesmo tempo escolhendo a estrutura textual, lexicogramatical, fonética (ou grafológica) e fonológica em que ele será instanciado. Ao modificarmos qualquer um desses elementos, estaremos, por conseguinte, modificando o significado.

¹⁵Obrigado ao Prof. Dr. Edson do Prado Pfutzenreuter por chamar minha atenção para essa relação

É por esse motivo que a escolha tem um papel tão importante dentro desta teoria: o falante ou escritor por estar optando por realizar um significado pode constituir um discurso¹⁶ inteligível apenas a um determinado grupo, ou que tenha uma função social específica dentro de uma determinada comunidade. Ao escolher, o falante está desprezando outras possíveis formas aceitáveis em um determinado contexto, motivado pela função que uma dada escolha pode realizar (THOMPSON, 1994). Por isso ela é um dos princípios básicos da GSF; o uso da linguagem humana se estrutura por meio de um sistema, que pode ou não estar em um nível consciente: a linguagem seria um sistema probabilístico de escolhas (HALLIDAY, 1991; 1992a; b; 1993b), que está disponível de forma condicional pelo ambiente social em que o falante se encontra. Apesar de não se aplicar diretamente à temática desta tese, uma rápida discussão sobre o "Princípio Idiomatista/Idiomático" (*Idiom Principle*),¹⁷ como definido por Sinclair (1991), poderia ser bastante ilustrativa. Para o autor, um dos fundadores da Linguística do Corpus na Inglaterra, o conceito de lacuna e preenchimento (*Slot and Filling*) tradicionalmente defendido pelos linguistas gerativistas e estruturalistas não é válido. Isso porque, segundo Sinclair, a pesquisa com corpus revela que os significados tendem a ser realizados por estruturas relativamente estáveis, o que é facilmente verificável em uma pesquisa de Colocados e N-gramas. Ou seja, o que há na verdade é um número de escolhas já pré-definidas socialmente que já estão ligadas ao significado que tradicionalmente expressam.¹⁸

¹⁶Para uma definição do termo discurso e de seu papel nessa abordagem, ver seção 1.2.2.1, na página 57, e na seção 5.1, na página 164.

¹⁷A Linguística do Corpus e GSF possuem um parentesco teórico o que as faz muitas vezes complementares em diversos aspectos. Isso advém do fato de Halliday e Sinclair terem sido, ambos, contemporâneos em sua formação, orientandos do linguista britânico J. R Firth (1890-1960), o primeiro a discutir essas questões motivacionais da língua em seus trabalhos de fonologia e fonética do inglês (STUBBS, 1996). Ver Beaugrande (1991) para uma rápida discussão.

¹⁸Voloshinov (Bakhtin), em *Marxismo e Filosofia da Linguagem* (BAKHTIN/VOLSHINOV, 1929/1992), discute a necessidade de estudar as palavras como signos dinâmicos e sociais, que têm significados diferentes para as diversas classes sociais em seus mais diversos contextos históricos. Essas palavras, não estariam sujeitas a simples compreensão passiva, mas seriam elementos dinâmicos com motivações sociais. Apesar de ecoar essa questão social, a GSF e a linguística do corpus não partem da palavra isolada para definição de seu significado. Em ambas as escolas, acredita-se que as regularidades de linguagem expressam significados diferenciados: poderíamos determinar o significado de uma palavra por meio das estruturas nas quais ela se manifesta, observando quais palavras estão a sua volta em um dado texto — o que se nomeia co-texto.

Para Martin e Rose (2007, p. 04-05), a atividade social, o discurso e a gramática são fenômenos, diferentes interfaces da GSF, sendo cada um deles operando em um nível distinto de abstração. Essa interface se torna possível porque culturas se manifestam por meio de textos, realizações sociais de linguagem, que se manifestam por meio de unidades menores como frases, períodos e palavras. Para os autores, da mesma forma que uma cultura não é um amontoado de textos, um texto não deve ser visto como uma simples soma de unidades menores, o processo de instanciação (ou realização) seria exatamente isso: é um processo de recodificação das estruturas sociais em estruturas gramaticais que formam o texto, permitindo que níveis mais abstratos se transformem em escolhas palpáveis no nível da linguagem. Por isso, simbolizar é um aspecto importante e indispensável, uma vez que a gramática seria responsável por codificar o discurso, que, por sua vez, é responsável por codificar a atividade social (MARTIN; ROSE, 2007; KRESS, 2010). A codificação se dá sempre pela criação de um sistema simbólico, pelo qual o contexto geral da cultura e a simples escolha de um processo — verbo na terminologia de GSF — estejam diretamente interligadas.

Esse construto teórico baseado na não arbitrariedade fez com que Halliday (HALLIDAY, 1973; 1985; 1994; HALLIDAY; MATTHIESSEN, 1999; 2004) criasse uma teoria que buscasse dar conta de diversos níveis de expressão: seria ingênuo pensar que dentro da construção do significado houvesse apenas uma relação direta entre forma e estrutura, naturalmente o contexto, a situação de enunciação são elementos a serem levados em conta. A GSF, então, define uma característica importante da linguagem, seu princípio composicional (HALLIDAY; MATTHIESSEN, 2004, p. 20). Cada texto é o resultado da operação de escolha que ocorre em diferentes níveis (ou *strata*), cada um responsável por um aspecto do significado e traduzível em um sistema próprio.

Como podemos ver na figura 1, página 31, os níveis (ou *strata*) são organizados a partir dos mais abstratos — que representam o nível do conteúdo — em direção dos mais concretos, representando o nível da expressão. Algo importante a ser notado é que a lexicogramática, responsável pela escolha nos níveis menos abstratos paira ainda entre os responsáveis pelo

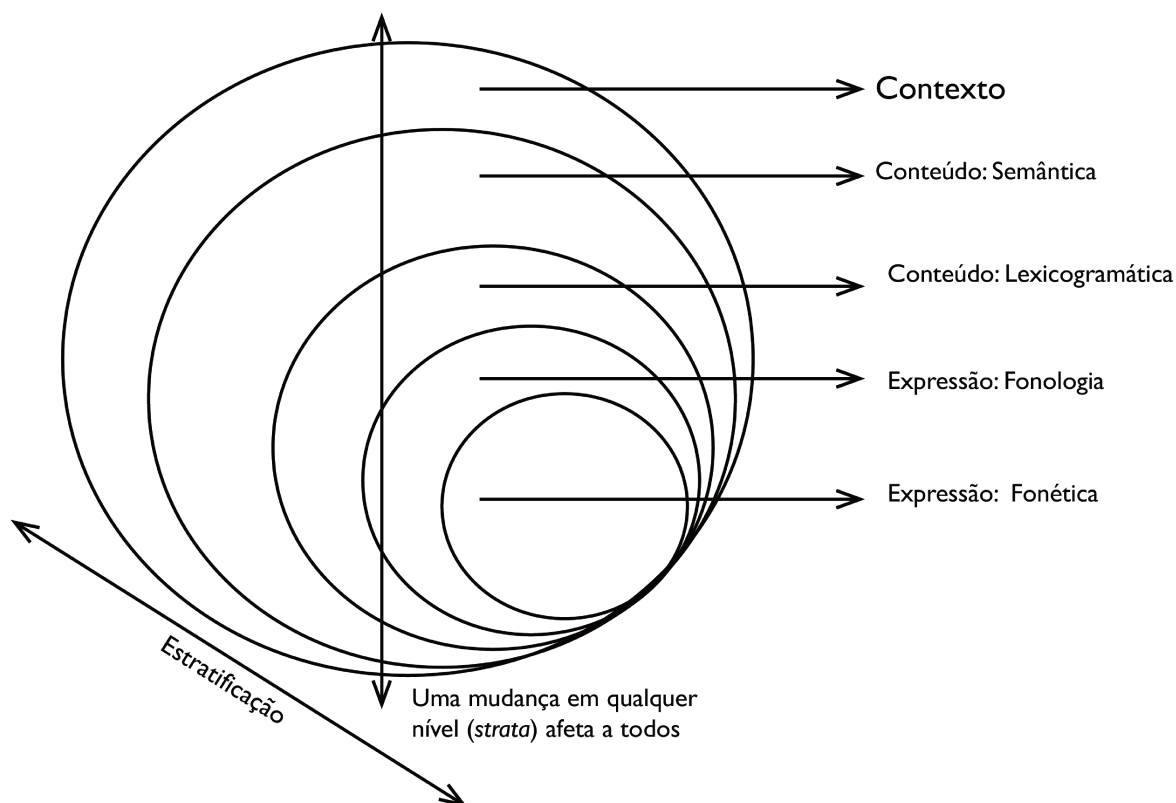


Figura 1: Níveis ou *Strata* na GSF, traduzido de Halliday (1994) e Halliday e Mathiessen (2004)

conteúdo; isso se dá porque, dentro dessa teoria, ela não é vista como um repositório de estruturas, mas como um construto semântico-organizacional que ajuda a reger as escolhas do falante.

A transposição teórica realizada pela teoria da multimodalidade pressupõe que essa teoria estratificadora dos diferentes níveis de escolha também é aplicável a outros sistemas semióticos. Acredito que a linguagem visual, videográfica, gestual, ou qualquer outra, pressupõe a expressão de significados e de escolhas em níveis que vão desde a simples realização das unidades de significado responsáveis pela ação de linguagem até seu significado em uma dada situação ou contexto. Dessa forma, uma tradução deste princípio para os textos videográficos poderia ser representada pela figura 2, na página 32.

Na figura 2, na página 32, vê-se que a estruturação dos níveis é muito similar à proposta hallidiana, mas com uma diferença, aqui tentei adaptar o conceito de níveis a uma possível

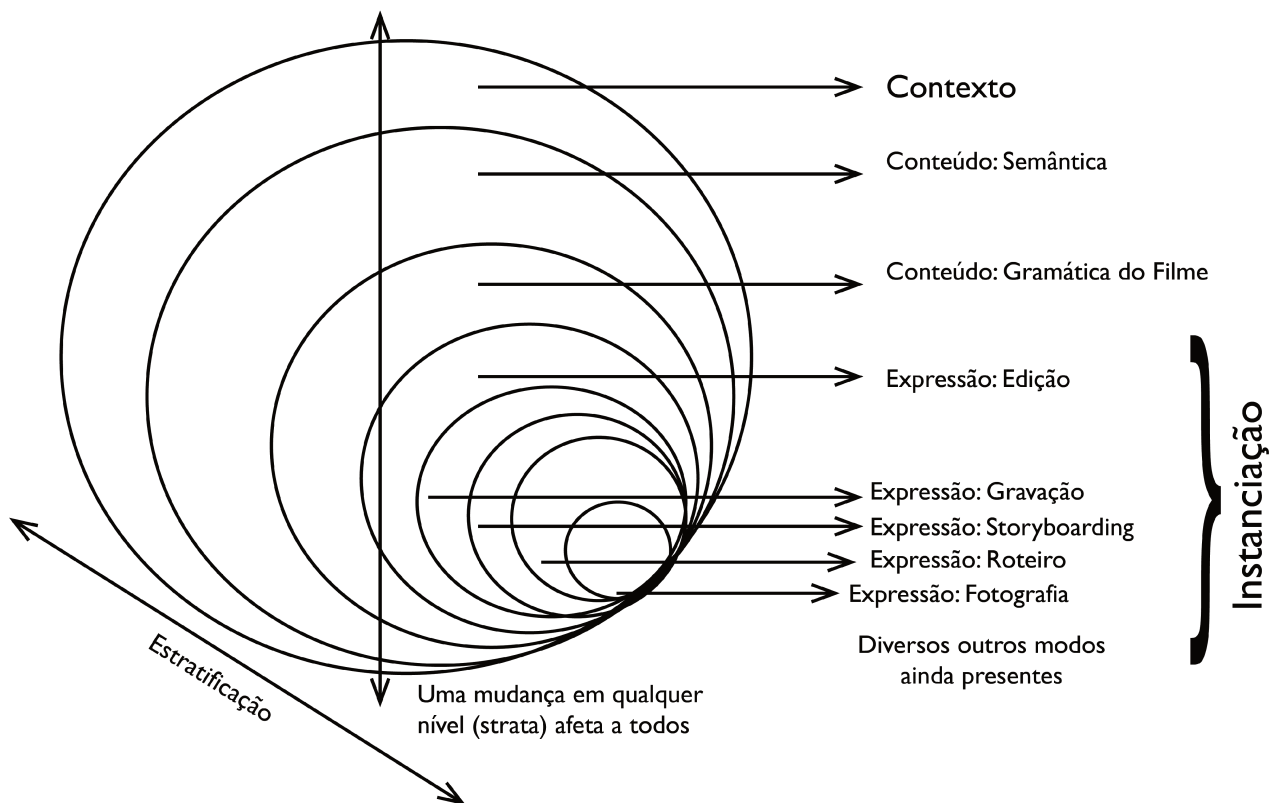


Figura 2: *Strata* e sua tradução para o Corpus inspirado em Halliday (1994) e Halliday e Mathiessen (2004)

gramática do vídeo. Cada modo de linguagem deve, de acordo com uma teoria multimodal abrangente da comunicação, ter uma gramática própria, responsável pelos construtos de expressão que ele teria em cada uma das culturas em que se manifestasse. Modos de linguagem (ou semioses) poderiam ser definidos como o resultado de um processo histórico que levou à utilização de determinados recursos materiais por uma sociedade em um processo de representação e criação de significados (KRESS, 2010, p. 11). Em outras palavras, para Kress (2010, p. 79), cada um desses recursos possui diferentes potenciais semióticos, que também seriam moldados pelo uso significativo que uma determinada cultura faz deles; um importante papel é dado às práticas sociais e históricas, pois são elas que ditarão o formato que um determinado modo de linguagem possui. Tomemos por exemplo o vídeo,¹⁹ material de

¹⁹Nesta transposição optei por manter a gramática do vídeo no nível do conteúdo, a exemplo da lexicogramática nos textos verbais. Isso se dá porque esses níveis representam elementos combinatórios e de

estudo desta tese; em nossa cultura ele acaba sendo um fenômeno ímpar por poder tanto ser um modo de linguagem como um mero registro do cotidiano (ELWES, 2005). No primeiro caso ele é um sistema elaborado que depende de uma série de outros modos de linguagem para acontecer. Ao se observar a figura 2, na página 32, perceberemos que ao mesmo tempo que cada um de seus níveis — tais como edição, roteiro e fotografia — instancia o vídeo, também poderiam ser estudados como um sistema à parte. Em outras palavras, ao mesmo tempo que um roteiro, por exemplo, permite que um vídeo seja rodado de acordo com as intenções de seu escritor e de seu diretor, ele também poderia ter uma gramática própria de instanciação, uma vez que o texto final é a materialização em linguagem de intenções de caráter textual. Essa ambiguidade permite que textos façam parte de uma gramática sem, contudo, deixarem de ter uma gramática própria; algo que me parece comum a todos os textos ditos multimidiáticos. Se tomarmos, agora, o vídeo como simples registro do cotidiano, todos esses elementos deixam de ser planejados, sem, contudo, deixar de existir de uma maneira mais intuitiva e inconsciente, parte-se para um grau inconsciente de realização. Alguns níveis, como *storyboard*, que garantem uma realização previamente controlada do vídeo, não existiriam, haveria apenas os níveis que poderiam ocorrer espontaneamente no momento efetivo do registro, como gravação, fotografia etc. Os textos videográficos analisados nesta tese pertencem ao primeiro tipo.

Isso traz à baila outro conceito também bastante importante para esta tese, o de constituição, que se aplica a todos os níveis da gramática. Nesse princípio, cada nível (ou *strata*) seria constituído pela imediatamente menor que ele: a lexicogramática seria constituída pelos níveis de expressão, como a fonética e a grafologia, que seriam imediatamente compostos pelos fonemas e pelos grafemas. Essa relação seria um dos elementos responsáveis por inter-relacionar todos os níveis na construção do significado; ao modificarmos, por exemplo, alguma variante de constituição fonética, estaríamos transformando não apenas um conteúdo lexi-

organização que, dentro da visão teórica aqui adotada, também seriam responsáveis pela produção de significado. Tal posição advém do fato de, como já colocado anteriormente, a GSF não ser uma abordagem que observe a estrutura apenas como uma expressão organizacional neutra de significado.

cogramatical, como também semântico e contextual, pois qualquer mudança pode também afetar o relacionamento entre os participantes, a expressão de suas experiências e a forma como organizam a linguagem. Naturalmente, modos de linguagem diferentes do verbal teriam constituintes específicos; um exemplo seria os gestos, que são constituídos por ações físicas com significados culturalmente estabelecidos. O vídeo é um modo de linguagem que além de uma linguagem própria, também é composto por uma série de outros modos de linguagem: ao assisti-lo também percebemos modos como a fala, os gestos, o olhar, as roupas e uma infinidade de outros, o que incluiria elementos formais e abstratos que constituem a imagem..

1.1.3 Função, contexto e metafunção

A Sociossemiótica hallidiana é, então, uma abordagem que une significado e estrutura, que busca explicar como escolhas socialmente significativas são realizadas em detrimento de outras. A gramática, para Halliday, é um sistema paradigmático que permite ao falante criar fraseados, *wordings*, (HALLIDAY, 1973; 1978; 1985; 1994; HALLIDAY; MATTHIESSEN, 2004), ou arranjos, *arrangements*, dentro de uma teoria multimodal (KRESS, 2010; VAN LEEUWEN, 2005). Fraseados ou arranjos poderiam ser definidos como combinatórias de signos utilizadas com sentido específico dentro de um ato comunicacional; a forma como os constituintes estruturais são organizados são resultado de uma demanda do contexto e das funções que realizam (KRESS, 2010, p. 145).

Para van Leeuwen (2005, p. 75-76), Halliday distingue dois tipos de função: a) função estrutural, relacionada a questões sintáticas, o que é um tema já largamente explorado na linguística do século XX, e b) a função social de proposição, definida como sendo seu impacto, valor e significado no contexto de produção. A teoria de Halliday parte de ambas como válidas, determinante das escolhas que realizamos em todos os níveis (HALLIDAY; HASAN, 1989, p. 5), e sua relação com a primeira: estruturas linguísticas geram significados sociais, os quais, por sua vez influenciam a estrutura. Por isso, para que se possa analisar um texto, é indispensável que seja observado aquilo que acontece à sua volta, encarando a ele

e ao contexto como interdependentes (THOMPSON, 1994, p. 9); outrossim, influenciando escolhas de palavras, de estruturas (EGGINS; MARTIN, 1997, p. 232) e de modos de dizer. De fato, segundo Halliday (1994, p. 29) o propósito de uma gramática funcional é fornecer meios de interpretação para as diferentes funções que uma mesma estrutura pode realizar. Tais funções são o resultado da interpretação analítica da língua, sendo responsável pela estruturação de nossa compreensão de todos os níveis (ou *strata*).

Dentro dessa tradição sociossemiótica, dois níveis de contexto são estabelecidos: o de situação e o de cultura.²⁰ O contexto de cultura pode ser entendido como o pano de fundo no qual a interação está inserida (HALLIDAY; HASAN, 1989) disponibilizando um potencial de realização de significados em uma cultura. Esses significados influenciam na forma como textos são organizados em nível estrutural e informacional, formando o que chamamos de gêneros. É de especial importância para esta pesquisa o trabalho desenvolvido por Martin (MARTIN, 1991; 1992; EGGINS; MARTIN, 1997; MARTIN, 1997; MARTIN, MATTHIENSEN et al., 1997; MARTIN, 2000), que parte do conceito de gêneros desenvolvido por Hasan (HALLIDAY; HASAN, 1989). Para Martin e Rose (2007, p. 08), o gênero é uma ação social de linguagem, orientada de forma a realizar um objetivo específico e é estruturado em partes, chamadas de estágios, que levam o leitor a conhecer o propósito de um texto. Isso levaria a padrões previsíveis e reconhecíveis aos falantes de uma determinada comunidade, sendo que estes podem variar desde um simples aperto de mão a complexos relatórios e apresentações. Todavia, apesar dessa variabilidade, esses padrões seriam consistentes a ponto de possibilitar a um falante prever como um gênero provavelmente será em sua realização, permitindo-o reconhecer a informação nova e interagir de forma apropriada e estratégica (MARTIN; ROSE, 2007, p. 08).

O contexto de situação (ou registro) está relacionado à situação imediata de realização do texto. Halliday (HALLIDAY; HASAN, 1989) define três variáveis de registro, aqui apresentadas de acordo com a reformulação de Eggins e Martin (1997, p. 238):

²⁰Diferentemente de outras abordagens linguísticas, a noção de contexto utilizada deriva dos trabalhos do etnógrafo Malinowski.

- **Campo**, ou ação social: o que acontece, a natureza da ação social, aquilo em que os falantes estão engajados;
- **Relações**, ou estrutura de papéis: quem está participando do evento e com qual função, quais são suas relações de solidariedade;
- **Modo**, ou organização simbólica: qual o papel da linguagem, o que cada interactante espera da linguagem enquanto função.

Dentro da teoria de gêneros proposta por Martin (1992, 2000), o registro funcionaria como uma forma de realização do Gênero. O registro seria responsável pela realização das escolhas do falante no nível da linguagem instanciada pelas escolhas lexicogramaticais, ao passo que o gênero estaria relacionado às escolhas nas esferas dos processos sociais (VIAN JR.; LIMA-LOPES, 2005, p. 34). Isso faz com que o registro seja um nível de contexto necessário para a materialização do gênero, além de imprimir nele particularidades referentes à situação em que ocorre. Para Eggins (1994, p. 34), o gênero (contexto de cultura) é necessariamente mais abstrato que o registro (contexto de situação) por um simples motivo: o primeiro tem sua instancição obrigatoriamente mediada pelo segundo.

Assim, os textos variam não só em relação a seu contexto de produção imediato, mas também em relação a seu objetivo social. Essa forma de estruturação dos planos de contexto, chamados também de planos semióticos,²¹ sugere que as escolhas realizadas em um determinado plano devem se refletir nos demais. Por exemplo, ao escolhermos realizar uma negociação de compra em uma loja, essa negociação não poderá ocorrer de forma idêntica em todas as possíveis lojas: o preço e o tipo de artigo determinam parte das escolhas a serem realizadas, tanto no nível lexicogramatical como nos estágios. Por outro lado, ao se escolher fazer uma negociação e não uma reclamação condiciona a natureza de algumas escolhas, que deverão ser prototípicas.

²¹Na GSF o conceito de semiótico está relacionado à produção de significados; planos semióticos seriam níveis de abstração no processo de instancição de um significado. Como já colocado, a expressão de um determinado sentido não mora em apenas um dos planos, mas em todos simultaneamente.

A figura 3, retirada de Eggins e Martin (1997, p. 243), explica como esse sistema simbólico se organiza na GSF. O gênero é o nível mais abstratamente semântico e responsável pelas convenções de caráter social, o registro estaria relacionado ao contexto imediato de produção da simbologia da linguagem, sendo condicionado por variáveis específicas, como o Modo, o Campo e as Relações, já explicadas acima. Por fim, temos a língua, que em sua realização conta com os sistemas em cada uma das metafunções.

Ao se manifestarem linguisticamente, essas escolhas são instanciadas em Metafunções de linguagem; elas não são as únicas que realizamos no uso da língua, mas segundo Halliday (HALLIDAY, 1973; 1978; 1985; 1994; HALLIDAY; MATTHIESSEN, 2004), são funções essenciais para a realização da comunicação humana, estabelecendo uma espécie de núcleo duro (*main core*). Dada a natureza semântico-funcional da linguagem, escolhas condicionadas por essas Metafunções não são apenas de caráter estrutural, todas elas refletem fatores ideológicos e sociais do falante e seu contexto imediato e de cultura. Halliday propõe um

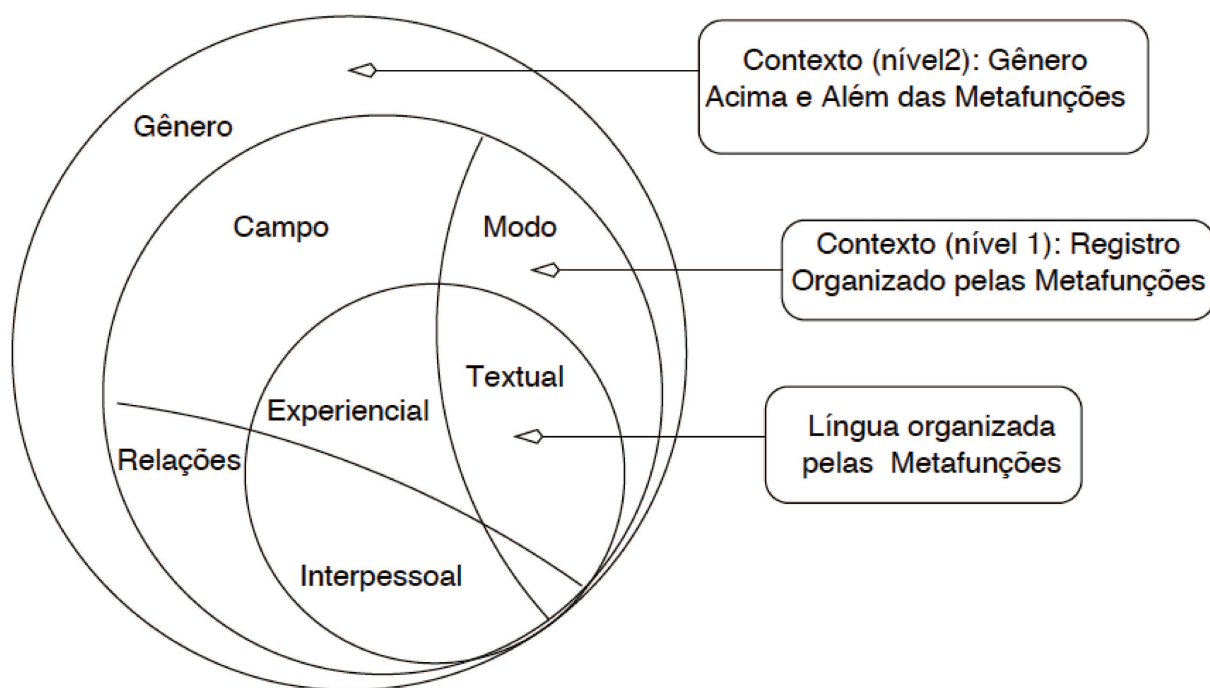


Figura 3: O gênero em relação ao registro e à Língua

modelo de análise tridimensional (IEDEMA, 2003), que não se fia apenas em um conjunto de regras exatas a serem seguidas ou refinadas. Essas três Metafunções representam, na verdade, sistemas disponíveis para realização das escolhas dentro das proposições, as quais, posteriormente, serão instanciadas. Essa característica leva a GSF a possuir uma grande flexibilidade analítica, que se dá graças à sua característica paradigmática: as escolhas não são analisadas apenas a partir de sua relação com outros elementos, mas sim como elas se relacionam comparativamente com as demais escolhas dentro de uma proposição. Esses contextos se manifestam em três grandes funções que realizamos com a linguagem, funções que expressam a forma como representamos o mundo, as relações de poder e solidariedade entre os indivíduos participantes da interação e a forma de como organizamos a linguagem para o seu uso. Essas Metafunções poderiam ser definidas da seguinte forma (aqui reescritas de acordo com a definição de Martin e Rose (2007, p. 07):

- **Metafunção Interpessoal:** engloba as atitudes em relação ao outro e os papéis sociais assumidos.
- **Metafunção Textual:**²² organiza o texto.
- **Metafunção Ideacional:** refere-se à representação de nossa experiência.

O método de análise marxista adotado por Halliday leva qualquer análise a estar profundamente calcada na realidade social do homem. Se, por um lado, essa característica dá à GSF uma forte capacidade analítica, uma vez que a semântica passa a ser o construto que explicaria a organização da linguagem humana, por outro, sua gramática é de difícil operacionalização. Cada uma das Metafunções possui um complexo sistema que visa a classificar os elementos simbólicos da linguagem. Dependendo do contexto de trabalho, o analista pode se perder na superficialidade da análise das estruturas sem necessariamente discutir as implicações sociais e culturais dessas escolhas.

²²A Metafunção Textual é de especial interesse para este trabalho, dado seu propósito de analisar o fluxo de informações nos textos videográficos que compõem o corpus de estudo. Por conta disso, gostaria de convidar o leitor para ler o capítulo 2, na página 85.

A importância desse modelo reside no fato de ele dar conta simultaneamente de uma série de significados e ações que, por vezes, extrapolam o componente linguístico. Isso integra a linguagem verbal a um amplo campo de análise que leva em consideração todos os elementos envolvidos na interação, integrando-a como parte da ação social (STUBBS, 1996). Essa possibilidade de análise simultânea entre diferentes formas de significado faz da GSF uma poderosa ferramenta, uma vez que as escolhas realizadas e instanciadas pelos falantes não são aleatórias (THOMPSON, 1994, p. 30). Como coloca Iedema (2003), o trabalho de Halliday permitiu que a análise linguística fosse além da mera aplicação de critérios sintáticos, oferecendo materialidade para diversas abordagens que visam à análise da ideologia. Todavia, o autor ressalta que um dos princípios mais importantes da gramática hallidiana é o fato da SEMIOSE não ser analisada em termos de blocos e estruturas, mas sim a partir de tensões sociais significativas, as quais podem ser instanciadas de diversas formas (IEDEMA, 2003, p. 31). Naturalmente, cada tipo de linguagem e modo deve contar com diferentes elementos que seriam responsáveis por essa realização. Kress e van Leeuwen (1996) estão entre os primeiros trabalhos a discutir como o princípio metafuncional poderia ser adaptado a outros sistemas semióticos que não as línguas naturais. Para os autores, que estão preocupados especialmente com o design da visualidade, apesar do modelo oferecido pela sociossemiótica ser abrangente, adaptações de caráter teórico, advindas de pesquisas sobre os diversos modos de linguagem, necessitam ser realizadas. Por conta disso, autores como Lemke (2006c, p. 1-2) sugerem que essa nomenclatura Metafuncional seja revista, de forma a ser mais abrangente. Entretanto, o trabalho de Lemke parece um tanto superficial no que tange à definição dos elementos que parecem compor cada uma dessas metafunções, talvez isso seja fruto do objetivo principal de seu trabalho, que é o estabelecimento de uma identificação entre as metafunções inicialmente pensadas para linguagem verbal e suas possíveis aplicações na análise da imagem. Neste trabalho, utilizo a terminologia de Lemke, todavia a enriqueço com algumas definições retiradas do trabalho de Dondis (1973):

- **Metafunção presentacional:** Os processos e o conteúdo informativo de uma semiose, como seria o caso das relações em nível representacional.
- **Metafunção orientacional:** Elementos abstratos que permitem a compreensão da relação entre os participantes de uma ação multimodal, ligada à nossa interpretação dessa composição.
- **Metafunção organizacional:** Estruturas de coesão e construção do significado da mensagem, como é o caso da composição de um quadro ou em uma foto.

De certa forma, o que garantiu a GSF se expandir e alcançar modos de linguagem que vão além da verbal²³ foi exatamente sua conexão de mão dupla entre significado↔estrutura↔ação social. Isso porque aqui a língua é vista como parte de um meio social mais amplo e, por conseguinte, da construção da realidade. Logo, seria esperado que esse mesmo nível de interação se desse entre diferentes tipos de sistemas; todos estariam integrados, influenciando uns aos outros. Dentro desta proposta a análise do discurso não mais privilegia a linguagem verbal, como principal via de análise das ideologias que permeiam nosso cotidiano. Como veremos em diversos momentos desta tese, um dos pressupostos dessa análise é que todos os sistemas semióticos presentes em um texto, seja ele predominantemente verbal, visual ou sinestésico, são igualmente importantes para construção do significado nele expresso. A multimodalidade busca ir além do plano estético, aprofundando-se no plano discursivo por se beneficiar da experiência adquirida nos estudos de análise do discurso e tornando possível estabelecer uma relação entre questões ideológicas em manifestações textuais que representam comunidades discursivas específicas. Visa-se, assim, ao levantamento de escolhas socialmente motivadas (HALLIDAY, 1992a; b; 1993b, 1994; HALLIDAY; MATTHIESSEN, 2004). A essa transposição teórica, deu-se o nome de análise multimodal. Essa transposição se fez necessária pela consequência da nova realidade, na qual a linguagem se tornou um elemento multifacetado que não mais prima pela supremacia da linguagem verbal (IEDEMA, 2003, p.

²³Neste trabalho utilizo o termo "linguagem verbal" como um termo guarda-chuva que inclui tanto a língua em sua variedade escrita como oral.

33). Contudo, vale ressaltar que, dentro de uma abordagem efetivamente funcionalista, todos os textos seriam multimodais, isso porque, de acordo com Baldry e Thibault (2006), qualquer realização linguística está relacionada a eventos não linguísticos, como gestos, expressões faciais (no caso da linguagem oral) e questões relacionadas a layout e tipografia (no caso da língua escrita), o que efetivamente questiona a supremacia daquele tipo de sistema semiótico (IVINS JUNIOR, 1969).

1.2 Transformações no contexto social e linguístico: em busca de uma teoria Sociossemiótica Multimodal

1.2.1 Por uma teoria abrangente da comunicação

Nesta seção, discutirei alguns pressupostos básicos da análise multimodal.

Como poderemos observar a seguir, esse teoria parece se dividir em dois momentos. No primeiro, temos o desenvolvimento de uma teoria baseada na necessidade de se advogar em favor de mudanças na forma de perceber o lugar da língua, em especial de sua variedade escrita, causada por transformações importantes na arena comunicacional. Nesse contexto, dá-se especial destaque a maior valorização do código visual, que passa a influenciar decisões dentro do próprio código escrito, como seria o caso do hipertexto. Essa realidade parece transcender o ambiente da microinformática e influenciar o mundo *off line*, uma vez que diversos meios de comunicação impressos, em especial jornais e revistas, tiveram seus projetos gráficos redirecionados para serem visualmente similares a seus postais, como é o caso dos jornais *O Estado de São Paulo*, o *Globo* e *Folha de S. Paulo*. Essa transformação, de uma lógica baseada no texto para uma lógica baseada na imagem deveria trazer, segundo o autor, uma série de modificações na forma como vemos o ensino de produção textual, que agora deveria estar relacionado a estas novas formas de significação (KRESS, 2002; 2005).

Em segundo momento, como também veremos a seguir, a teoria de Kress se transforma, agora o autor está interessado em elaborar uma teoria abrangente da comunicação, que dê conta das escolhas dentro de determinados modos de linguagem, assim como a combinação entre eles. Para o autor, essas relações são motivadas pelo nosso contexto cultural, que será responsável não apenas por moldar nossa percepção de linguagem, como também de nossa própria definição de seus usos. Nessa nova perspectiva, a ênfase deixa de estar nas transformações sofridas nos processos de comunicação e passa para a construção de significados. Os objetivos mais específicos de uma teoria sociossemiótica seriam garantir, de acordo com Kress (2010, p. 18):

- Que membros de uma comunidade tenham acesso aos recursos semióticos de forma a se beneficiar de seu processo de constituição de significados.
- Que os membros de uma comunidade coesa sejam capazes de contribuir para os propósitos comuns de sua comunidade, construindo semioticamente suas contribuições e as comunicando.
- Que os membros de uma comunidade tenham consciência dos efeitos de suas ações semióticas de linguagem.

Ou seja, uma teoria sociossemiótica que se baseia em uma análise multimodal teria por objetivo analisar a utilização de modos de linguagem dentro e uma determinada comunidade de forma a produzir uma consciência dos meios de produção de significado. Isso poderia trazer à baila algumas reflexões sobre como a teoria multimodal poderia influenciar e transformar nossa forma de pensar o processo comunicacional. Para Kress (2002), o conhecimento sobre os diferentes potenciais semióticos dos modos de linguagem de uma cultura traduziriam sua ideia do que seria letramento e justificariam o desenvolvimento de uma teoria sociossemiótica que buscasse entender essas relações.

1.2.1.1 A imagem e a tecnologia

Kress inicia o desenvolvimento de sua teoria motivado pelo bipé tecnologia/imagem. No caso do primeiro, Kress começa a constatar que a transformação operacionalizada pela vulgarização da microinformática leva a uma digitalização contínua dos meios de comunicação e do cotidiano, substituindo formas de fazer e produzir, antes manuais ou mecânicas. Nesse sentido, uma nova lógica capitalista surge (GEE, 1996a), com efeitos muito nítidos sobre a produção textual, sobre a leitura e sobre a forma como indivíduos estabelecem suas relações socioafetivas. Segundo Kress (2002, 2005), haveria uma consequência quase lógica: a criação de sistemas computacionais baseados em imagens, não apenas em linhas de comando — como foram os computadores até meados dos anos 80 —, levou a uma valorização da imagem como principal meio de expressão. Uma consequência seria uma contínua desvalorização do texto escrito e de seu meio principal de publicação, o livro.

Por conta disso, Kress (2005) está especialmente preocupado em advogar em favor de uma teoria multimodal que dê conta das modificações que a tecnologia e as novas formas de significar possibilitadas não só por ela, mas também pela crescente valorização da linguagem visual. Consequentemente, sua análise está sempre centralizada na forma como a produção e compreensão de textos multimodais é diferente dos textos ditos monomodais, uma vez que, para o autor, essa modificação é responsável por transformar nossas formas de interação, trazendo perdas e ganhos nos processos comunicacionais de nosso cotidiano. O autor exemplifica as modificações no nível da leitura e compreensão desses novos textos comparando a página de um livro e um sítio de internet, demonstrando quais as principais diferenças em relação a ambos.

Kress dá um importante passo ao colocar que essas modificações na tipologia textual podem levar a profundas modificações na forma como, por exemplo, devemos pensar estratégias de leitura ou de produção textual. Essa constatação serve para introduzir um dos elementos centrais do argumento de Kress: para ele, entre os ditos ganhos da nova arena tecnológica está a possibilidade de se utilizar modos de linguagem e mídias que sejam mais

propícios para a comunicação de um determinado conteúdo. Além de uma possibilidade de aumento da capacidade midiática da comunicação, uma vez que mídias digitais possibilitam a inclusão desses infinitos recursos, temos uma grande mudança na forma como os indivíduos passam a encarar o texto e, conseqüentemente, os modos de linguagem que podem estar ali presentes. Isso implicaria uma grande transformação de paradigma: o fato de a escrita deixar de ser o eixo central da comunicação faz com que todas as nossas práticas educacionais e analíticas tenham de ser revisitadas, pois a desvinculação entre a língua e a propagação dos sistemas e eventos comunicativos cria uma nova lógica de pensamento para a qual devemos estar atentos (KRESS; VAN LEEUWEN, 1996; 2002; KRESS, 2003; 2006). Nesse sentido, acho que há um certo fetichismo tecnológico por parte de Kress. É certo que a liquidez e velocidade desse novo meio necessita de uma instrumentalização que permita ao indivíduo entender que essas relações se modificam em um ritmo bastante rápido; ao mesmo tempo não é o conhecimento técnico que parece garantir essa compreensão. Outrossim, pensar que esse dinamismo natural ao hipertexto é uma vantagem sobre os textos impressos me parece uma valorização demasiada. Penso que a resposta estaria em um modelo que, como sugerido por Lemke (LEMKE, 2006b), garantisse a diversidade das possíveis interpretações midiáticas, instrumentalizando o indivíduo para que sua relação com esse meio não se baseie em fetichismo tecnológico (MACHADO, 2007a).

É importante observar que a relação entre linguagem e tecnologia talvez não seja tão imediatista como a colocada por Kress nesse momento do desenvolvimento de sua teoria. Kress (2001; 2005) e Iedema (2003), em especial, colocam a década de 80 do século XX como ponto crucial para o crescimento da imagem como modo mais influente no processo comunicacional. De certa forma, poderíamos dizer que tais autores estão certos, dado que é a partir dessa década, com a invenção da computação pessoal pela IBM, do ambiente gráfico pelo ex-sócio da Apple, Steve Wozniak, e do protocolo WWW, pelo professor do MIT Tim Berners-Lee, que formas de produção textual com inspiração visual tiveram um grande salto.

Todavia, pode-se analisar essa questão com um pouco mais de atenção. O que parece falhar no argumento da teoria multimodal, nesse momento, é tanto uma observação mais apurada da relação entre texto e imagem, especialmente no âmbito histórico, como o reconhecimento da instabilidade interpretativa e maleabilidade do código visual (MCDONAGH, GOGGIN et al., 2005). Ou seja, eles estão certos em assinalar que esse período histórico como importante para essa relação, mas vale ressaltar que a produção textual está longamente relacionada com a tecnologia e com a linguagem visual. Essa relativização é necessária para que possamos efetivamente entender que essa tensão sempre ocorreu, em diferentes graus na sociedade ocidental.

De certa forma a relação: escrita↔imagem↔tecnologia sempre esteve presente. Uma observação dessa relação a partir de autores que acreditam em seu caráter histórico social, remeteria à escrita e à imagem como um sistema de construção ideológica, relacionado a própria lógica da construção de nossa sociedade ocidental; a segunda, a uma visualidade inerente à língua escrita (Ferreira, 2011).

Encarar a relação entre a escrita e a imagem sob uma perspectiva histórica nos ajuda a perceber que a realidade é um pouco mais complexa que a dicotômica separação que temos na literatura multimodal, uma vez que o papel social realizado por cada uma delas é fruto de um contexto socioideológico. Em seu livro *Filosofia da Caixa Preta*, o filósofo tcheco Villén Flusser (2002) constrói um interessante argumento sobre essa coexistência; para ele, os sistemas de comunicação imagético e o escrito representam duas formas diferenciadas de organizar o mundo a partir de diferentes concepções ideológicas. A imagem, anterior à escrita, é um processo de abstração do mundo em pelo menos dois níveis, primeiro por bidimensionalizá-lo, e segundo por nos obrigar a utilizar imagens em relações significativas com o mundo. Em outras palavras, imagens, que nascem como registro visual da experiência passam a ser a própria experiência, uma vez que elas são nosso único contato com determinadas situações de vida, o que parece ser comum a todos os signos. Contudo, essa experiência imagética, de acordo com o observado por Flusser, não pode ser considerada inequívoca ou denotativa,

todo processo imagético é aberto a múltiplas interpretações, dependentes em especial das referências de seu leitor. As imagens, teriam um tempo de leitura circular, sendo remetidas a si mesmas: seriam lidas em camadas de observação, cada uma construindo um nível mais profundo e mais peculiar. Além disso, talvez fosse correto afirmar que a imagem tem um grau de referencialidade que é um tanto aberto, pois existe uma menor convencionalidade em relação à escrita. A escrita, por sua vez (FLUSSER, 2002, p. 10) é apoiada em uma codificação mais exata, de caráter linear e conceitual, afastando-se ainda mais do mundo concreto: ela é um processo de representação no qual a significância não está ancorada em eventos particulares, mas em ideias que criam um metacódigo. Enquanto as imagens se baseiam na representação de cenas do mundo real, a escrita coloca-se como uma tradução desse mundo em conceitos. Logo, a escrita me parece algo mais codificado, com significação mais convencionalizada socialmente.

Flusser aponta a Idade Média como momento crucial em que essa oposição começa a se firmar, o código escrito e o visual passam a representar diferentes facções ideológicas, o primeiro estaria ligado à erudição de um cristianismo intelectualizado, ao passo que a última passa a significar um paganismo imagético a ser combatido pela nova cultura ocidental que se estabelece. Na Idade Moderna, essa oposição ganha novas cores: as sociedades de cultura escrita e cientificista contra as menos evoluídas e, por conseguinte, visuais.²⁴ A construção da nossa sociedade, para Flusser, estaria diretamente relacionada à suplantação do primeiro modelo sobre o segundo, a escrita e o cristianismo vencem decretando fim da idolatria em favor da textolatria, a ciência e a exatidão vencem as sociedades imagéticas inferiores (FLUSSER,

²⁴Tenho consciência de que essa afirmação pode ser questionável, uma vez que nem todas as comunidades sem escrita necessariamente recorrem a códigos visuais complexos para comunicação, o que levaria a necessidade de uma relativização dos conceitos de Flusser. Todavia, a reflexão de Flusser parece se firmar na dicotomia que se estabelece na Europa, em especial durante a Idade Moderna. Nesse sentido, haveria um grupo letrado — por conseguinte cientificado e sem a necessidade de imagens — e o não letrado — a maioria da população que por não ter acesso à linguagem escrita ainda dependia das imagens para sua educação litúrgica, na maioria das vezes católica.

Escolhi utilizar Flusser como parâmetro para questionamento do trabalho de Kress porque ambos parecem refletir a partir de uma perspectiva bastante semelhante, o mundo letrado ocidental de tradição eurocêntrica. Naturalmente, Flusser também necessitaria ser relativizado, afinal sua oposição entre a cultura escrita e imagética parece criar um universo de separação quase absoluto, sem nenhuma variação.

2002, p. 10). Isso significa que a palavra, torna-se o modo preferencial de definição de nosso raciocínio técnico-científico (FLUSSER, 2007a), pois ela é, em muitos aspectos, mais objetiva que uma imagem. Dessa forma, a palavra passa a buscar a si mesma como elemento definidor dessa abstração, procura-se conscientemente a língua como elemento de produção e criação de relação entre os diversos elementos, tudo isso em conformidade com a estrutura que ela oferece naquele momento histórico-científico. A construção da subjetividade nesse sistema de valores se dá pela frustração na realização das ideias articuladas pelo falante, é a língua em estado nascente em suas mais diversas reformulações (FLUSSER, 2007a, p. 138-142).

Mas apesar da vitória dessa forma de pensamento e construção da realidade, o texto nunca conseguirá sobreviver em isolamento da imagem. Eles passam a coexistir, a escrita ganha o status preferencial, ao passo que as imagens passam a ser um sistema ilustrativo e dependente da palavra. Isso se deve ao fato deles estarem em uma relação dialética: à medida que a sociedade, cristã, por exemplo, combate o paganismo, ele passa a absorver elementos imagéticos e se paganiza; enquanto a ciência combate as crenças, absorve elementos visuais em suas descrições e se torna idólatra (FLUSSER, 2002, p. 10). Essa complexa situação faz com que a língua inclua em si cada vez mais elementos visuais, em um processo que Flusser (2007a, p. 166-180) chamaria de amplificação da língua escrita; no caso das línguas alfabetais/não alfabetais e flexionais/isolantes, elas passariam a assimilar elementos de amplificação em pelo menos dois níveis: a música, responsável por adaptações prosódicas, e plásticas, relacionadas a adaptações visuais. Na primeira, estariam modificações fonéticas como acentuação tônica ou musical, tonalidade de fala etc., ao passo que na segunda estariam a criação de diversos tipos de sistema de escrita, novos símbolos gráficos, entre outras mais.

Mas por que essa necessidade de adaptação? Somente porque mesmo com todo o seu poder de abstração a língua não é, e talvez nunca tenha sido, capaz de suprir todas as necessidades significativas da comunicação, criando uma constante necessidade de criação de códigos semióticos cada vez mais complexos ou mesmo de produções significativas que agreguem mais de uma semiose na produção de seu significado (IVINS JUNIOR, 1969).

Nesse sentido, os textos nascem para tentar rasgar as imagens e objetivar nossa relação com o mundo, fato que não acontece, o que temos é efetivamente uma influência cada vez mais profunda das imagens nos textos (FLUSSER, 2002, p. 11). Essa questão também é discutida por diversos estudos, entre eles podemos citar o de Prior (2005), cujo trabalho faz uma extensa discussão de obras relacionadas ao design gráfico de forma a demonstrar que a relação entre textos e imagens existe antes do surgimento das tecnologias digitais. Segundo Prior, estabelece-se uma relação de construção de significados que é conjunta e dependente de ambos os códigos semióticos para sua construção. Na verdade, a instanciación de significados é algo que vai além da simples justaposição de diferentes modos, eles parecem trabalhar conjuntamente e de forma simbiótica. O design seria um exemplo disso, nele diversos sistemas semióticos sempre coexistiram (inclusive o escrito se pensarmos em design gráfico) e são combinados na produção de uma complexa forma de significar (FLUSSER, 2007b). Mas qual a consequência dessa perspectiva para a massa iletrada? Sabe-se que enquanto essas discussões que articulam o papel da escrita e da imagem estavam acontecendo, uma parte significativa da população europeia ainda não tinha conhecimento da linguagem escrita; por conta disso, a produção da imagem não desaparece, mas ganha um papel litúrgico e didático, elas estão presentes para servirem a um processo de facilitação, possibilitando o acesso daqueles que não podiam ler as escrituras sagradas. Se por um lado isso garante a presença da imagem, por outro isso parece subjugar a à escrita: ela se estrutura como meio intelectualmente superior de transmissão de conhecimento.

De fato, é sabido que a valorização da imagem ganhou um impulso avassalador desde o fim da Segunda Guerra Mundial (HORKHEIMER; ADORNO, 2002), a nova ordem capitalista cria uma indústria de consumo de bens culturais que vai, cada vez mais, se distanciando das formas de entretenimento não baseadas na transmissão para grandes audiências. É estabelecido, assim, um movimento que deixa de valorizar uma cultura local e unificada em favor de uma cultura destilada e remodelada de acordo com as vontades de um público massivo e impessoal, fazendo com que a antiga cultura local comece a perder espaço para

formas de propagação de informação baseadas em formatos pré-moldados por essa indústria, como foi do cinema e, posteriormente, da televisão.²⁵ A criação desses formatos, associada a uma produção em série dos bens culturais, barateia e acelera a produção, que também usa a imagem como o elemento facilitador; as adaptações cinematográficas de livros, contos ou histórias populares são vistas como preenchedoras de vazios interpretativos. Assim, alguns produtos audiovisuais — e os vídeos de publicidade desta tese nele se incluem — seriam responsáveis por traduzir conceitos de forma palatável para a massa consumidora desses produtos; em outras palavras, o público, recebe informação já pronta e facilitada pelos meios de propagação da informação. Como veremos nos capítulos referentes à análise de dados (ver capítulo 5, na página 163, e capítulo 6, na página 213), a campanha analisada traduz de forma simplificada conceitos computacionais de forma a incentivar a compra do produto/serviço anunciado.

É importante pensarmos que o papel da linguagem nessa nova realidade deve ser pensado além da simples compreensão de como indivíduos interagem com os meios digitais ou da utilização dessas mídias dentro do processo educacional (BUCKINGHAM, 2008, p. 73-74). Não que essas possibilidades não tenham sua importância, mas temos que efetivamente pensar esse processo a partir de uma série de questões éticas, tecnológicas e políticas (LEMKE, 2006b). Em outras palavras, não podemos pensar nas mídias digitais apenas como informação ou tecnologia, hoje indivíduos se desenvolvem em uma cultura que se mistura com esses meios.

1.2.1.2 Produção de significado

Em trabalhos como Kress (2010), por exemplo, o autor começa a desvincular a necessidade de uma análise multimodal do surgimento das chamadas novas tecnologias — que afinal

²⁵Naturalmente existiam modos de comunicação escrita que também participaram desse processo de industrialização, sendo representados em especial pelos jornais e semanários; todavia, com o crescimento da televisão e, posteriormente, com a internet, esses veículos perderam muito de sua força, chegando por vezes a descontinuar sua versão impressa. Exemplos desse tipo de transformação são o jornal *Notícias Populares*, que deixou de circular graças à pressão de programas sensacionalistas da televisão aberta (CAMPOS JR, MOREIRA et al., 2002), *Gazeta Esportiva*, antigo jornal de esportes paulistano agora disponível apenas na forma de portal, ou mesmo o renomado *Jornal do Brasil*, que anunciou recentemente que deve retirar sua edição impressa das bancas e permanecer apenas na internet (GASNIER, 2010).

nem são tão novas assim, uma vez que a internet já está completando 15 anos de existência — e a relaciona diretamente com as capacidades expressivas de cada modo de linguagem, dadas as possibilidades e limitações impostas pela cultura em que se encontram. O papel da imagem e o da língua também passam a ser revistos, agora a relação entre os meios é tratada com maior neutralidade, respeitando os fatos históricos dessa relação. Essa mudança é importante, pois mostra que Kress relativiza as relações entre os diferentes modos e passa a observar com mais cuidado a arena ideológico-capitalista. Ao mesmo tempo, Kress torna sua teoria mais abrangente, uma teoria da comunicação responsável por descrever o processo de construção do significado. Busca-se, assim, entender como os diferentes contextos sociais, assim como individuais, influenciariam no processo de significação. Esse processo passaria por uma instância social, no qual indivíduos de um determinado grupo dividiriam conhecimento e uso de determinados modos de linguagem. Esse conhecimento incluiria não apenas a forma como uma certa comunidade constrói os processos de combinação e construção da mensagem em um determinado modo, como também de como as relações entre os diversos modos computacionais poderiam acontecer. Aqui também estariam incluídos a valorização que cada um dos modos de linguagem possuiriam ao serem utilizados na construção do significado.

Dentro desse contexto, sua teoria, aos poucos, vai ganhando uma relativa independência do modelo hallidiano. Em Kress (2010), pode-se observar que o autor já busca uma teoria sociossemiótica quase autônoma, que ecoa a teoria fonte em quatro pontos-chave, todos já discutidos anteriormente: a) a motivação dos signos e das escolhas de linguagem, b) do arranjo como elemento de produção de significado, c) da existência de diferentes níveis (ou *strata*) que condicionam as escolhas em todos os modos de linguagem, d) da realização (ou instanciação) como construto que materializa níveis (ou *strata*) mais abstratos em escolhas efetivas no sistema e e) os conceitos de constituição, função e sistema, essenciais para a definição da gramática de cada modo de linguagem.

Kress justifica sua postura teórica observando a influência do contexto e dos usos situacionais e sociais que a construção de um signo possui. Isso advém do fato de Kress (2010,

p. 62-67) acreditar que em uma teoria sociossemiótica signos não são utilizados, mas criados a partir das necessidades comunicacionais por um falante que realizará uma conjunção entre significado e forma, de acordo com seu interesse (KRESS, 2003). Esse processo é contínuo, cada falante "dá nova forma" constantemente a sua construção simbólica, utilizando-se sempre dos recursos culturais de forma a melhor lidar com seu contexto social. Ou seja, o modelo de Kress não deixou de se fiar na existência de uma estrutura expressiva pré-codificada — definível como os diferentes modos de linguagem — que é abstrata e está à disposição do criador de signos (*sign maker*) responsável por utilizá-lo em um processo de significação. Uma diferença importante, todavia, parece ser o fato de o modelo desenvolvido por Kress não estar preocupado com a instanciamento dos significados em escolhas ditas universais, mas sim únicas em cada contexto.²⁶

O conceito de arranjo introduzido por Kress (KRESS, 2010; KRESS; VAN LEEUWEN, 2001; 1996) parte da noção de fraseado introduzida por Halliday (1985, 1994). Halliday define fraseado (*wording*) como a combinatória linguística de uma proposição, sendo esta o resultado de um processo motivado de significar. Não haveria, por conseguinte, escolhas não motivadas ou aleatórias (vistas como sinônimos dentro dessa teoria). Dessa forma, ao constituirmos uma proposição, estaríamos realizando fraseados que expressam funções que espelham nosso mundo exterior e interior, nosso relacionamento com indivíduos — estes presentes ou não na interação — assim como os princípios de organização da mensagem. A ideia central é mantida na teoria multimodal, mas cresce em importância. O arranjo é descrito como responsável pela materialização do significado durante seu processo de realização; o

²⁶ Apesar da grande importância que Kress dá a produção de signos não seria importuno observar que sua teoria, de certa forma, dá pouca ênfase ao cabedal de representações sociais já existentes e seu papel para a criação de significados localmente desenvolvidos. Kress parece deslocar esse "ponto de partida" para os diversos modos de linguagem, eles seriam nosso material comum para o processo de representação e comunicação: ao escolhermos o modo de linguagem, estaríamos também escolhendo convenções representacionais que nossa sociedade carrega. Esse acordo social que versa sobre o que deveria ser um modo de linguagem, não poderia, também ser considerado um repositório de signos-base para início da produção local de significados? Em outras palavras, acredito que esse conjunto de significados já comumente estabelecidos seja importante alicerce (*common ground*) para uma transformação de seu significado em nível local. Por conta disso talvez acredite que um modo de linguagem possa também ser encarado como um repositório de elementos significativos que dão o insumo para criação de signos no Registro de uma interação cotidiana.

significado é moldado em matéria de acordo com o ambiente social que é produzido. Os diferentes arranjos são formas de expressão que carregam consigo convenções sociais e ideológicas, como uma matriz que oferece possibilidades de realização. Esse sistema permitiria não apenas o processo de organização dentro de um modo, mas também entremodos. Ao escolher determinado modo de linguagem para significação, está-se realizando um arranjo que leva em conta as possibilidades semióticas do modo, as necessidades de expressão do contexto. Naturalmente, seria plausível pensar que o conceito de arranjo também poderia ser ampliado para uma relação entre os diversos modos de linguagem que constituem um ato comunicacional. Uma vez que o significado não é construído por apenas um sistema semiótico, a escolha em diferentes modos de linguagem deve ser pensada de forma constituir, também, um arranjo multissemiótico. Dessa forma, ao escolhermos quais elementos deverão compor uma determinada mensagem, estaríamos também escolhendo quais elementos de cada modos de linguagem serão combinados de forma a, juntos, constituir a mensagem.

As noções de escolha e sistema se tornam especialmente complexas quando pensamos em termos de uma análise multimodal. Isso porque leva-se em conta não apenas um sistema de escolha, mas diversos que ocorrem em paralelo; cada modo de linguagem tem um próprio sistema de escolhas. Um exemplo desse tipo de realidade são os vídeos que analiso neste trabalho, como já discutido anteriormente (ver figura 2, página 2): o vídeo congrega uma série de modos de linguagem; escolhas em diferentes modos começam a influenciar umas as outras, criando uma rede de significados complexa. Ao se realizar uma escolha no roteiro, está-se também influenciando o enquadramento, a fotografia e a gravação; se tal escolha não for traduzível para algum desses meios, ela terá de ser repensada. Por conta disso, se criam roteiros pós-gravação, para que se tenha uma real tradução daquilo que foi realizado em vídeo para a linguagem escrita.

Entretanto é importante observar que mesmo outorgando uma grande importância ao componente social, essa teoria dá ao indivíduo, criador de signos (*sign maker*), um papel bastante importante. É ele que, em última instância, toma a decisão composicional única e

motivada que um signo representa. Isso é fruto de uma complexa operação que parece ser resultado de:

- Processos simbolizantes
- Decisões composicionais
- Interesse do *sign maker*
- Processos de tradução

Para Kress (2010, p. 49) o processo de simbolização e o de comunicação são práticas distintas. A primeira estaria relacionada ao universo interno do criador de signos, sua forma de compreender o mundo que o cerca, fazendo dele sentido. Esse processo é eminentemente mental e condicionado por questões geracionais, de gênero, profissionais de educação formal, entre muitos outros. Ele está diretamente relacionado ao conceito de interesse que poderia ser definido como fruto dessas relações internas, conjugadas com a história de experiências pessoais, as quais necessariamente estão ligadas ao tempo de vivência em uma comunidade. São esses fatores que, segundo a teoria, farão com que o falante (criador de signos) configure suas decisões composicionais, num esforço de tradução do significado para o outro, ação que se dá em um processo externo ao falante. Ele também é responsável pela determinação da atenção que os interactantes terão em relação aos processos simbólicos e comunicacionais que participam, fazendo com que a simbolização e a comunicação sejam únicos de cada contexto.

Em outras palavras, essa teoria assume que o processo de simbolização é um constante ato de tradução e metaforização.²⁷ Nos trabalhos de Kress (2010, p. 124) a tradução é uma ação pela qual o significado é transportado de um contexto ou de um modo de linguagem para outro, algo muito similar ao conceito de ressemiotização já introduzido por Iedema (2003). Quando a rearticulação de significados envolver mais de um modo de linguagem,

²⁷A despeito de citar a metáfora como princípio básico de sua teoria, Kress parece não definir um conceito claro, deixando muito a cabo do leitor a definição desse conceito. Por conta disso, o presente trabalho, mesmo não tendo como parte do seu escopo uma reflexão sobre o conceito de metáfora, utiliza-o como definido por Lakoff; Johnson (1980), que é tratado com mais detalhes na seção 5.1, na página 164.

esse processo é chamado de transdução, termo que parece ser importado ou da física ou da biologia, significando a transformação de energia em outra de natureza diferente ou a transferência de DNA de uma célula para outra, como fazem os vírus. Já o processo de transformação é a tradução que envolve apenas um modo de linguagem.

A metaforização, por seu turno, seria responsável por ajudar na simbolização e comunicação, por tornar os significados palpáveis em ambos os casos. Apesar de ser colocado como importante na teoria multimodal, pouco é discutido por Kress ou outros autores, mesmo teorias que poderiam ser adotadas como ancilares pouco são discutidas.

Kress e van Leeuwen (1996) desenvolvem um trabalho que visa à interpretação de elementos visuais, seu objetivo principal é observar que sistemas semióticos visuais e verbais são organizados de forma semelhante, demonstrando como o código visual é utilizado na produção de significado. De maneira geral, os autores conseguem estabelecer um paralelo entre o código verbal e o imagético, seu trabalho foi seminal para os estudos de multimodalidade, uma vez que os autores conseguem demonstrar que diferentes sistemas sociossemióticos possuem uma estrutura organizacional semelhante. Isso se dá porque nossos códigos são expressões dos nossos valores e culturas, fazendo com que uma série de escolhas similares sejam realizadas dentro desse contexto. Uma das principais contribuições desses autores é mostrar que o conceito de Metafunção aplicável a uma gramática estratificada estabelecido por Halliday (HALLIDAY, 1973; HALLIDAY, 1978; 1985; HALLIDAY; HASAN, 1989; HALLIDAY; MATTHIESSEN, 2004) é aplicável à linguagem imagética. Apesar de, logicamente, imagens e língua possuírem suas especificidades, a iniciativa tem impacto por mostrar que usamos recursos muito similares na construção de qualquer tipo de mensagem utilizando qualquer tipo de sistema.

A teoria desenvolvida por esses autores dá ao analista a possibilidade de observar os diferentes modos de linguagem de uma forma bem peculiar: composições passam a ser proposições visuais que surgem a partir de um conjunto de escolhas possíveis, governadas por ideologias e convenções (MACHIN, 2007, p. xi). Entre as principais semioses que já fo-

ram estudadas na literatura, encontramos: imagens estáticas (Kress; van Leeuwen, 1996), artes plásticas (O'TOOLE, 1994), arquitetura (MENG, 2004; O'TOOLE, 2004), filmes e vídeos (VAN LEEUWEN, 1985; IEDEMA, 2001; VAN LEEUWEN e JEWITT, 2001; BALDRY, 2004; O'HALLORAN, 2004; BALDRY; THIBAUT, 2006; KLINGER, 2006; MACHIN; JAWORSKI, 2006; ALMEIDA, 2006; BÖHLKE 2008; MOURA, 2011), legitimação de práticas sociais (VAN LEEUWEN, 2007), análise de interações em ambientes profissionais e não profissionais (NORRIS, 2002; 2006), hipermídia (LEMKE, 2002; BALDRY; THIBAUT, 2006; BRESSANE, 2006; TOLEDO FILHO, 2008) e relação entre textos e imagens na grande mídia (Ferreira, 2011), não podendo se esquecer das óbvias consequências que o estudo deste e diversos outros modos de linguagem e mídias podem ter para os processos de letramento (BEZERRA et. al. 2011).

Para Machin (2007), a abordagem multimodal nasce da necessidade de ir além do estudo dos efeitos que determinados modos causam em seus leitores. O autor coloca que a multimodalidade dá ao pesquisador a possibilidade de ampliar o horizonte de pesquisa, permitindo a análise de signos em combinação em um determinado contexto, ao invés de se preocupar apenas com os significados gerais dos diferentes componentes de um sistema de representação. Isso significa que a multimodalidade não trabalha com sentidos estanques e determinados: o valor de cada signo advém de sua associação com outros, levando-se sempre em conta o momento em que essa associação é realizada e percebida. Por conseguinte, o significado de um dado elemento não está apenas nele, mas sim no sistema de que faz parte, uma espécie de gramática que lhe dá o potencial de significar (MACHIN, 2007, p. ix).

Logicamente existem diferenças e questões específicas de cada sistema semiótico que influenciam na forma como esses códigos são combinados na construção da mensagem. Um erro de interpretação comum é pensar que a análise multimodal se presta apenas ao trabalho com manifestações textuais relacionadas ao que comumente chamamos de multimídia (vídeos, CD-ROMs, sítios de internet etc.); em termos gerais, pode-se afirmar que essa abordagem sociossemiótica satisfaz o anseio de pesquisadores que procuravam ferramentas para

analisar esses novos textos, vinculados por meios ainda pouco estudados e de características nada perenes. Todavia, um fato a ser levado em conta é que dentro da definição do que seja uma interação multimodal pressupõem-se, também, interações cotidianas, uma vez que esses textos também são multimodais em sua essência (KRESS; VAN LEEUWEN, 2001; BALDRY; THIBAUT, 2006). Para Halliday (2005/1979), cada forma de expressão linguística ou fraseado é repleto de potencial de significados (*meaning potencial*), conceito já definido anteriormente, os quais podem ou não ser trazidos à tona em uma interação. O que determinará quais deles efetivamente estarão presentes é o conjunto de funções, normalmente representado pelo componente semântico da linguagem. Isso quer dizer que o significado e a estrutura daquilo que falamos estão profundamente interligados, havendo uma interação quase que determinante entre eles (VAN LEEUWEN, 2005, p. 3-5).²⁸

1.2.2 Refletindo sobre a articulação de significados em textos multimodais

Kress e van Leeuwen (2001, p. 4-22) discutem a articulação no processo de construção do significado para textos multimodais. Para os autores, o modelo linguístico tradicional parte do conceito de dupla-articulação — inicialmente estabelecido por Marinet, esse construto determina que a linguagem articular-se-ia enquanto forma e conteúdo — de forma a criar um modelo de construção de sentidos monomodal, no qual o significado seria produzido, nas palavras de Kress e van Leeuwen, “de uma vez”. Ou seja, ambos os níveis articulatórios estariam trabalhando simultaneamente, numa abordagem em que o sentido não necessariamente

²⁸Um bom exemplo seria o encarar: na cultura brasileira, olhar diretamente nos olhos do indivíduo é um requisito básico para o estabelecimento da interação, pessoas que não olham nos olhos são vistas como esquivas, insinceras e não confiáveis. Em outras culturas, como a japonesa ou a inglesa, olhar continuamente nos olhos de alguém pode ser visto de forma negativa. O mesmo pode ser dito em relação à nossa proximidade física ao conversar: notoriamente observamos que estrangeiros tendem a se afastar sistematicamente de nós em uma interação informal, isso porque nós valorizamos essa proximidade física e não a vemos como uma espécie de invasão da individualidade alheia. Ao mesmo tempo, em uma discussão, essa proximidade física pode ser vista como um ato de agressão, mesmo entre nós brasileiros. Isso faz com que a escolha em qualquer modo de linguagem não seja aleatória, nossos recursos semióticos são condicionados pelos significados que tentamos expressar.

se daria pelo acúmulo dos significados específicos dos diferentes níveis. Os significados em uma perspectiva multimodal poderiam seguir esse princípio, uma vez que ela tem por pressuposto a interação entre os diferentes níveis e códigos de linguagem, criando uma espécie de múltipla articulação.

Num primeiro momento os autores questionam a hierarquização dos níveis sugerida pelos trabalhos de Halliday, argumentando em favor de um sistema no qual esses níveis se integram de forma semelhante na produção do significado (KRESS; VAN LEEUWEN, 2001, p. 4). Apesar dessa opção, Kress e van Leeuwen continuam a se utilizar do termo nível (ou *stratum*), por acreditarem na potencialidade descritiva do modelo de Halliday. Mesmo assim, coloca-se que o modelo hallidiano, fortemente arraigado na produção verbal, necessitaria de adaptações a fim de abarcar um número maior de semioses. Na proposta dos autores, a criação de textos multimodais passaria por um processo de articulação quádrupla que seria responsável por constituir quatro tipos de significados simultaneamente, eles seriam o Discurso, Design, Produção e Distribuição.

1.2.2.1 Discurso

Para Kress e van Leeuwen (2001, p. 4), o discurso pode ser definido como conhecimentos socialmente construídos a respeito da realidade, ou, pelo menos, de algum aspecto dela. Esses conhecimentos seriam típicos de determinados espaços sociais, refletindo os interesses dos indivíduos que neles interagem. Esses espaços podem variar de acordo com seu tamanho e amplitude: podemos falar tanto da cultura ocidental como um todo, assim como do discurso familiar. Dentro desse bojo, também encontramos contextos bastante institucionalizados, como empresas. Um exemplo desse tipo de *stratum* seriam as empresas de mídia: não é incomum observar que diferentes jornais realizam diferentes coberturas do mesmo fato, dando interpretações políticas bastante distintas para eles. Esse tipo de fenômeno faz com que cada indivíduo possua uma variedade grande de discursos para se filiar, escolhendo aquele que é mais adequado para sua situação comunicacional (KRESS; VAN LEEUWEN, 2001, p. 21).

O discurso talvez seja o mais abstrato de todos os *strata* definidos pelos autores. Todavia, é relevante observar que a experiência de linguagem não é abstrata, ela se manifesta em produções pontuais, produções essas que, de diferentes formas, espelham esse recorte de mundo. É, por conseguinte, correto afirmar que o discurso não apenas se manifesta, mas é também modelado por práticas sociais (VAN LEEUWEN, 2005, p. 104).

Partindo da análise de revistas britânicas, principalmente da área de decoração, e presumindo que toda a ação semiótica é uma ação social, Kress e van Leeuwen demonstram que uma associação entre textos e fotos criam uma situação ideológica que reafirma determinados valores defendidos por essas publicações. Entre esses valores, está uma mudança na forma de classificação e organização de seus leitores: ao invés de dividir os indivíduos por classe social — um conceito notadamente capitalista de inspiração marxista —, os indivíduos agora são definidos pelo estilo de vida que levam (KRESS; VAN LEEUWEN, 2001: 34). Isso gera a definição dos produtos como signos de sua individualização e de um estilo determinado, sendo que isso cria uma pressão por diferenciação em cada indivíduo através de suas práticas semióticas.

Da mesma forma que essa busca pelo individualismo gera a criação de espaços semióticos individualizados, esses indivíduos estão passando por um processo de desindividualização, uma vez que a busca por esses espaços semióticos faz com que ele esbarre em outros que buscam a mesma solução em um espaço similar. Dentro da análise de mídia, conceito é denominado segmentação (MORAES, 1998), tendo sido inicialmente criado para definir a oferta de produtos de mídia (como canais de TV a cabo, rádios etc.) voltados para grupos de interesse específico (como esportes, filmes cult, culinária, moda etc.). Hoje tal conceito também nos ajuda a explicar o fato de as pessoas que consomem esse tipo de mídia formarem comunidades que se definem pelos traços semióticos de seu consumo informacional. Por conseguinte, para Moraes (1998), há um fenômeno de massificação segmentada, na qual grupos massificados são reorganizados e redivididos em subgrupos de consumo de determinado estilo de vida, propiciando a criação de comunidades específicas para cada segmento. Segundo Moraes, esse

tipo de sistema cria uma impressão de individualização no espectador, o que, na verdade não ocorre, uma vez que ele está participando de um novo grupo de pseudoindividualizados.

Se pensarmos nos vídeos que compõem o corpus analisado nesta tese, eles têm por objetivo produzir um efeito muito similar ao descrito pelos autores, ou seja, diferenciar e definir os consumidores de produtos Apple a partir da sensações de superioridade e de exclusividade que a aquisição de máquinas e softwares dessa marca pode significar. Mais do que uma simples aquisição de bens materiais, adquirir produtos Apple significa participar de um seleto grupo de indivíduos que agregam esses valores a sua definição de identidade. Kahney (2004) em seu livro *The cult of Mac* explora o impacto que a empresa e seus produtos possuem na cultura popular do ocidente. Para tanto, o autor faz um extensivo levantamento da forma como ela tem transcendido a simples condição de produtora para se tornar uma empresa determinante das principais tendências tecnológicas na área de informática.

Kahney (2004) faz um profundo estudo da comunidade de usuários de computadores Macintosh, demonstrando que ela, até certo ponto, parece algo muito similar a uma comunidade religiosa. Essa realidade cria um processo de fidelização à marca, que parece extrapolar a relação que nossa cultura esperaria que se desse entre um consumidor de computadores e quem os fabrica; Kahney documenta a existência de indivíduos que realizam obras artísticas como vídeos, quadros ou músicas para homenagear seus computadores, até mesmo os que criam cortes de cabelo e tatuam o corpo com o logotipia da empresa ou ícones de seus programas (KAHNEY, 2004, p. 31-45).

O estudo de Kahney (2004) vai mais além, demonstrando que essa seita de fanáticos pela Apple se caracteriza por sacralizar os fundadores da empresa — em especial Steve Jobs e Steve Wozniak — como símbolos da cultura yuppie e da contracultura. Apesar da contradição dessas duas filiações — a primeira sendo o símbolo do capitalismo bem sucedido e a segunda uma representação da luta contra o *status quo* —, cria-se uma nova representação ideológica: comprar produtos Apple implica manter-se no mundo da informática, mas permitindo-se quebrar os padrões corporativamente estabelecidos. É a criação do que Steve

Wozniak (apud KAHNEY, 2004, p. contracapa) define como: “The macintosh is more than 80 a computer. It’s a way of life (...) it is to be a Macintosh person (...)”. A utilização dessas máquinas ganha cores políticas, uma vez que as usar parece simbolizar não estar diretamente subordinado aos monopólios capitalistas representados por outras empresas, como é o caso da Microsoft. Nas palavras de Kahney: “(...) *Apple represents creativity and individuality; Microsoft represents business and conformity* (...)” (KAHNEY, 2004, p. 247).

1.2.2.2 Design

Design é um *stratum* (ou nível) que se encontra no meio do caminho entre conteúdo e expressão, podendo ser caracterizado como o lado conceitual da expressão bem como o lado expressivo da conceitualização (KRESS; VAN LEEUWEN, 2001, p. 05), definido como usos de recursos semióticos; o design é um meio para realização de um discurso em uma determinada situação comunicacional. O design seria o espaço no qual os interactantes seguiriam padrões relacionados ao hábito, tradição, prescrição ou mesmo inovação e quebra, levando a produção de criações que expressam o senso comum, passando pela inovação ou mesmo chegando a subversivo. Todavia, é importante observar que o design não é a instanciación efetiva do produto semiótico, ele é um *stratum*, ou nível, ainda abstrato, capaz de ser realizado em diferentes materialidades (KRESS; VAN LEEUWEN, 2001, p. 06). Ele estaria no espaço do planejamento e concepção da forma como aquele discurso seria realizado e estaria diretamente conectado, por conseguinte, com as práticas sociais estabelecidas.

Em seu trabalho, Kress e van Leeuwen (2001, p. 45-49) fazem uma longa discussão de como mudanças na conceitualização de design em novas mídias influenciam os indivíduos em seus contextos comunicativos. De forma geral, novas mídias criam espaços movediços nos quais as fronteiras entre as diferentes práticas sociais não são tão claras. Tomemos como exemplo a área de ensino-aprendizagem de línguas estrangeiras. Se em um mundo monomodal, o professor de idiomas e o produtor, ou mesmo editor de manuais didáticos, tinham papéis bem estabelecidos; isso hoje já não é bem verdade, uma vez que o acesso a

diferentes mídias e a possibilidade de interação entre elas faz com que virtualmente qualquer indivíduo possa se servir desses meios de produção: o professor pode escrever, imprimir e publicar seu material. Com os avanços em EAD e, dependendo do tipo de mídia que se queria usar, até mesmo a impressão é dispensável: o material pode estar disponível em diversos servidores. No mundo da arte isso também tem sido verdade e o cinema seria um exemplo: com a variedade de equipamentos disponíveis e a digitalização das ilhas de edição, virtualmente qualquer pessoa pode produzir um filme com baixo orçamento, o que era outrora impossível. Isso fica claro se voltarmos nossos olhos para as comunidades online de distribuição de conteúdo, como YouTube, Vimeo, entre outras; nelas indivíduos podem postar suas produções originais, sem restrições em relação ao conteúdo ou formato.

Os autores encaram o design como uma forma de estabelecer uma sequência/estrutura, "*a blueprint for production*" (KRESS; VAN LEEUWEN, 2001, p. 50- 54), sofrendo variações de acordo com o contexto em que é instanciado. Poder-se-ia afirmar que o design seria a expressão de um dado discurso, o que inclui valores e preferências, além de oferecer um modelo no qual aquilo que está prestes a ser produzido deve seguir. Então, se estamos falando de regularidades e comunidades, seria o design um espaço análogo ao gênero?

A resposta para essa pergunta é complexa e precisa de uma discussão específica. Se nos debruçarmos nas abordagens sobre gênero dominantes na LA brasileira nos dias hoje, perceberemos que duas parecem ser as principais. A primeira seria derivada dos estudos do filósofo russo Mikhail Bakhtin (1953; 1929/1992) e está direcionada para um estudo sociológico das condições de produção e do impacto sociodiscursivo dos gêneros do discurso. Bakhtin pressupõe que a produção dos diversos enunciados deve refletir as condições de sua produção, criando enunciados que carregam uma carga significativa atrelada ao seu objetivo e ao contexto social mais amplo que estão relacionadas. Esses enunciados são relativamente estáveis, proporcionando ao falante utilizar produções padronizadas para cada esfera (ou domínio) discursivo. Naturalmente, por serem diversas as esferas sociais, o são também os gêneros do discurso utilizados em cada uma delas.

Dentro dessa abordagem, Bakhtin estabelece uma diferenciação importante entre os gêneros primários e secundários. Os primários seriam aqueles produzidos de maneira espontânea, fruto da interação social cotidiana; ao passo que os secundários seriam típicos de ações comunicacionais mais complexas; formados a partir de combinações de gêneros primários. Cada falante teria o conhecimento dos gêneros existentes nas esferas (ou domínios) em que atua, sendo essencial.

A segunda abordagem estaria relacionada aos estudos desenvolvidos pela chamada escola de Sidney. Esses trabalhos têm como origem os estudos de GSF desenvolvidos por Halliday e pela noção de gênero estabelecida por Hasan (HALLIDAY; HASAN, 1989) e estão presentes nos trabalhos de Martin (1992; 1997; 2000), Eggins (1994; 2006) e Eggins e Martin (1997). Segundo Herrington e Moran (2005), essa abordagem desenvolve-se com o objetivo de usar o gênero como um princípio de inclusão social para nativos australianos, que não possuam inglês como sua L1, e para imigrantes com dificuldade de inserção social. Como resultado, teve uma grande influência em contextos de EAP e ESP.

O trabalho dessa escola se fia, principalmente, na busca por regularidades lexicogramaticais e estruturais que estruturam o gênero como uma produção textual estável (VIAN Jr.; LIMA-LOPES, 2005)²⁹, que tende a ser constituído por um número determinado de estágios, cada um com um objetivo específico dentro do processo de instanciação textual. Esses estágios possuem uma ordem relativamente estável, podendo ser classificados de acordo com sua obrigatoriedade e essencialidade (LIMA-LOPES, 2005).

Ao se comparar as diversas peças publicitárias em vídeo que nos deparamos em nosso cotidiano, perceber-se-á que há algumas dificuldades de aplicação de ambos os modelos brevemente discutidos acima. Não existe uma regularidade estrutural, e às vezes nem mesmo discursiva, entre os diferentes vídeos publicitários exibidos na televisão, muitos nem mesmo mostram o produto ou serviço nela anunciado: a variedade de recursos e de estruturas pos-

²⁹O artigo de Vian Jr. e Lima-Lopes (2006) foi publicado em uma coletânea organizada pelo GT de Gêneros Textuais da AMPOLL, cujo objetivo foi discutir um estado da arte da teoria de gêneros até então. Cf.: MEURER, J. L.; BONINI, A., et al (Ed.) **Gêneros: teorias, métodos e debates**. Campinas: Parábola, 2005. p.19-44.

síveis é grande. Isso parece ser resultado de uma característica importante desses textos: diferentemente da linguagem jornalística, da empresarial, da acadêmica e de esferas de atendimento do setor público, domínios muito comumente estudados pela LA, a publicidade não procura organizar seu discurso por meio de regularidades. De fato, o que chama a atenção em uma peça publicitária, seja ela em vídeo, impressa ou televisiva, é exatamente até que ponto ela se destaca em relação às demais. Isso leva a uma busca frenética pelos produtores por formatos que não se repitam. Naturalmente, algumas peças mais simples, como um jornalzinho de ofertas de supermercado ou um *flyer* de restaurante podem ser similares uns aos outros em diversos aspectos. Todavia, quando nos debruçamos sobre produtos comunicacionais mais complexos, e de produção mais cara como é o caso de vídeos, peças impressas de revistas, PDVs e uma série de outras mídias, esse preceito da regularidade parece cair por terra.

Apesar de reconhecermos os vídeos que compõem esta tese como peças publicitárias, eles não são necessariamente similares a outras da mesma natureza. Um exemplo muito interessante seria uma rápida comparação entre alguns anúncios que vendem a mesma categoria de produtos anunciados por eles. As empresas Dell e Lenovo, que vendem notebooks e computadores no Brasil e são concorrentes diretas da Apple, produziram peças em vídeo completamente distintas para sua última campanha. A Dell centra seus vídeos em ofertas de preço, muito similares ao realizado por supermercados, ao passo que a Lenovo, em uma campanha que anunciava um notebook com extensão de tela, mostrava como Geeks conseguiam seduzir mulheres graças às características de seu computador; ou ainda, uma simpática senhora que fazia suas tarefas de cozinha (assar, picar, cortar legumes) sobre um *Think Pad* (nome dos computadores da marca), de forma a demonstrar sua resistência.

Não obstante o modelo problema vs. solução ser um conceito muito comum em vídeos publicitários — nesse modelo o produto anunciado, representando a solução, é contraposto ao da concorrência, representando o problema —, a forma como esse processo se dá também é bastante peculiar. Por exemplo, os vídeos de "Get a Mac" mostram de forma clara quais

são os problemas técnicos que a plataforma se propõe a resolver, mostrando inclusive qual plataforma é a causa da infelicidade do consumidor. Fato que nem sempre se confirma em todas as campanhas que utilizam esse modelo.

Van Leeuwen (2005) discute essa questão, se inspira na pragmática de Austin (1962) e nos trabalhos de Eggins (1994) e Martin (1992) para caracterizar textos multimodais. Para o autor, é possível pensar a multimodalidade como um ato de fala, ações com a linguagem não se resumiriam apenas à linguagem verbal: ao observarmos uma imagem, por exemplo, também seríamos afetados por ela, o que faz dela um ato ilocucional (sua caracterização como ação), perlocucional (efeito no indivíduo) e locucional (apresentação da imagem). Ao mesmo tempo, esses atos devem se organizar de forma a serem reconhecidos pelos indivíduos que fazem parte de uma comunidade. No caso da linguagem multimodal, em especial em textos visuais, o conceito de gênero se relaciona à identidade visual de uma determinada mensagem (VAN LEEUWEN, 2005, p. 136), responsável por fazer com que ela seja reconhecida. Como já coloquei em Lima-Lopes (mímeo), os templates seriam um exemplo de como poderíamos discutir essa questão: eles oferecem algum tipo de organização para sítios de internet, por exemplo, mas mesmo assim não os deixam totalmente similares. Ali também discuto como dois jornais e duas revistas, veículos que são aparentemente muito similares, possuem inúmeras particularidades visuais que os fazem únicos. De certa forma, a aparente similaridade faz com que os reconheçamos, sem que sejamos incapazes de observar suas diferenças. O mesmo aconteceria com peças publicitárias em vídeo: é uma regularidade de propósito que me parece ser capaz de identificar esse produto audiovisual: apesar de não haver dois vídeos idênticos — afinal seu enredo, a construção de suas personagens, o tipo de linguagem usada, entre outros elementos constituidores tendem a ser completamente diferentes —, todos procuram vender uma ideia de que necessitamos de um determinado produto ou serviço.

O que parece possível no contexto destes vídeos, tanto os que pertencem à campanha que estudo nesta tese como a outras, é observar que eles possuem uma identidade interna. Todos os vídeos aqui pesquisados seguem um padrão temático e de estrutura, que não se

repete em nenhuma outra campanha, seja da Apple, seja de outro fabricante; talvez isso esteja relacionado, como já dito, à pressão constante por inovação sobre os profissionais responsáveis pela produção desses textos. Deve-se adicionar a isto, também, questões relativas ao direito autoral de criação e produção dessas peças; similaridades estruturais ou temáticas podem levar a processos judiciais, ao pagamento de indenizações. Anúncios publicitários não podem ser similares aos já existentes, parte do processo de sedução caminha exatamente pela inovação que eles devem trazer para o público. Dessa forma, a campanha “Get a Mac” é identificável aos seu público graças a suas características, sem que, contudo, seja similar às demais que vendem o mesmo tipo de produto. Essas regularidades poderiam formar uma espécie de EGP (Halliday; Hasan, 1989), mas ela não seria necessariamente aplicável a outras campanhas, seja da própria Apple seja de seus concorrentes. Por conta disso, acredito que a própria estrutura de problema vs. solução citado anteriormente, pelo menos no caso dos textos estudados nesta pesquisa, esteja relacionado à forma de organização do discurso.

Dado o grande número de possibilidades composicionais que esses textos possuem, não me sinto à vontade para discuti-los em termos de “gêneros”. Isso porque a regularidade dos enunciados parece ser estabelecida apenas localmente, e não no nível de toda a sua comunidade (em termos da escola de Sidney) ou de toda a sua esfera (em termos bakhtinianos). Por conta disso, prefiro assumir uma abordagem que procure definir suas regularidades composicionais — o que eles tem em comum no que tange à organização das imagens presentes nestes comerciais — e discursivas — que remeteriam aos conhecimentos e fatores ideológicos neles representados. Elas existem mas não são da mesma natureza que textos verbais: pensar na análise multimodal de um sítio de internet, tal qual o que distribuiu os vídeos desta tese, é muito diferente de analisar um artigo de opinião. Enquanto o primeiro segue uma lógica interna própria, que não se repete, uma vez que o layout é registrado em nome da Apple e certamente não pode ser usado por outra empresa, artigos de opinião tem uma estrutura semântica rígida, que se repete de forma quase invariável em todos os jornais que lemos.

Por conta disso, na análise a ser realizada neste trabalho, estarei me referindo constantemente a regularidades composicionais ou discursivas. Isso porque, como essas regularidades parecem ocorrer apenas em um nível local, ou seja, dentro da própria campanha "Get a Mac", eles não serão características definidoras ou caracterizadoras de todos os comerciais em vídeo. Em outras palavras, não busco analisar esses vídeos de forma a desenvolver um potencial de generalização para caracterizá-las como um gênero do discurso.

Sdjic (2008) discute o design a partir de sua conceitualização como linguagem e como arquétipo de prática artística e social. Para o autor, que tem seu foco no design de bens de consumo, ele evolui da simples criação de objetos funcionais, passando a agregar valor ao conceito que subjaz ao objeto. Sdjic (2008, p. 18) parte da já conhecida ideia de que o tempo transforma o design, sendo ele um indicativo da passagem do tempo, o que pode, em alguns contextos, fazer com que um objeto tenha seu conceito descartado ou valorizado. Essa relação cria uma linguagem específica de cada tipo de objeto, que nos ajuda a reconhecê-lo e a localizar seu lugar dentro da cada vez mais rica variedade de conceitos e projetos de design de hoje em dia. Essa forte ligação entre o design de uma peça e seu propósito social faz com que nasçam verdadeiros arquétipos de design, esses arquétipos existem graças a necessidade: objetos não existem no vácuo, eles são parte de uma complexa coreografia de interações (SDJIC, 2008, p. 54).

Dessa forma, um design — tratado não apenas como realização industrial de um objeto, mas como sua concepção artístico-funcional — segue uma espécie de ciclo, que deve se encerrar quando o arquétipo conceitual representado por aquele objeto não for mais socialmente necessário. O autor usa como exemplo o Mini Cooper, carro inicialmente produzido no Reino Unido a partir dos anos 50. Ele não só continua em produção, mas ainda possui diversos outros fabricantes que o imitam como fonte de inspiração para uma série de outros modelos. De fato, a combinação entre facilidade, desempenho e espaço fez com que o Mini superasse seu status de veículo e ganhasse o de arquétipo, o tema básico que diversos outros carros seguiriam (SDJIC, 2008, p. 56). O Fusca (*Beetle* em inglês), já descontinuado depois de um

ciclo de produção recorde (64 anos), talvez seja um outro exemplo de arquétipo, mas que corre em outra direção: apesar de o carro já não mais existir, ele é uma clara inspiração para outros modelos, como o Ford Ka, o Citroën C3 e o VW New Beetle. No caso específico da Apple, seus produtos e campanhas publicitárias, a partir da primeira metade dos anos 90, passaram a servir como referência da quebra desses arquétipos. Até 1993, ano de lançamento da primeira versão do iMac, o arquétipo de produção de computadores servia a um design que privilegiava o bege como cor padrão em todos os seus módulos (monitor, impressoras, CPU e periféricos). A revolução que os iMac propiciaram permitiram àquela empresa inovar no design de outros produtos, como foi o caso do lançamento do iBook e dos periféricos fabricados por ela. A partir dessa época, a Apple passa a ser uma marca de computadores desejada não apenas pelo desempenho de suas máquinas, mas também pelo design de seus produtos (LINZAMAYER, 2004), se transformam em símbolo da quebra desses arquétipos, pela beleza que representam. Ao mesmo tempo, seus produtos deixam de ser simplesmente computadores e passam a ser admirados como peças de arte, até mesmo pelos que não são seus consumidores. É natural que essa inovação em design fosse fruto de um discurso mais amplo que marcasse a diferença entre a empresa e seus concorrentes.

1.2.2.3 Produção

Para Kress e van Leeuwen (2001, p. 66), a produção pode ser o uso comunicativo dos diferentes recursos materiais. A produção pode ser definida, inicialmente, como um trabalho pertencente ao mundo físico e realizável por humanos ou por máquinas que são, assim, responsáveis pela materialização de um texto semiótico, sendo que a interpretação de seu significado sempre esteve relacionada à utilização dos sentidos. Por essa razão a produção sempre esteve ligada a diferentes mídias — ou meios, aqui usados como sinônimos — e, por conseguinte, relacionada a diferentes canais sensoriais. Esses canais sensoriais formam uma determinada configuração (KRESS; VAN LEEUWEN, 2001, p. 66), necessária para a interpretação de um texto; por isso não se pode pensar na interpretação da produção como

uma mera recepção passiva de conteúdo. Cada vez que usamos um de nossos sentidos, esse sentido pode ser considerado uma atividade de percepção de alguma produção multimodal e multissensorial: não incomumente varremos nosso ambiente com nossos órgãos sensoriais e gradativamente tomamos conhecimento de significados expressos em diferentes modos, como a variação de luz, a existência de sons e cheiros, a percepção de texturas etc. (KRESS; VAN LEEUWEN, 2001, p. 67). Naturalmente, nossa cultura privilegia a utilização de alguns canais sensoriais em detrimento de outros na compreensão e recepção das diferentes produções, o que nos leva a valorizar formas de produção que lidem com esses canais sensoriais.

Para os autores, o conceito de meio (ou mídia) inclui não apenas o corpo e a voz, mas todas as ferramentas que podem estar além da comunicação puramente corporal, incluindo ferramentas e materiais utilizados para a produção de artefatos diversos. Essa forma de encarar a produção, a coloca em um ponto de destaque, dando-lhe, por conseguinte, um papel semiótico específico na realização de significado de uma dada mensagem e não a categorizando apenas como um simples meio de instanciação do design.

Textos videográficos possuem um processo de produção bastante complexo, incluindo diversos profissionais, cada um com uma função bastante específica. Essa complexidade acaba por dividir o processo em três fases:

Pré-produção: normalmente relacionada à escritura do roteiro, seleção dos atores, produção dos cenários e locações, contratação de profissionais como fotógrafo, maquiador, continuísta, cinegrafista, entre outros, além do aluguel ou compra de todos os equipamentos a serem utilizados na gravação.

Produção: é a gravação propriamente dita; vale lembrar que qualquer produto audiovisual não é gravado na ordem sequencial da narrativa: costuma-se a gravar todas as cenas que ocorrem em um determinado cenário ou locação, evitando o deslocamento constante de equipamentos barateando o processo.

Pós-produção: nela ocorrerá montagem do filme — a grossíssimo modo a montagem é o processo de colocar as cenas na ordem —, aplicação de efeitos especiais, dublagem de cenas mais complexas, mixagem do áudio, correções de cor, colocação de efeitos especiais, entre outros processos.

Essa forma de entender a produção faz com que ela seja tão importante como o discurso ou o design. Isso se dá pelo fato de que os recursos que utilizamos para compreender, articular e interpretar o mundo residem tanto nos modos semióticos como em suas mídias (KRESS; VAN LEEUWEN, 2001, p. 68). Por essa razão é importante diferenciar modos de linguagem de mídia, uma vez que esses são conceitos chave para a interpretação das escolhas e dos significados em diferentes níveis. Kress e van Leeuwen (2001, p. 67) exemplificam essa diferença analisando o Rádio — classificado pelos autores como um meio de distribuição e não de produção —, que é multimodal por definição — uma vez que envolve a fala, música paisagens sonoras e uma série de outros tipos de sons — mas monomidiático, uma vez que esses significados são apenas ouvidos, não envolvendo outros sentidos. O simples fato de uma informação ser distribuída por um sítio de internet não a faz multimidiática, se apenas a leitura estiver envolvida ela será monomidiática; mas se torna multimodal à medida que temos imagens e tipografia juntos para a construção desse significado. Isso faz com que a definição de produção carregue consigo um potencial experiencial de significados, que seria definido como a expressão de sistemas de valores que são expressos por meio de qualidades materiais dos artefatos produzidos em uma cultura (KRESS; VAN LEEUWEN, 2001, p. 77). Isso incluiria todas as qualidades definidoras dos elementos que constroem uma cultura, entre elas a cozinha, a forma de se vestir, a arquitetura, a forma como indivíduos se comportam fisicamente etc. Como resultado, os participantes de uma determinada cultura já teriam como pressuposto uma série de características que deveriam ser típicas da produção de certos significados e/ou sistemas semióticos. Quando pensamos na imagem em movimento, essa questão pode ser um pouco mais complexa. Isso porque além da composição visual estática,

como na fotografia, temos ainda a transição entre os diferentes frames, além da interação entre o vídeo e o áudio, que é um elemento construtor de significados.

Hashimoto (2010) discute a importância que o avanço tecnológico tem para a concepção de novas formas de pensar a publicidade. Em sua pesquisa, ele argumenta a favor da importância de coletivos e formatos alternativos de produção e distribuição de filmes que, tradicionalmente, estão inseridos apenas em modelos de distribuição tradicionais, como a televisão. O interesse de Hashimoto (2010) é desenvolver uma nova plataforma, cuja significação é uma nova forma de pensar a produção desses vídeos. Entretanto, ao percorrer esse caminho, o autor faz uma análise importante da trajetória das tecnologias de produção e distribuição de audiovisual que influenciam na construção dos significados. Efetivamente, pode-se afirmar que a tecnologia é um elemento constituinte importante para a produção de qualquer tipo de criação audiovisual, uma vez que ela é influenciadora da linguagem utilizada — formatando-a e limitando-a de acordo com suas possibilidades —, pois a técnica produz uma diferença na expressão (MACHADO, 2007b).

1.2.2.4 Distribuição

Na maioria das vezes, tecnologias de distribuição são criadas para propósitos de reprodução ou de transmissão de informações específicas, um exemplo desse tipo de realidade é a internet. Inicialmente inventada para a comunicação militar, a Arpanet (versão militar da Internet) tinha por objetivo garantir a comunicação entre as bases militares estadunidenses em situações críticas, garantindo a reprodução e o recebimento de mensagens sobre ataques e lançamento de mísseis. Posteriormente, a rede foi ampliada para as universidades e seu potencial semiótico permitiu que ela se tornasse o que é hoje. Para KRESS e VAN LEEUWEN (2001) o potencial semiótico é um elemento importante para a compreensão das funções que os diferentes meios possuem, esse potencial não está encerrado nas possibilidades que o design de um sistema de distribuição inicialmente previa; ele é ampliado à medida que esse meio se

populariza e é tecnologicamente modificado, criando possibilidades infinitas de significação e estruturação dos textos neles veiculados.

Ao se realizar qualquer processo de distribuição de um texto (seja ele escrito, verbal ou multimidiático) temos duas operações ocorrendo a (re)codificação e transmissão (KRESS; VAN LEEUWEN, 2001, p. 87). Conversas telefônicas são apenas transmitidas — gravações são esporádicas e, normalmente, fogem do propósito inicial do meio —, CDs são apenas gravados, ao passo que mensagens de e-mail são gravadas, enviadas e apagadas de servidores — no caso da tecnologia POP3 —, ou mesmo simultaneamente sincronizadas — no caso da tecnologia IMAP. Ao se decidir redistribuir um CD, transmitindo-o em uma rádio ou disponibilizando-o em uma rede P2P para compartilhamento de arquivos, estamos realizando um novo processo de distribuição, e, portanto, recodificando a mensagem; o mesmo seria válido para mensagens de e-mail que sejam impressas para uma reunião, ou que tenham seu conteúdo anexado a agendas de compromissos.

Seria ingênuo pensar que o processo de distribuição não influencia a compreensão, uma vez que a forma de acesso e o potencial semiótico de cada meio de distribuição trazem perdas e ganhos para a mensagem contida em um dado texto (KRESS; VAN LEEUWEN, 2001, p. 88-89). Um dos primeiros ganhos trazidos pelas transformações nos meios de distribuição e, portanto, no potencial semiótico que em cada um deles se encerra, é a acessibilidade. Um avanço considerável foi trazido pela imprensa que popularizou em grande escala a linguagem escrita e, posteriormente, o acesso às artes plásticas, como a pintura e a fotografia (KRESS; VAN LEEUWEN, 2001, p. 89); outro seriam as transmissões em nível radiofônico e televisivo que nos permitem, por exemplo, assistir a concertos sem necessariamente irmos à ópera. Nesse sentido, apesar de não apontado pelos autores, a internet novamente tem um papel deveras importante: as possibilidades de retransmissão e recodificação oferecidas por ela possibilitaram o acesso a informações que antes, muito provavelmente, uma grande massa de internautas não teria. Em minha opinião, a distribuição via rede internacional de

computadores está relacionada não apenas à acessibilidade, mas também à globalização e propagação de discursos e designs outrora locais.

Já a questão da perda é tratada por Kress e van Leeuwen (2001, p. 89) em dois níveis: aura e modo. A primeira estaria relacionada à noção de vulgarização e massificação que a distribuição em grandes quantidades gera, após a prensa de Gutenberg, por exemplo, uma série de textos outrora restritos passam a ser consumidos por um grupo muito maior de indivíduos. Nos dias de hoje, o mesmo acontece no universo da música alternativa: comunidades fechadas se orgulham de conhecer uma determinada banda de rock antes que ela faça sucesso e chegue até as massas por meio de uma gravadora *mainstream*. Já a segunda, estaria ligada a transformações que uma mensagem sofreria para que sua recodificação ocorra; presenciar um concerto, por exemplo, é muito diferente de ouvi-lo em um aparelho de som, ao passo que baixar um CD via redes P2P torna impossível o acesso a certas facetas artísticas como boxes especiais para distribuição, desenho de capas, entre outros.

Será que essa diferença de experiências deve ser considerada uma perda? Ou deveríamos pensá-la como uma diferente forma de expressão de conteúdos, agora com características novas? Em minha opinião, pensar em diferenças na constituição e não em perda de conteúdos seria algo teoricamente mais adequado. A visão de Kress e van Leeuwen (2001) pressupõe uma espécie de hierarquia entre os diferentes modos, sendo pressuposto pelos autores que o modo de distribuição temporalmente mais antigo é o que traz a mensagem dita "original". Isso implica que seria apenas nela que residiria sua completude e plenitude. No mundo multimidiático em que vivemos, nos deparamos com diferentes canais de distribuição e nosso acesso à informação depende de uma série de fatores, entre eles dos canais disponíveis no momento em que temos acesso a ela. Por conta disso é difícil afirmar que os indivíduos que têm acesso a, por exemplo, filmes de cinema distribuídos por redes de TV sob demanda como — Joost, Veo entre outras diversas — tenham uma experiência incompleta. Essa mesma questão pode ser abordada de um outro ângulo: alguns filmes, por exemplo, após sua estreia nos cinemas nunca mais são relançados em salas comerciais, sendo que o contato dessas

obras depende exclusivamente de sua distribuição em DVD ou Blu-Ray. Nesse ponto, se pressupormos o colocado por Kress e van Leeuwen (2001), todos os que não prestigiassem a exibição nas salas de cinema teriam recebido a mensagem com algum grau de perda. Kress e van Leeuwen (2001, p. 90) afirmam que o fascínio pelas novas mídias e formas de distribuição pode gerar uma sensação de ganho, que está ligada às novas necessidades criadas por ela. O que parece haver, efetivamente, é a criação de uma nova forma de relacionamento com o texto que é distribuído em uma nova mídia: há uma forma de interação que difere daquela originalmente pressuposta pelo produtor, sendo que essa forma não é inferior tampouco superior a nenhuma outra, ela gera significados próprios e que contribuem de forma única para sua interpretação.

Um exemplo disso seria a forma como os vídeos estudados neste trabalho são disponibilizados: por meio do sítio da empresa responsável por estes anúncios. No nível puramente técnico, essa forma de distribuição faz com que os vídeos sofram um processo de compressão para diminuição de seu tamanho e possibilidade de serem assistidos online. Do ponto de vista ideológico, a Apple, uma empresa estadunidense, disponibiliza seus valores ideológicos a um número muito maior de indivíduos, possibilitando uma hegemonia que vai muito além das fronteiras de Cupertino, cidade em que a empresa se situa, ou mesmo dos EUA.

1.2.3 Analisando diversos modos simultaneamente

A análise conjunta de diversos modos de linguagem vem demonstrando ser uma das grandes contribuições da teoria multimodal para a análise dos discurso. Pela primeira vez, foi permitindo aos linguistas ir além e analisar elementos significativos que estão relacionados a um mundo no qual a língua é apenas mais um modo de linguagem (NORRIS, 2002; 2006). Norris (2002) estuda a interação entre 5 crianças que brincam juntas, assistindo à televisão e jogando jogos de computador, com o objetivo de mostrar que a incorporação de diversas modalidades na transcrição para análise pode mudar o foco do estudo da linguagem verbal para a ação humana como um todo. A autora consegue demonstrar que, da mesma forma

que as palavras têm um importante papel na comunicação dessas crianças, as imagens as quais elas se referem durante sua interação são igualmente importantes para a compreensão de sua interação e dos significados por elas articulados. Já em outro trabalho (NORRIS, 2006), a autora estuda a interação em um escritório de contabilidade, demonstrando como a análise de uma ação complexa pode extrair significados que revelam informações sobre a interação como um todo: a quem a real atenção é dada, qual o real foco da interação e que tipo de relação os indivíduos possuem. Norris identifica algumas das principais limitações da análise linguística tradicional: o foco único e exclusivo de transcrições baseadas na linguagem verbal parecem fragmentadas, não dando conta da totalidade de significados que as crianças parecem exprimir. A mudança para uma abordagem que leve em consideração as imagens vistas pelas crianças, os filmes assistidos, seus gestos, olhares ajudam a criar uma reflexão mais profunda sobre a interação, compreendendo melhor os significados ali criados. O estudo de Norris explora essa totalidade, demonstrando que a interação entre essas crianças só pode ser efetivamente compreendida se todos os modos de linguagem forem levados em conta.

Outros trabalhos relevantes seriam o de Jones (2005), que estuda a multimodalidade na interação entre homossexuais masculinos em videochats, Baldry e Thibault (2006), que oferecem um importante guia de análise multimodal para meios impressos e digitais, Martinec e Salway (2005) que realizam um estudo multimodal em anúncios de mídia impressa, Lemke (1998; 2002), que estuda a relação entre imagens e linguagem escrita no discurso científico e em hipertextos, respectivamente, além de Ferreira (2011), que faz um estudo comparativo entre a relação texto vs. imagem em anúncios da revista *Times*. Em seu estudo, Jones (2005), mostra que o ato de utilizar imagens para a comunicação e negociação é um elemento que serve para a criação de identidade e associação de valores. No caso da interação multimodal estudada pelo autor, uma série de riscos é assumida; por conta disso, há um controle da natureza da informação disponibilizada na interação, o que certamente afeta seu fluxo. Para Jones, o ato de liberar a câmera de vídeo do computador durante a negociação do encontro garante a continuidade da conversa, mas também agrega valores que dependem de alguns

fatores, entre eles está a parte do corpo que se mostra. Por seu turno, Martinec e Salway (2005) observam a interação entre linguagem verbal e visual em anúncios publicitários. Os autores observam, entre seus resultados, que textos e imagens podem trazer significados independentes e complementares, contribuindo assim de forma diferente para a construção dos significados. Eles também observam que textos e imagens podem possuir uma relação de interdependência, sendo que o texto pode trazer informações que complementam os significados imagéticos ou vice-versa. No caso do primeiro, texto e imagem expressam diferentes significados que, ao serem justapostos, trabalham de forma complementar para a mensagem. No segundo, é demonstrado que as teorias sobre coesão textual, desenvolvidas por Halliday (HALLIDAY; HASAN, 1976), são elementos essenciais nessa interação entre sistemas semióticos. Dessa forma, tanto o código imagético como o verbal se autorreferenciarão em um processo de interação e construção conjunta dos significados. Essa relação de referência mútua mostra-se algo sistemático e observável como discute Kong (2006), cujo trabalho estabelece uma taxonomia da relação entre ilustrações e linguagem verbal em jornais de Cingapura. Kong consegue demonstrar que grande parte das relações estabelecidas por Halliday em termos de hipotaxe e parataxe também são aplicáveis ao seu contexto de pesquisa, sendo necessárias apenas algumas poucas adaptações devido às especificidades dessa relação.

Lemke (1998) analisa a interação entre elementos gráficos (tabelas, figuras e quadros) e a linguagem escrita em textos acadêmicos. Entre seus resultados, o autor observa que a construção dos significados nesses textos está efetivamente calcado em ambos os códigos (visual e escrito). Na verdade, para o autor, a base da possibilidade de interpretação destes textos está em uma formação temática dependente de uma contextualização intertextual e interpictorial. Já em outro artigo, o mesmo autor (LEMKE, 2002) discute a questão da multimodalidade no hipertexto. Sua pesquisa traz importantes resultados no que tange à interação entre linguagem verbal e visual, mostrando que a relação entre esses dois meios também é um elemento determinante da produção de significados. Os resultados de Lemke são bastante similares aos de Macken-Horarik (2004) que estuda essa mesma interação em

obras de arte expostas em uma exposição temática na Austrália. Contudo, a autora chama a atenção para a necessidade de observarmos os pontos em que esses dois meios não necessariamente coincidam. Ela ainda coloca que, uma vez que vivemos em um mundo multimodal, há necessidade de desenvolvimento e um aparato metodológico que permita a análise dessas complexas interações.

De certa forma é isso que Baldry e Thibault (2006) tentam realizar. Os autores dão uma importante contribuição ao mostrarem a aplicabilidade dos critérios de análise verbal em textos multimodais. Os autores observam que é a interação entre gestos, imagens e linguagem verbal que proporciona a possibilidade de compreensão da mensagem, sendo que todos esses modos exigem o mesmo tipo de conhecimento dos contextos de situação e de cultura. Além disso, eles mostram que todos esses meios devem ser observados paralelamente para que haja uma completa apreensão dos significados e funções presentes nesses textos.

Com relação aos textos vídeo-fílmicos, Baldry e Thibault (2006) demonstram que os diferentes modos compõem sua construção de forma igualmente importante. Uma implicação disso é que, ao analisarmos uma peça audiovisual, não podemos prescindir de analisar todos os modos de linguagem, dado que eles conjuntamente formam um mesmo signo. Caminhando pela mesma seara, Eisenstein (1947/2002) reflete sobre a produção audiovisual de sua época, sendo que sua preocupação é entender como a conjunção de imagem e som cria uma obra de significado único. Dessa forma, parte-se da edição (ou montagem) como princípio básico que permite a criação desse ícone de caráter único e indivisível, cada unidade de uma obra audiovisual funciona em um movimento unificado visando não apenas à construção de uma linha que integre os diferentes quadros que se sucedem, mas também os diferentes modos que ali estão presentes (1947/2002, p. 55). Eisenstein, que tinha forte influência do estruturalismo, faz uma analogia entre as relações dos eixos sintagmático e paradigmático como definidos na linguagem verbal: em textos audiovisuais teríamos uma montagem vertical, responsável por encadear a trilha sonora e a imagem em seu caráter sequencial, que forma uma superestrutura e rege a estrutura horizontal de cada quadro ou *frame* (1947/2002, p. 57).

Naturalmente esse processo de conjugação ganha aspectos mais complexos, em especial se observamos que uma unidade sonora nem sempre corresponde a uma visual e vice-versa. Inicialmente, a união entre essas diferentes unidades aconteceria por questões rítmicas e de sincronização física entre as bandas visuais e sonoras. Todavia, segundo o autor, modos expressivos e significativos de linguagem podem nascer de construções nas quais essa sincronização não necessariamente esteja presente. Isso gera a criação de obras que se baseiam em contrapontos, baseado em uma coerência interna entre a imagem tangível e os sons percebidos (Eisenstein, 1947/2002, p. 57-61). Iedema (2001) também parte da edição como um pressuposto para a construção de significados dentro de um texto vídeo-fílmico e mostra como o processo de escolha de imagens, planos e trilha sonora pode tentar levar a uma interpretação daqueles que assistem a essa obra (ver discussão mais acima sobre o conteúdo do documentário analisado pelo autor). No caso da obra estudada pelo autor, um documentário que explora os problemas do sistema de saúde australiano, determinados personagens são nitidamente favorecidos pelo processo de montagem, mostrando que esse tipo de texto também pode ser influenciado por escolhas de caráter ideológico.

Apesar de todos esses esforços em analisar diferentes modos de linguagem, é importante discutir questões relativas às contribuições que uma análise dessa natureza pode trazer para este trabalho.

1.3 Resultados parciais

Um elemento importante para essa tese foi pensar se seria possível utilizar as categorias já propostas e explicar a realidade de linguagens multimodais, de forma a entender qual o potencial e o limite dessas, verificando o poder explicativo. Por conta disso, neste capítulo, busquei elaborar uma visão ampla e crítica da análise Multimodal, ou Sociossemiótica Multimodal. Meu ponto de partida foram os estudos de Halliday (HALLIDAY, 1973; 1978; 1985;

1994; HALLIDAY; MATTHIESSEN, 2004), uma vez que essa análise parece ter evoluído diretamente de seus trabalhos.

Halliday constrói uma teoria de caráter social, com características semântico-discursivas que desloca a análise para o eixo paradigmático. A linguagem é, então, encarada como um sistema de escolhas de signos motivados, graças ao potencial semiótico que cada escolha carrega consigo em uma cultura ou comunidade. Isso leva à observação de como cada cultura ou comunidade organizaria seu discurso por meio de ações de linguagem. Aos olhos da Gramática Sistemico-Funcional — que dá as bases teórico/epistemológicas para análise multimodal — a linguagem e o contexto estão em uma relação intrínseca. A ideia central dessa relação é que ambos estão ligados de forma tão profunda, que qualquer ato de linguagem refletiria o contexto social que foi realizado, da mesma forma que esse contexto reflete as ações de linguagem. Para Halliday e Matthiessen (1999, p. 1), a construção do significado se dá enquanto experiência, sendo esta vista não como um produto, mas como um recurso potencial para a compreensão, interpretação e ação sobre nossa vida cotidiana e construção da realidade. Por essa razão que a teoria encara as práticas sociais e linguageiras como indissociáveis: os contextos em que agimos pré-selecionam os significados de nossos atos comunicacionais em relação ao contexto. Isso se dá porque nem todo significado é igualmente comunicável nas diferentes situações, dado que as relações entre os falantes, a forma socialmente aceita de estruturação da mensagem, assim como as diferentes ações que realizamos para efetivação da comunicação serão diferentes nas distintas realidades.

Ou seja, para Halliday (HALLIDAY, 1985; 1994; HALLIDAY; MATTHIESSEN, 2004) as noções de significado e conhecimento, enquanto elemento constitutivo do discurso, são construtos equivalentes (HALLIDAY; MATTHIESSEN, 1999, p. 3). A definição da experiência é, logo, dada em termos de sua constituição por meio da linguagem, é ela que organiza e ordena os elementos do mundo em um sistema de escolhas, criando o que Halliday chama de *meaning base*. Essa base é definida pelo autor como uma gramática, na acepção descritiva e não normativa, semanticamente motivada e que se opõe a uma abordagem formalista. Em outras

palavras, o signo perde seu caráter de arbitrariedade, o que se dá graças à forte conexão que a linguagem tem com a ação social: as motivações sociais dadas ao uso de determinados recursos linguageiros fazem com que eles sejam reflexos motivados de nossa experiência e da cultura em que estamos (KRESS, 1993; STUBBS, 1996).

Poder-se-ia, assim, afirmar que a GSF tem por objetivo explicar como as diversas possibilidades de instanciação de linguagem se organizam em sistemas não fechados de escolha, os quais estão ligados de forma funcional ao seu significado socialmente motivado. Aqui não existem modelos de combinação preestabelecidos, a abordagem se preocupa em realizar um extensivo levantamento da instanciação dos diferentes significados em seus contextos específicos. Esses sistemas de escolha abertos se transformam dentro dessa dinâmica em pelo menos 3 momentos: um primeiro chamado filogênese, que segundo Halliday leva ao nascimento desse sistema, assim como sua maturação; o segundo, chamado ontogênese, relacionado a seu desenvolvimento pelo indivíduo; e a logogênese, que seria a instanciação efetiva do sistema nas diferentes manifestações textuais (HALLIDAY; MATTHIESSEN, 1999, p. 18). Nesse sentido, todo sistema de escolhas é um sistema de significados potenciais, os quais são instanciados nos diferentes textos que produzimos.

É importante retomar dois princípios-chave da GSF: o de realização e o de constituição. O primeiro poderia ser definido como uma espécie de relacionamento interníveis (ou interestráticos). Na mesma medida que o nível lexicogramático realiza o nível semântico, este é realizado pela lexicogramática; o mesmo pode ser dito em relação a todos os níveis, o fonético é, logo realizado pelo fonológico e todos, em suas interações instanciativas, realizam o contexto. O que se coloca é que a relação de um nível com os imediatamente menos abstratos é aquilo que lhe permite existir enquanto expressão de linguagem; cria-se uma relação de copresença e dependência entre níveis, cada um deles reflete e é refletido em suas escolhas (HALLIDAY; MATTHIESSEN, 1999, p. 33). Exemplos disso são a figura 1, na página 31, e a figura 2, na página 32, nas quais pode ver uma exemplificação de como esses níveis (ou *strata*) se relacionam tanto na linguagem verbal (que parte dos pressupostos hallidianos) como na linguagem

videográfica (uma tentativa minha de adaptação do modelo para a realidade desta tese). Isso no leva ao segundo princípio, a constituição. Como podemos inferir, cada um dos níveis (ou strata) é constituído pelo nível que está imediatamente abaixo dele. Tomando a linguagem verbal em sua modalidade oral, à guisa de exemplo, observaremos que o nível fonético existe dentro do fonológico; imediatamente acima, perceberemos que a lexicogramática, que representa o nível das escolhas de fraseado, abarca tanto o fonético como o fonológico, e assim sucessivamente até o contexto, que é, por conseguinte, o produto da interação entre todos os níveis de estratificação. Tal princípio operacionaliza, em termos funcionais, o da realização, ajudando a determinar quais são as possíveis unidades imediatamente constituintes em cada um dos níveis de contexto. Diferentemente de outras abordagens que primam apenas por ser ou botton-up ou top-down, a GSF parte de um sistema de análise no qual a observação de níveis mais abstratos ou mais concretos são complementares. Nesse sentido, tanto o processo como o produto são igualmente importantes: ambos mostram faces diferentes da mesma construção funcional de significados.

Apesar de existirem diversas funções de linguagem, algumas parecem se destacar por serem essenciais à constituição do enunciado. Assim, ao falarmos estamos agindo sobre o outro (Metafunção Interpessoal ou Orientacional), representando nossa experiência (Metafunção Ideacional ou Presentacional) e organizando a linguagem para o uso (Metafunção Textual ou Organizacional). Um resumo dos principais conceitos da GSF que irão ser continuamente retomados pelos teóricos de multimodalidade é:

Sistema de escolha: a linguagem é composta por um sistema de escolhas, as quais são realizadas em detrimento de outras; as escolhas interagem entre si, formando uma rede.

Níveis (*strata*): a linguagem é composta por níveis, que vão desde os mais abstratos até os mais concretos, qualquer alteração em qualquer um desses níveis levará a uma transformação em todo o sistema.

Instanciação (realização): é o ato de transformação de um conceito abstrato em linguagem.

Potencial semiótico: relacionado à capacidade de significar de um elemento linguístico ou modo, dada sua função social.

Contexto de cultura: é o mais abstrato, relacionado ao gênero e aos acordos culturais mais abstratos no que tangem a nossa comunicação.

Contexto de situação: é o registro, relacionado ao contexto imediato de produção

Metafunções: funções indispensáveis na realização de qualquer enunciado..

Função: em GSF é sinônimo de significado e de ação social.

Seguiu-se então uma discussão sobre algumas questões relacionadas à abordagem multimodal, que parece se dividir em dois núcleos centrais de argumentação. O primeiro, representado por autores como Kress e van Leeuwen (KRESS, 1993; KRESS; VAN LEEUWEN, 1996; 2001; 2002; KRESS, 2003; 2005; 2006; 2010) e Lemke (LEMKE, 1998; 2002; 2004; 2006a; b; c) atestam que o desenvolvimento tecnológico gerou modificações em nossa forma de pensar e usar a linguagem, tendo consequências óbvias para o papel da escrita. O segundo, que inclui autores como Norris (2002; 2004; 2006) e O'Toole (1994, 2004), que partem desses pressupostos para desenvolver uma teoria da multimodalidade que não dê conta apenas das linguagens tecnológicas, mas também das ações em nosso cotidiano.

Adotou-se como linha condutora os trabalhos da primeira corrente, por serem autores seminais que desenham as teorias desenvolvidas pela segunda. Nesse ponto, observa-se que o trabalho de Kress, que serviu como base para a argumentação parecia seguir por duas linhas. Uma, que está em seus primeiros trabalhos, argumenta a favor de um declínio da valorização da escrita, algo que seria fruto tanto dos avanços tecnológicos da microinformática e da consequente valorização do código visual. Nesse ponto, críticas foram tecidas a Kress, especialmente, em relação ao que pareceu ser uma espécie de imediatismo de sua teoria. Esses

dois fatos parecem ter sido fruto de uma evolução histórica, de um processo de construção de visualidade que traz em si um caráter ideológico e estrutural. Também discutiu-se a necessidade de observação de fatores sociais, políticos e econômicos como sendo muito relevantes para compreensão da maior valorização dada um determinado modo de linguagem.

Seguiu-se então a reflexão sobre um segundo momento da teoria, agora preocupada em discutir como os diferentes modos semióticos são utilizados para a construção do significado. Ali há uma preocupação perceptível em adaptar os elementos da teoria hallidiana de forma mais clara, ao mesmo tempo que a multimodalidade parece ganhar mais independência, tornando-se uma teoria geral da comunicação. Em termos gerais, a discussão colocada neste capítulo ajuda a observar alguns pressupostos importantes da análise multimodal, os quais serão importantes para a análise de dados realizada nos capítulos a seguir: 1) que cada um dos modos possui um sistema próprio; 2) o sistema de cada um dos modos de linguagem se relaciona de forma própria em cada um dos textos instanciados, eles diferem em termos das variáveis de registro e das metafunções em cada um dos vídeos analisados; 3) não existe modo principal de linguagem, todos são igualmente relevantes para produção do significado ali instanciado e 4) as escolhas em cada um dos modos podem influenciar os demais, numa complexa rede de escolhas e de criação de significado.

Se observarmos a figura 3, na página 37, notaremos que o conceito de *strata* usualmente desenvolvido dentro das análises funcionais de origem hallidiana, prima pela hierarquização da relação entre os níveis de contexto: o gênero, o registro e as metafunções representam uma escala de abstração, na qual valores relativos a nossos códigos sociais (gênero) se manifestam nas variáveis da situação em que interagimos (registro), que são finalmente instanciadas em forma de escolhas no sistema semiótico (metafunções). Não é à toa que esses níveis de contexto são representados de forma a pairar uns sobre os outros. Dentro dessa teoria (HALLIDAY, 1994; HALLIDAY; MATTHIESSEN, 1999; 2004), qualquer modificação em quaisquer dos níveis de contexto afeta as demais, ou seja, qualquer variação, no nível mais abstrato (contexto de cultura), ou no mais concreto (instanciação e realização linguística),

seria sentida e acarretaria em transformações nos demais. Isso significaria uma indissociabilidade entre o sistema linguístico e a sociedade que o cerca, levando o primeiro a ser, logo, motivado pelo segundo (KRESS, 1993). Essa teoria sugere que a nova realidade multimodal faz com que uma escolha agora afete não apenas o próprio sistema — representado dentro do modelo hierarquizante descrito acima — como também os sistemas que estão sendo paralelamente presentes naquele ato comunicacional. Em consequência disso, há a criação de uma rede Metassemiótica de escolhas, linguagens virtualmente distintas estão conectadas e interagem entre si graças à motivação criada pelo objetivo comunicacional do autor. Em outras palavras, ao escolhermos elementos específicos na instanciação de um significado, estamos selecionando não apenas os elementos que os acompanham dentro de um modo de linguagem, como também nos demais modos que constituem aquela proposição.

Logo, a rede de co-ocorrências e os mapeamentos de significados possíveis perpassam todos os modos de linguagem presentes em um determinado ato comunicacional. O que sugiro aqui é que essa relação pode se manifestar de igual forma entre os modos que compõem uma interação: apesar de eles estarem num mesmo nível hierárquico, todos refletiriam escolhas que juntas construiriam um mesmo significado, qualquer alteração em um desses modos de linguagem leva a uma reconstrução quase que total desse sistema. Em termos gerais, a relação probabilística da linguagem como sugerida por Halliday (1991; 1992a; b; 1993b; a; 2005/1979), se ampliaria agora para uma realidade hipermodal: cada nova escolha se manifesta em uma teia virtualmente infinita de significados e modos que juntos constroem uma determinada situação comunicativa.

Alguns modos de linguagem possuem uma formação significativa naturalmente híbrida, causando uma dualidade na forma como podem ser analisados, sendo o vídeo (BOLTER, 2006; HOBBS, 2006), ou mesmo a fotografia, casos típicos de tal hibridismo; ambos podem ser analisados ora como uma mídia tão somente ora como um modo de linguagem. No caso da primeira concepção, o vídeo e a fotografia podem ser encarados como forma de registro do cotidiano, que implicaria uma produção não necessariamente preocupada com questões

estéticas e de linguagem foto/videográfica. Nesse caso o vídeo e a fotografia não teriam muito mais influência na mensagem, além da natural tradução do mundo tridimensional em que vivemos para a bidimensionalidade da tela, da ampliação fotográfica e da composição do frame ou fotograma. Assim, essas duas mídias poderiam ser encaradas como registro de outros modos, formas de manter uma interação que antes poderia estar perdida devido sua ocorrência apenas imediata. Na segunda concepção, o vídeo e a fotografia seriam, cada um, uma linguagem e um sistema semiótico possuidor de uma estrutura própria. Isso levaria a padrões de escolha bastante complexos, que utilizariam uma série de subsistemas.

No contexto geral desta tese, essa discussão foi importante por oferecer uma melhor base de reflexão teórica para o desenvolvimento da alternativa metodológica de análise a que me proponho. De maneira geral, ela parece deixar claro que a análise multimodal não pode deixar de observar todos os modos de linguagem como um amálgama de significados único. Ao mesmo tempo, entender esse significado, assim como sua instancição, passa por um processo analítico que inevitavelmente desconstrói o texto, observando quais escolhas dentro de cada modo de linguagem instanciam um determinado significado. É a consciência desse processo que permite o tipo de abstração analítica desenvolvida nos capítulos 4, na página 127, 5, na página 163, e 6, na página 213. Além disso, pareceu-me também claro que seria impossível pensar em uma análise multimodal sem uma reflexão interdisciplinar, que traga elementos de outras áreas como comunicação, design e filosofia. Isso é importante por nos ajudar a compreender os elementos típicos de cada um dos modos de linguagem, além de ser bastante útil no processo de compreensão de suas especificidades discursivas e estruturais; o hibridismo estrutural dos textos multimodais acaba por demandar uma abordagem teórico e metodológica também híbrida.

Metafunção organizacional e multimodalidade

[...] transposition is not just a matter
of finding "semiotic equivalents" ...

Rick Iedema

Introdução

Este capítulo tem por objetivo discutir a Metafunção Organizacional (ou textual) a partir de três perspectivas. Na primeira, 2.1, Conceitos iniciais, na próxima página, realizo uma exposição dos conceitos de Tema e Rema a partir do exposto por Halliday (HALLIDAY, 1994; HALLIDAY; MATTHIESSEN, 2004) para a linguagem verbal. Nesse ponto, reflito sobre a origem desse conceito-chave, destacando o resultado de diversos estudos que relacionam as escolhas nessa metafunção a fatores culturais, coesivos, linguísticos e estruturais do idioma. Em seguida, exploro a proposta de Martin e Rose (2007), que sugere a existência de HiperTemas que poderiam ser definidos como Temas que condicionam as escolhas em textos verbais dando unidade coesiva e criando coerência interna.

Segue-se então 2.2, Metafunção organizacional e a linguagem visual, na página 92, que explora a adaptação desses conceitos gerais aplicados à linguagem visual. Para tanto, inicio minha reflexão em 2.2.1. A imagem estática, na página 92, de forma a discutir como essa

adaptação se dá em imagens estáticas, como gráficos, fotografias e desenhos. Tal reflexão é relevante por trazer alguns conceitos que também se aplicam ao estudo de textos vídeo-fílmicos, uma vez que, como discutirei na seção 3.2, Noções básicas de linguagem vídeo-fílmica, na página 118, esses textos têm muito em comum com a fotografia, dado que a constituição do movimento na tela se dá graças a projeção contínua de imagens estáticas. Por fim, 2.2.2, A linguagem videográfica, na página 96, trata da adaptação dessa teoria para a imagem em movimento. Partindo dos estudos de van Leeuwen (2001), observo como a relação Tema vs. Rema se manifesta em textos audiovisuais de forma a sugerir que não apenas o código visual deve ser levado em conta, o código verbal, a edição e a manipulação técnica do vídeo, assim como gestos, olhares e outros modos de linguagem que dependem da instanciação por meio dos atores, ou indivíduos que participam das interações, também devem ser levados em conta.

O Capítulo se encerra com os Resultados parciais, na página 107.

2.1 Conceitos iniciais

Dentro das três Metafunções definidas por Halliday (HALLIDAY, 1985, 1994; HALLIDAY; MATTHIESSEN, 2004) a Metafunção textual está relacionada a organização da mensagem no nível da sentença. Para Halliday e Matthiessen (2004, p. 64), todas as línguas possuem algum sistema de organização e contribui para a constituição do significado e, consequentemente, para o fluxo de informações em um texto; sistema que Halliday chama de estrutura temática. Assim, toda a mensagem estaria organizada de forma a relacionar o seu ponto de partida, o Tema; e a parte na qual essa informação é desenvolvida, o Rema.

Siqueira (2000, p. 17-18), partindo dos trabalhos de Fries (1983, 1995a, 1995b), afirma que o conceito de Tema deve seu desenvolvimento a Mathesius (1939). Posteriormente essa noção é ampliada por outros intelectuais do Círculo Linguístico de Praga, que estudam a dinâmica da sentença como um evento comunicativo, chamada de Perspectiva Funcional da

Sentença. Dentro dessa perspectiva, o Tema seria identificável como a informação conhecida pelo falante no momento em que instancia sua proposição, um lugar de partida, até certo ponto óbvio para ele (FRIES, 1983; apud SIQUEIRA, 2000). Influenciado por esses estudos Halliday desenvolveu uma noção muito semelhante, mas com uma diferença básica: na GSF Tema e informação dada são conceitos distintos. Essa diferença é importante porque, em cada uma dessas abordagens, a própria definição do que pode ser considerado Tema se modifica (SIQUEIRA, 2000, p. 18).

Para a GSF a sinalização da estrutura temática em inglês acontece pela atribuição de um *status* especial para a primeira parte da sentença; este é então definido como Tema (*What the message is about*). Um exemplo dessa estrutura representada pela transcrição 1, retirada dos vídeos que compõem o corpus de pesquisa, ali o Tema seria representado por "I" ao passo que o rema por "I'm a PC". No caso do inglês e do português, o que também pode ser considerado correto para boa parte das línguas de origem indo-europeia, o Tema se manifesta de forma posicional, sendo o primeiro elemento da oração. Halliday considera como limite para observação dessa unidade o primeiro elemento da oração com valor experiencial (ou seja, que realiza alguma função dentro do sistema de transitividade). Contudo, a sua realização do Tema não se dá apenas pelo sujeito oracional: deve-se ter em mente que embora a forma de sinalização seja posicional, o Tema é antes de tudo um conceito funcional.

- | | | |
|-----|-------------|-------------|
| (1) | I | 'm a PC |
| | Tema | Rema |
-
- | | | |
|-----|-------------|---|
| (2) | You | were made to instimulate a ten-year-old brain |
| | Tema | Rema |
-
- | | | |
|-----|-------------|----------------------------------|
| (3) | Ow! | I have to go, listen to some Emo |
| | Tema | Rema |

Nas transcrições acima (retiradas do vídeo *Meant for Work* (V-10)³⁰), observa-se que a função de Tema é exercida pelo pronome *you* (transcrição 2) que é ator de um processo material. Já em 3, o Tema é uma interjeição que demonstra a desaprovação de PC a respeito do fato de necessitar ouvir música com seus usuários infantis.

Existem inúmeros estudos sobre o Sistema de Tema/Rema nas mais diversas línguas, entre eles estão os de Fries (1983; 1994, 1995a; 1995b), Downing (1991), Trillo (1994), Souza (1997), Siqueira (2000a, 2000b), Gouveia e Barbara (2001; 2003), Barbara e Gouveia (2001), Ventura e Lima-Lopes (2002) e Netz (2007) para citar apenas alguns. Entre os resultados relevantes que esses trabalhos trazem para este estudo temos:

Padrões de escolhas temáticas. Cada gênero parece ter padrões próprios de escolhas temáticas, esses padrões demonstram como a linguagem é organizada e como ela se distribui pelas proposições de um texto.

Valor coesivo de Tema. Se o Tema é um recurso que funciona na organização de informações, incluindo seu fluxo, ele pode ser considerado um elemento coesivo, tornando-se, assim, uma ferramenta de que o autor/falante pode se servir para construção da tessitura textual.

Diferenças entre idiomas. Há variação das escolhas temáticas entre as diferentes línguas, mesmo quando tratamos de um mesmo texto. Essa realidade persiste mesmo quando lidamos com questões de tradução, ou com variações dialetais em uma língua.

Motivação social. A sociedade e seus valores são os elementos motivadores dessas escolhas, uma vez que elas tendem a ser diferentes em cada um dos contextos, mesmo em textos de estrutura similar.

Tema e oralidade. Marcadores conversacionais e marcas de passagem ou retenção de turno podem ser considerados Tema na linguagem oral.

³⁰Todos os vídeos que compõem o corpus desta pesquisa têm a nomenclatura de nomeação explicada na seção 3.1.1, na página 110, além de receberem uma sinopse na seção 3.1.2, na página 114. Para uma rápida referência, aconselho o leitor a dirigir-se ao quadro 1, na página 112.

Marcação X não marcação. A noção de Tema marcado (HALLIDAY, 1994) passa a ser discutida à luz de levantamentos em diferentes gêneros. Em alguns contextos, a ocorrência mais frequente de estruturas consideradas não convencionais ou atípicas por Halliday levou o Tema marcado a ser pensado não pela não canonicidade da estrutura, mas sim pela sua frequência: um tema marcado seria o menos comum, a escolha menos provável.

Estrutura da oração. O Tema pode ser, mas não necessariamente, o sujeito da oração, em muitos casos temos advérbios, interjeições e uma série de outros elementos linguísticos podem realizar essa função.

É dentro desse contexto que Martin e Rose (2007, p. 188-192) refletem sobre o conceito de fluxo de informações. Para os autores, a repetição e as escolhas temáticas são formas de "empacotar" os significados em ondas de informação, sendo que esses pacotes formariam padrões temáticos criados por escolhas que determinam nossa referência, estabelecendo expectativas de leitura. Martin e Rose fazem uma analogia com um conceito normalmente aprendido durante as aulas de redação/composição: o Tópico Frasal, o período inicial de cada parágrafo, responsável por trazer o conteúdo a ser ali desenvolvido. Para os autores, esse fenômeno não ocorre apenas no nível do parágrafo, mas também em nível textual: ele se manifestaria pelo acúmulo de escolhas temáticas, cada acúmulo de escolhas determina um grupo de informações, um "pacote" que pode ser responsável por uma onda informacional. Seguindo esse caminho, segundo Martin e Rose (2007: 193), um estágio de um gênero, por exemplo, seria caracterizado pelo acúmulo de escolhas Temáticas que visa à realização de um Tema maior, ao qual os autores chamam de HiperTema. Dentro desse universo, todas as formas de comunicação verbal seriam condicionadas por esse tipo de fenômeno: os HiperTemas determinam nossa organização de significados. Temas não marcados seriam, assim, definidos por sua tipicidade, estando de acordo com o HiperTema daquele estágio ou texto e sinalizariam a continuidade de uma onda, ao passo que, Temas marcados seriam responsáveis por sinalizar incidentes e, portanto, a descontinuidade.

A partir desse pressuposto, se analisarmos, então, um texto que possui diversas subdivisões, como seria o caso de um artigo acadêmico ou uma tese, deveríamos ser capazes de observar uma relativa estabilidade nos padrões temáticos de cada uma delas. O texto, por conseguinte, seria composto por um HiperTema que se subdivide em diversos SubTemas, cada um representado por um conjunto de Temas "reais", efetivamente instanciados. Algumas divisões formais passam ser, por conseguinte, também manifestadas em proposições que nos ajudam a mapear seu conteúdo. De fato, é importante também ressaltar que nem sempre esse processo de construção de padrões temáticos obedece as divisões formais de um texto; muitas vezes sua observação demonstra que não há necessariamente uma equivalência entre os elementos formais e os funcionais.

Apesar de não desfilar dentro da Metafunção Textual, o fenômeno da coesão textual é deveras importante para o estudo da construção textual e do fluxo de informações. Halliday e Hasan (1976), em seu livro *Cohesion in English*, fazem a análise de sistemas de coesão em textos de linguagem verbal. Seu trabalho é seminal, sendo uma das abordagens mais sistemáticas e importantes desse fenômeno textual, inspirando o trabalho de linguistas como Eugene Winter e Michael Hoey. Dentro do estabelecido pelos autores, a coesão é um fenômeno que permite a estruturação de um texto como tal, diferenciando uma unidade de sentido de um amontoado de frases justapostas. Dentro desse contexto, repetição e sequencialização são os elementos responsáveis por esse fenômeno: o primeiro estaria relacionado as maneiras pelas quais realizamos um mesmo significado no texto, sem que necessariamente a mesma palavra esteja presente inúmeras vezes; já o segundo estaria ligado às relações semânticas que estabelecemos entre as ideias. Exemplos do primeiro seriam a referência — anafórica e endofórica, — a substituição, nominal, verbal e clausal, a colocação, justaposição sistemática de palavras, e a elipse, substituição por zero; já na segunda, teríamos os diversos tipos de marcadores textuais (como aditivos, explicativos etc.), aplicados em seus diversos contextos (HALLIDAY; HASAN, 1976). Essas escolhas coesivas não aleatórias, como demonstra Fries (1992), a escolha e repetição de elementos experienciais são responsáveis pela criação de pa-

drões nos eixos sintagmático e paradigmático, sendo que esses padrões parecem querer levar o leitor a ter uma determinada interpretação de um texto ou de suas partes. Nos resultados do autor, essas repetições se dão especialmente pelos elementos experienciais, as quais ocorrem conjuntamente a relações lógico-conjuntivas, criando um caminho de interpretação preferencial que é oferecido ao leitor.

Todavia, a não presença de alguns elementos visuais ou verbais pode ser tão importante para esse processo coesivo como a repetição. Isso é referente a uma omissão de algum signo, de um significante que deveria estar ocupando uma determinada posição dentro da mensagem, algo que é propositadamente não-dito; a esse fenômeno dá-se o nome de elipse, ou substituição por zero (HALLIDAY; HASAN, 1976, p. 142). Muito do que significamos em uma interação permanece não presente, subtendido pelo contexto comunicacional em que nos encontramos. Para Machado (2007b, p 83-122), a coesão e a coerência em textos vídeo-fílmicos é dependente do ponto de vista e a situação privilegiada e onisciente do espectador. Assim, ao produzir esses textos, seu autor está guiando o olhar do espectador-leitor através de sua objetiva: é ele que conscientemente escolhe quais *takes* e ângulos serão utilizados e em que ordem serão dispostos e, em um certo nível, é ele que nos oferece com quais aspectos das personagens nos identificaremos. Da mesma forma que textos verbais, textos vídeo-fílmicos também ocorrem no tempo x espaço e dependem de ações concatenadas que se acumulam no processo interativo. No caso desses textos, essas ações são representadas pelo que acontece nos diferentes *takes* e *sequências*,³¹ sendo que o estabelecimento dessa coesão se dá não apenas por aquilo que fica, que se mantém, mas também por aquilo que se esquece (MACHADO, 2007, p. 91).

³¹Para definição do vocabulário técnico sobre vídeo, ver a seção 3.2, na página 118, e a seção 4.2, na página 138.



Figura 4: Fluxo de informação em fotografias esportivas (retirado de <http://veja.abril.com.br>)

2.2 Metafunção organizacional e a linguagem visual

2.2.1 A imagem estática

Kress e van Leeuwen (1996, p. 181—229) discutem a possibilidade de adaptação de uma teoria funcional da língua para a linguagem visual. Para os autores é possível realizar uma analogia entre a teoria hallidiana e a análise de outras semioses, centrando a questão da organização da imagem em uma reflexão sobre a composição visual. Nesse momento Kress e van Leeuwen estão interessados em entender como a constituição do quadro em uma imagem estática pode nos levar a compreender como os elementos visuais são escolhidos e combinados na significação. A composição é, então, discutida a partir de três princípios:

1. valor da informação (*information value*);
2. saliência (*salience*);
3. emolduramento (*framing*).³²

³²Talvez emolduramento não fosse a tradução mais adequada para *framing*, mas a escolhi por temer que enquadramento pudesse gerar confusões com a linguagem técnica de vídeo, que discutirei mais a frente (ver seção 3.2, na página 118, e a discussão a respeito da figura, 15, na página 121. Logo, minha opção foi para evitar uma possível ambiguidade terminológica.

O item 1, valor da informação, estaria relacionado à colocação de elementos — participantes, objetos e sintagmas que se relacionam entre si ou com o leitor — de forma a atribuir valores conectados a diferentes partes da imagem: direita vs. esquerda, topo vs. base entre outros. O item 2 se refere a arranjos dentro da imagem cujo objetivo é atrair a atenção dos leitores de forma diferente para cada elemento que a compõe. Isso seria instanciado por meio de recursos da linguagem visual como posicionamento em primeiro ou segundo plano, tamanho e proporção, contraste (incluindo cores) e nitidez. Por fim, o item 3 é caracterizado pela presença de vetores de enquadramento (tais como linhas e objetos) que conectam ou desconectam elementos entre si em quadro; como resultado teríamos elementos que parecem pertencer mais a uma parte da tela do que outra. Segundo Kress e van Leeuwen (1996, p. 183), esses conceitos podem ser aplicados em todas as formas de manifestação multimodal que incluem imagens, tais como fotografias, gráficos e infográficos; além da possibilidade de aplicá-los em imagens em movimento.



Figura 5: Propaganda de notebook Sony (retirada de <http://www.sony.com.br>)

Isso, tem consequências lógicas para a compreensão de como o fluxo de informações nessa semiose, como se pode observar na figura 4, na página anterior. Ali o jogador que está no chão dá um carrinho com o objetivo de tomar a bola logo à frente, essa mesma bola é o foco do olhar do jogador que está de pé. Todos os processos visuais levam em direção a ela, a bola é o centro da disputa entre os dois jogadores e o ponto de partida da nossa leitura dessa imagem. Em segundo lugar temos o jogador de pé que, por estar em uma postura mais ereta em relação ao outro, acaba sendo o segundo elemento mais importante da imagem, tanto

o jogador caído quanto os demais elementos, sem foco e perdidos em segundo plano, são visualmente menos importantes. Essa interpretação é possível porque os princípios descritos por Kress e van Leeuwen (1996) estão presentes na imagem da seguinte forma:

Valor da informação: as relações direita vs. esquerda e em pé vs. deitado são importantes pois são elas que definem a bola como elemento central e o jogador em pé como segundo elemento em importância.

Saliência: os jogadores e a bola, que estão em primeiro plano, são mais salientes que os elementos desfocados; o jogador em pé também é mais saliente que aquele que está deitado.

Emolduramento: a linha do campo, logo à frente, e as pernas dos jogadores formam um vetor que vai em direção à bola. Além disso, o fundo, vemos uma pessoa emoldurada pelo gol.

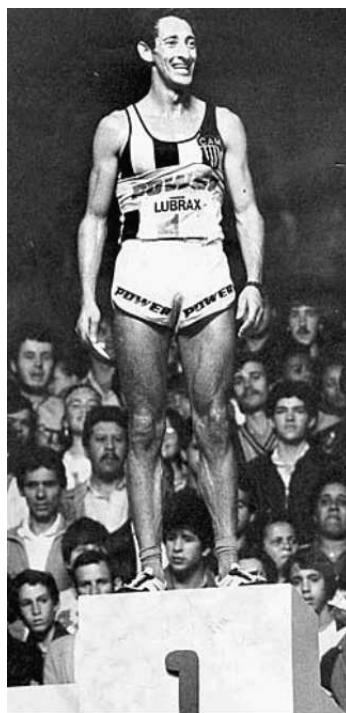


Figura 6: João da Mata no pódio da São Silvestre de 1983. Retirado de <http://www.gazetapress.com.br/>

Apesar do diferente peso composicional, o mesmo pode ser dito em relação a imagem representa pela figura 5, página 93, retirada do sítio de comércio online da Sony do Brasil. No primeiro retângulo temos o anúncio de um notebook representado pela imagem à direta e em posição temática. À esquerda, temos o Rema, uma série de informações verbais que trazem características mais detalhadas do produto, informações técnicas que desenvolvem e ampliam nosso conhecimento sobre o notebook apresentado.



Figura 7: Sequência de animação em Flash — anunciando câmera fotográfica Sony. Retirada de <http://www.sony.com.br>

Não obstante, isso nem sempre parece ser dessa forma, um exemplo é a figura 6, na página 94, que traz uma fotografia do pódio da corrida de São Silvestre em 1983, vencida pelo brasileiro João da Mata. Se pensarmos na área de varredura que normalmente nossos olhos realizam em uma tela ou fotografia, perceberemos que João e o pódio estão centralizados, bem na área nobre de observação. Isso talvez signifique que a informação de sua vitória, que seria nova para quem visse a foto na época de sua publicação, ocupe posição temática, aqui representada pelo primeiro plano. Os populares em segundo plano representariam, assim, o

Rema. Isso parece ser um resultado de um processo de saliência que ajuda a dar o maior valor da informação que está no primeiro plano. Apesar de Kress e van Leeuwen estarem certos, acredito que existem elementos técnicos relacionados à produção imagética que são necessários para construção desses significados. Como veremos a seguir, em diversos momentos desta tese, acredito que esse tipo de análise deve estar relacionada a uma observação dos elementos de linguagem audiovisual de forma a entender como o fluxo de informações se dá no nível da instanciação.

2.2.2 A linguagem videográfica

Como já discutido anteriormente, os princípios organizacionais que governam diferentes sistemas sociossemióticos, sozinhos ou em interação, deverão ser necessariamente diferentes. No caso específico deste trabalho, que se debruça sobre a linguagem vídeo-filmica, a questão é ainda mais complexa, isso porque um filme ou vídeo, como discutiremos na seção 3.2, na página 118, é caracterizado por uma série de imagens em movimento. Essas imagens sofrem transformações na linha do tempo, podendo tanto contar uma história, no caso de peças ficcionais ou documentários, ou de puro valor estético e contemplativo, como é o caso das vídeo-artes. Baldry e Thibault (2006, p. 189-190) também levantam questionamentos similares, mostrando que a solução mais interessante para esse tipo de linguagem seria considerar a característica topológico-contínua da linguagem em textos visuais, que contrasta diretamente com o caráter predominantemente topológico-categorial da linguagem verbal. Em outras palavras, os autores colocam que é necessário levar em conta a questão da organização espacial dos textos visuais: muitas vezes é ela que determina qual é o real fluxo de informações. Além disso, segundo Baldry e Thibault (2006), o texto vídeo-filmico sofre uma transformação dinâmica, fazendo com que a estrutura posicional não seja uma solução satisfatória para todas as situações, sendo pertinente considerar também a saliência informacional. Para os autores³³

³³Baldry e Thibault não discutem a fotografia como linguagem autônoma. De fato, ela é um ponto de partida tanto para o surgimento do vídeo e do cinema, ambas seriam suas subsidiárias e dela herdariam boa parte de seu cabedal de constituição de linguagem. Por conta disso, uma vez que boa parte dos elementos sugeridos por Baldry e Thibault (2006) — e por outros autores como Van Leeuwen (2001), Kress e Van

esse é um fenômeno relacionado ao ganho de ênfase, ou seja, a informação gradativamente vai se tornando aparente, ganhando foco e importância em uma dada tomada. Essa saliência, ou estado focal de informação, seria o que determinaria se uma informação é ou não nova, se desfilará como Tema ou Rema. Van Leeuwen (2001, p. 95-96) coloca que a saliência pode ser alcançada por vários recursos, entre eles estariam a mudança de foco, tamanho, maior detalhe de textura, contrastes na cor, aproximação ou distanciamento, perspectiva além de valores culturais específicos de uma dada comunidade.

Tomemos como exemplo o anúncio representado pela figura 7, página 95, um vídeo em *flash* que anuncia uma câmera fotográfica digital. A cada giro do seletor que se encontra à esquerda, uma nova fotografia, contendo um ângulo diferente da máquina aparece. Posteriormente, essa imagem se desloca para o centro e depois para a esquerda. Simultaneamente uma mensagem verbal, que contém novas informações sobre o produto, surge e também vai se deslocando gradativamente. O movimento da câmera chama nossa atenção devido à sua variação de tamanho e de posição, fazendo com que nossos olhos acompanhem sua trajetória pelo *banner* animado. Até o fim desse movimento, ela é a informação nova, o que só se modifica à medida que as cores que compõem a mensagem verbal no centro se tornam mais opacas e se fixando na imagem. Essa mensagem verbal se torna estática depois da câmera, chamando assim nossa atenção e se caracterizando como a segunda informação nova. Se fôssemos pensar em termos da estrutura temática, teríamos um Tema (o seletor), como informação dada, e dois Remas, a câmera e os dizeres, que se caracterizam como tal graças a seu movimento e sua saliência.

Van Leeuwen (2001, p. 94-104) caminha pela mesma seara, discutindo uma série de recursos que podem ser utilizados para garantir o fluxo de informações em textos vídeo-fílmicos. Além da saliência, o autor introduz outros conceitos: 1) enquadramento; 2) conjunções temporal, espacial e comparativa. Ao tratar do enquadramento, o autor discute as formas pelas

Leeuwen (1996, 2001), Lemke (1998) entre outros —, também seriam aplicáveis, com certas limitações, à fotografia, seria importante em trabalhos futuros partir desta linguagem como um elemento de apoio para análise multimodal da visualidade.

quais o fotógrafo — termo geral, utilizado tanto para o operador de máquinas fotográficas como para o iluminador de cinema e vídeo — cria a composição de seu quadro ou revela a informação desejada por meio de movimentos de câmera dentro desse mesmo quadro. Assim, a informação dada vs. nova/Tema vs. Rema será bastante dependente de questões estéticas e da interação entre elementos em cena e a objetiva. No caso específico do enquadramento, ele pode ser o responsável por conectar diferentes entidades dentro de um mesmo quadro, o que acontece por intermédio de linhas. Essas linhas podem ser representadas por algum elemento físico dentro da cena ou mesmo serem abstratas. Ao mesmo tempo, diferentes linhas e usos de cores similares podem indicar a conexão entre as entidades.

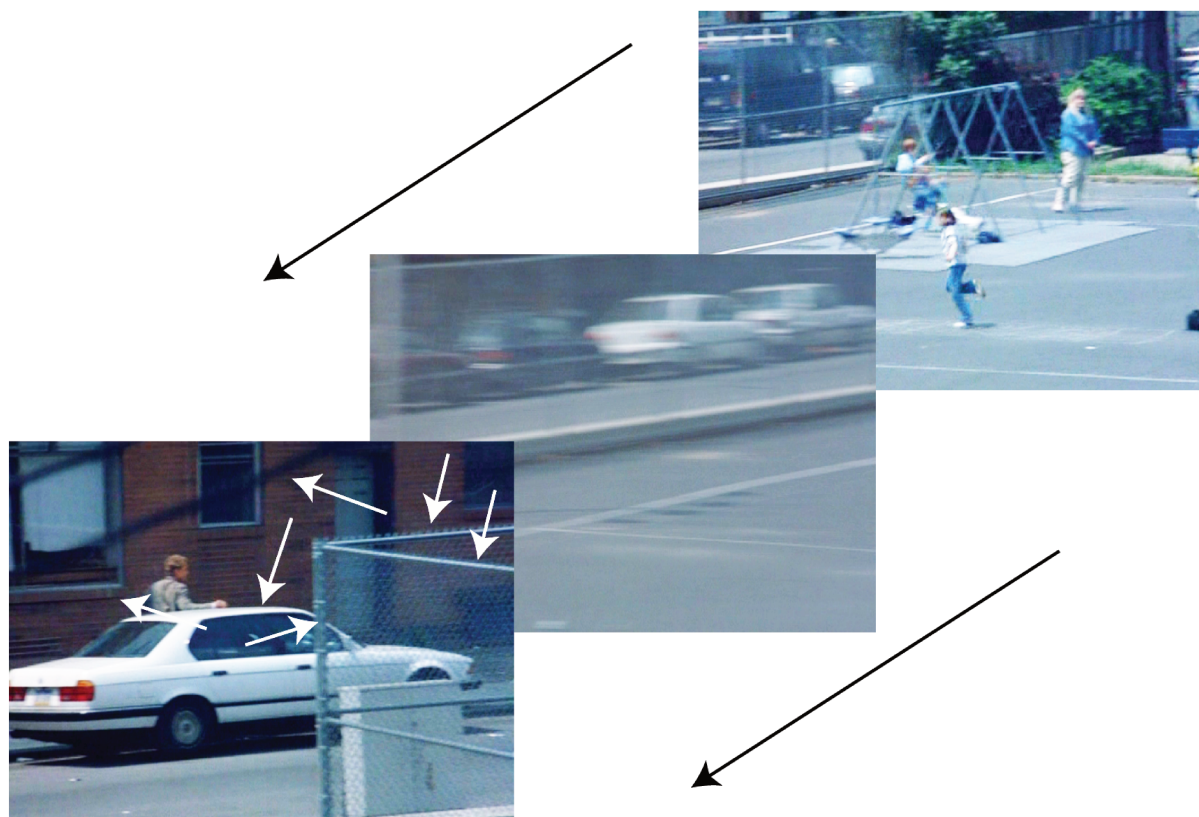


Figura 8: Sequência de *O lenhador* (KASSELL, 2004)

Um exemplo de como essas relações podem ocorrer estão na figura 1, na página 31, que traz frames³⁴ retirados do filme *O lenhador* (KASSELL, 2004), um caso no qual a saliência

³⁴Nesta tese, as terminologias de origem francesa e anglo-saxã são utilizadas de forma equivalente no que tange às unidades videográficas.

e o enquadramento contribuem juntos para o estabelecimento da informação em relevância. Ali há três quadros distintos, no primeiro, um playground de uma escola primária, a câmera está em plano geral,³⁵ todos os participantes — diversas crianças e uma mulher adulta — dividem o espaço em posição de quase igualdade, apenas uma criança mais à frente ganha algum destaque pela sua posição. Logo em seguida temos uma rápida panorâmica para a esquerda (ver setas), a qual revela um homem saindo do carro. Nesse quadro, uma série de vetores formados por objetos dentro da cena: as grades, o carro, os fios de alta-tensão, as janelas, todos excluem outros espaços, deixando o homem em uma situação focal. Ele é a única entidade em movimento no quadro final. Uma questão importante e que não pode ser esquecida é relativa aos movimentos corporais dos atores na figura 8: tanto o distanciamento da criança como a saída do carro nos ajudam a construir a percepção de que a criança e o homem se encontrarão nos próximos frames, o movimento é um fator determinante do nosso preenchimento daquilo que não está dito no quadro. Em outras palavras, a informação parece fluir não apenas porque há um enquadramento momentaneamente estático, ela é dependente de movimentos humanos e técnicos de câmera, que nesse caso incluiria a passagem do playground para a rua. Como veremos no capítulo 4, Reflexões sobre a transcrição multimodal, na página 127, tanto os movimentos corporais como a linguagem técnica parecem ser igualmente importantes para a construção de significados nos textos audiovisuais, isso porque, ao instanciarmos um vídeo estamos, na verdade, interferindo nos dois tipos de ação.

Na figura 9, página 100, temos a formação de diversas linhas que enquadram a atriz como informação principal dessa sequência, entre elas estão a janela que dá fundos para a cozinha, o sofá, ela se encontra no vértice em um dos cantos. No caso da saliência, observamos que ela está centralizada na imagem, sendo a única entidade completamente em foco. Ao mesmo tempo, pode-se observar que outro elemento importante parece nos dar pistas sobre como a informação flui nesse quadro: o olhar. É nítida a formação de um vetor entre o olhar dos atores: pelo fato de a mulher estar em foco, presumimos que ela é o Tema da imagem,

³⁵Para definição do vocabulário técnico sobre vídeo, ver a seção 3.2, na página 118, em especial a discussão a respeito da figura 15, na página 121.

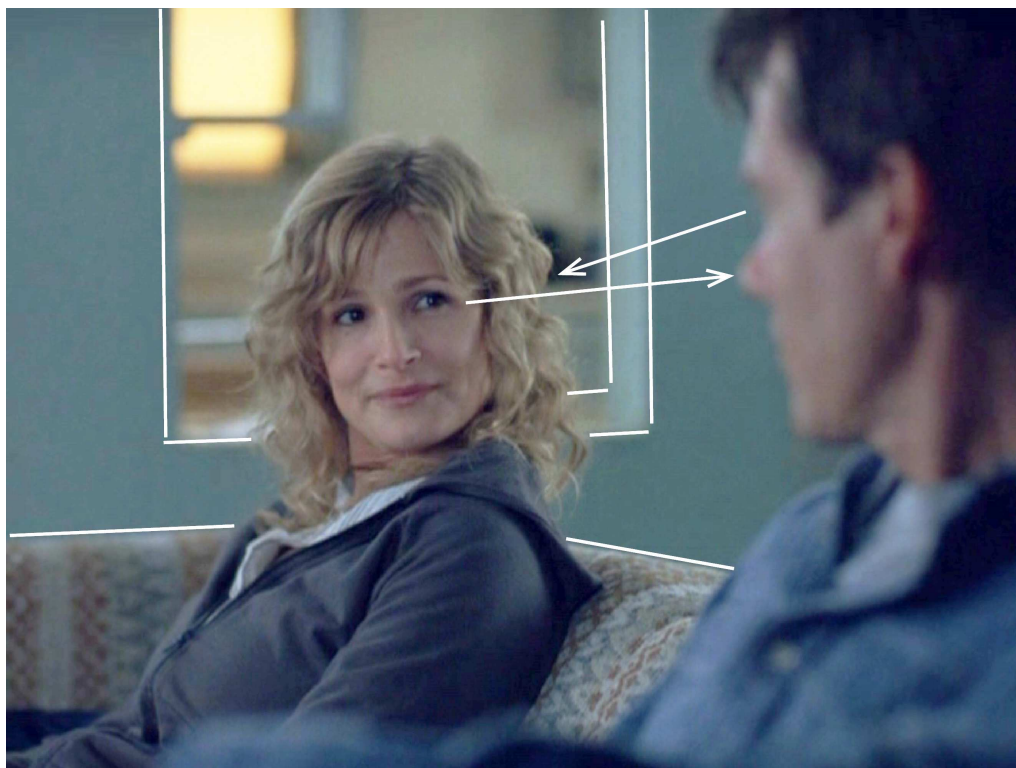


Figura 9: Frame isolado de *O lenhador* (KASSELL, 2004)

e que dela emana a informação que efetivamente precisamos nos ater. Ao mesmo tempo, é seu olhar dirigido que parece mostrar que essa informação não está sendo dirigida para o espectador, mas sim para a outra personagem em quadro, parcialmente representada e fora de foco. A importância disso recai sobre o fato de que esses vetores deixarem claro, também, que o nosso papel nessa interação é de olhar a distância, o foco da interação recai exclusivamente sobre as personagens que são retratadas no vídeo. Apesar da relevância das ações das personagens, elas não desfilam como um elemento importante na literatura referente ao fluxo de informações; em outras palavras, apesar de muito importantes para esta tese, trabalhos como o de van Leeuwen (2001) tratam a imagem em movimento de forma muito similar à estática. Uma abordagem que, por conseguinte, parece limitar a análise em alguns pontos. Por conta disso, sugiro que tantos os elementos técnicos como as ações efetivas dentro da imagem em movimento sejam levadas em conta no processo de análise e

compreensão de como o significado é construído em textos audiovisuais.³⁶ Ao se observar a figura 9, perceber-se-á, também, a existência de uma série de outros fatores que também contribuem para a construção desse significado, entre eles estaria a existência da relação entre plano e contra plano,³⁷ que nos permite uma visão subjetiva do ator que olha esses vetores, além da profundidade de campo, algo que o vídeo e o cinema herdam da fotografia e que ajuda a definir a saliência da informação.

Saliência e enquadramento são valores posicionais, os quais dependem de duas variáveis: o lugar que os elementos ocupam na tela e o movimento realizado pela câmera. Contudo esses não são os únicos elementos que garantem o fluxo de informação em textos vídeo-fílmicos, questões temporais também devem ser levadas em conta. No caso da conjunção temporal, transições e eventos podem ser conectados (VAN LEEUWEN, 2001, p. 97); somos capazes de perceber que os diversos quadros de uma cena estão ocorrendo em um determinado lugar no tempo, formando assim uma sequência. Entre as estratégias, o autor destaca três: 1) sequência temporal de ações; 2) simultaneidade e 3) *flashback* (ou lembrança).

No caso da primeira temos estratégias que, quando utilizadas, parecem levar o espectador do texto vídeo-fílmico para a próxima ação. Entre elas estariam movimentos panorâmicos, *travellings* ou mesmo planos contínuos, além de cortes, normalmente partindo de um plano mais aberto para um mais fechado. Mas isso não seria uma regra, podemos observar exemplos em que um corte seco³⁸ entre diferentes imagens podem já realizar esse tipo de transição. No caso da segunda, tem-se um recurso no qual duas ações são retratadas como acontecendo ao mesmo tempo, o que é retratado, normalmente, por *takes* intercalados. Um exemplo de ambas pode ser ilustrado pela figura 10, página 102, retirado do *Life on Mars* (NALLURI et.

³⁶ Uma discussão mais apurada de como isso pode ocorrer está no capítulo 4, Reflexões sobre a transcrição multimodal, na página 127, e nos que se seguem.

³⁷ Define-se uma relação plano vs. contra plano quando há uma sequência de *takes* em que temos a perspectiva das duas personagens que travam o diálogo se alternando na tela. Apesar de os atores, dependendo do enquadramento, não precisarem estar ali — pode-se contracenar apenas com a objetiva —, dá-se a impressão de que as personagens estão uma diante da outra, uma vez que essa variação respeita seu ponto de vista. Essa variação de enquadramentos costuma ser de 180º, respeitando convenções de perspectiva de cada personagem.

³⁸ Corte de câmera sem a utilização de efeitos.

al., 2006). Se observarmos o frame A, percebemos que um grupo de furgões está passando por um beco, ao passo que no frame B temos o mesmo beco já a distância, sendo que todos os furgões já estão estacionados. Ali, o que permite a percepção da mudança de enquadramento não é, apenas, a justaposição de *takes*, mas também a presença de diversos elementos que se repetem em diferentes posições, demonstrando a modificação no tempo mas não no espaço, como os furgões, a escada e a parede de tijolos do beco, dispostos em um plano muito mais fechado em A do que em B. O frame C também indicaria uma sequência, agora marcada pela mudança de plano — de geral para médio-fechado — e de espaço — externo para dentro do furgão. Nesse mesmo frame, observa-se que o policial da esquerda (jaqueta escura) olha para a direita e faz um movimento com o rosto (frame D), como se acompanhasse algo. Nesse momento, um texto verbal dito por essa personagem chama a atenção dos demais (frame E) que passam a olhar na mesma direção. Há, assim, um corte para o próximo frame (frame F), no qual vemos um novo furgão entrando no mesmo beco do qual eles há pouco saíram. Nessa sequência, teríamos um caso de simultaneidade: enquanto estão com seu furgão parados, discutindo, os policiais presenciam e observam um outro automóvel que lhes parece suspeito.

Duas são as estratégias utilizadas nessa sequência de *takes*: 1) os vetores formados pelo olhar dos policiais, que apontam para algo que acontece fora do espaço que se encontram e 2) o corte seco, que mostra um utilitário a passar em um ângulo em um eixo compatível com



Figura 10: Sequência de *Life on Mars*



Figura 11: Sequência de *Ônibus 174* (PADILHA, 2003)

o olhar dos policiais. Em minha opinião, ambas as estratégias são igualmente importantes, ao manter o eixo ótico do corte, podemos imaginar os atores observando o beco, que nos foi mostrado menos vezes: cabe, no fundo, ao espectador dessa sequência ser capaz de observar e fazer a conexão entre elas. De fato, não parece ser apenas a justaposição e o ângulo dos olhares que pode nos informar qual a relação entre os cortes, acredito que a repetição funciona como uma importante estratégia, a noção de simultaneidade se estabelece ao imaginarmos que os vetores do olhar se dirigem à imagem justaposta, que com algumas modificações se repete. Não se pode esquecer, também, o papel que da linguagem verbal na cena representada pela figura 10, página 102, uma vez que as falas anunciam a chegada do furgão esperado pelos policiais.

Já a conjunção espacial, por seu turno, pode ser entendida como a descrição de um objeto por meio da revelação de suas partes. Nesse sentido, van Leeuwen (2001, p. 100) sugere

duas estratégias, uma relacionada à conjunção de partes que constituem um todo, e outra à mudança entre planos, do geral para o detalhe ou vice-versa, em um mesmo objeto. Como exemplo da primeira, poderíamos citar uma sequência de *takes* que descrevem diferentes espaços dentro do Rio de Janeiro, retirado do documentário *Ônibus 174* (PADILHA, 2003), na qual temos o mar, os morros e as ruas da zona sul carioca (figura 11, página 103). Essa sobreposição de imagens definiriam elementos constituintes da cidade, dando ao espectador uma ideia dos espaços que ela contém, numa clara migração entre eles.

Com relação à segunda estratégia, podemos tomar como exemplo a sequência em que Ray Milland está prestes a realizar a ligação que levará ao assassinato de sua esposa Gracy Kelly em *Disque M para Matar* (HITCHCOCK, 1954/1982/2004), aqui representada pela figura 12, página 105. Como pode-se observar, a figura 12 traz uma sequência de imagens que caminha de um plano mais aberto (frame A), que representa o momento no qual a personagem entra na cabine telefônica, passando por um plano intermediário (frame B) que demonstra a intenção de utilizar o telefone e se encerram em dois planos mais fechados (frames C e D), cuja função é detalhar o telefone já apresentado nos dois planos anteriores, além de mostrar a discagem da letra M, ação que dá nome ao filme. Um ponto importante a ser lembrado é que uma cena como a retratada em 12, na próxima página, não é simplesmente o resultado de uma justaposição de imagens, há a necessidade de um grande planejamento prévio que faz com que essa construção imagética seja significativa. Percebe-se que o fluxo de informações se baseia no fechamento contínuo da objetiva, que elimina todos os objetos e até mesmo o ator do quadro, direcionando nosso olhar para aquilo que o diretor acha importante: o disco do telefone e, posteriormente, o dedo que lho disca.

Ao mesmo tempo, seria correto pensar que cabe a nós, espectadores, sermos capazes de perceber, por meio de um processo metafórico, alguns processos representacionais importantes: o dedo representa o homem que realiza a ação, ao mesmo tempo que o telefone é o que nos resta de todos os demais elementos que estão naquele contexto. Apesar de não vermos toda a sala ao redor, imaginamos que a personagem esteja ali, que o telefone esteja na ca-

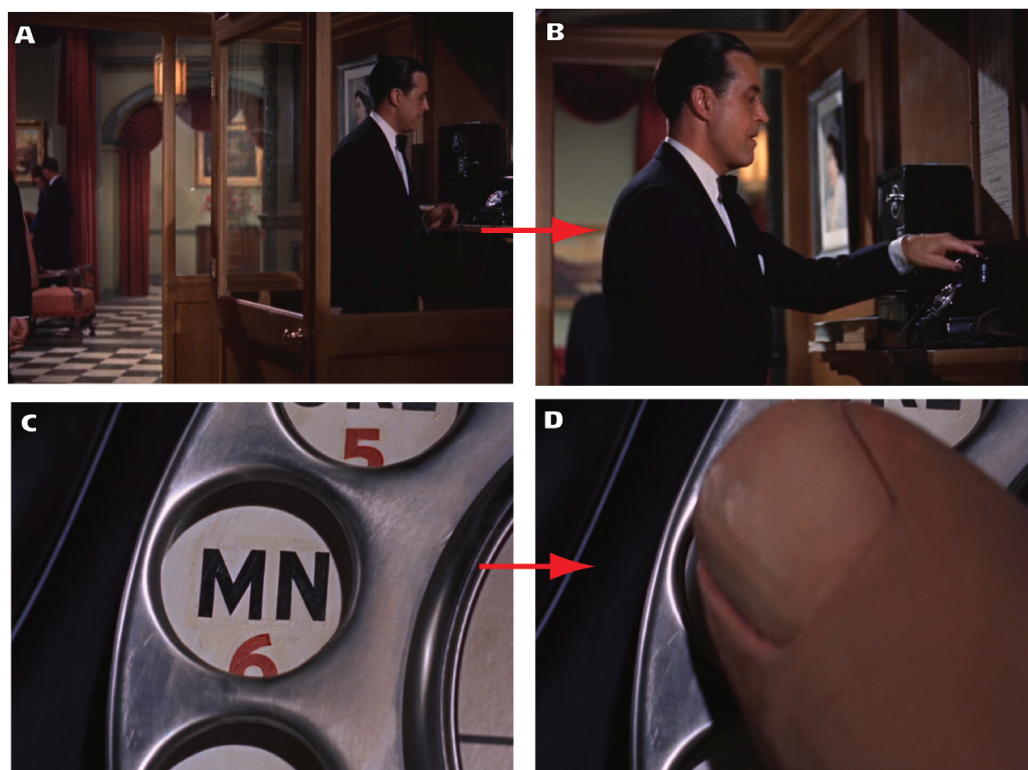


Figura 12: Sequência de *Disque M para Matar* (HITCHCOCK, 1954/1982/2004)

bine e que o disco pertença àquele telefone. Isso é nitidamente um processo de coesão por apagamento que se estabelece, as imagens fechadas eliminam redundâncias visuais da mesma forma que nós o fazemos na escrita. Tal fato permite ao diretor chamar nossa atenção para a informação que ele julga conveniente em um dado momento.

Já o *flashback* seria uma estratégia utilizada para mostrar uma ação que deve ser entendida como ocorrendo antes do que a antecede (VAN LEEUWEN, 2001, p. 99). Esse tipo de estratégia pode ser utilizada para estabelecer a relação não apenas entre *takes* ou sequências, mas também em relação a cenas inteiras. Apesar de van Leeuwen (2001) não mencionar em seu trabalho, esse tipo de conjunção normalmente se dá pelo uso de efeitos modificadores da cor — como utilização de preto e branco — ou de transições de vídeo ou sinalizações em áudio — o que incluiria temas musicais próprios para *flashback*.

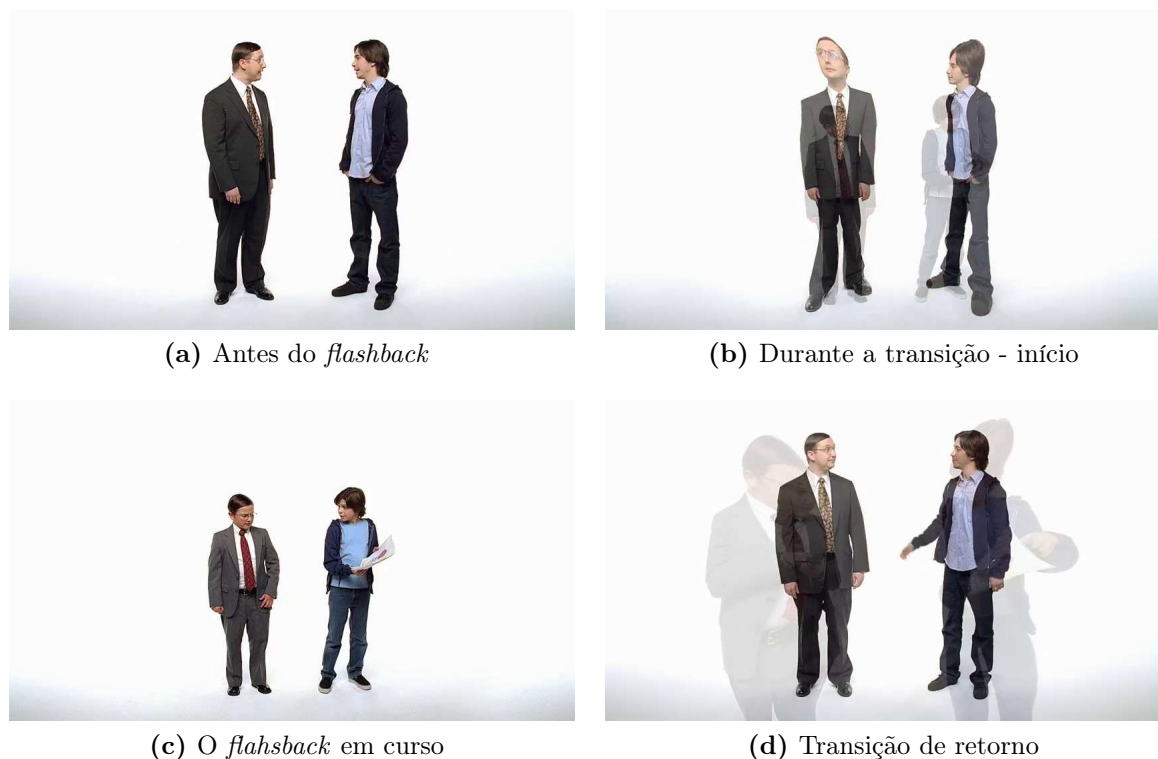


Figura 13: Sequência de comercial em *Flashback*

Um exemplo seria a figura 13, página 106, que traz um vídeo publicitário da campanha "Get a Mac". Ali temos um efeito muito convencionalizado para sua introdução: uma transição de efeito líquido. Inicia-se a sequência com o frame representado em 13a, que mostra os atores em nossa contemporaneidade; segue-se então o efeito que vemos em 13b, a imagem parece se desmanchar em ondas e temos, agora, dois atores infantis vestidos a exemplo dos adultos de outrora. Essas crianças interagem em 13c, na página corrente, até que um novo efeito ocorra e tenhamos o retorno para a cena anterior 13d, nesta página. É importante ressaltar que essa transição foi introduzida por uma proposição em linguagem verbal, que se lê transcrita em 4, a seguir. Esses recursos lembram a artificialidade do código ou a materialidade do filme, trazendo à baila algumas convenções presentes nessa linguagem.

(4) **PC (de paletó e gravata):** Some things never change Mac... never change...

Mac (de roupas esportivas): What's that supposed to mean?

A questão do *flashback* merece algumas considerações. Em primeiro lugar, no caso do vídeo representado na figura 13, na página anterior, o elemento verbal é importante por ser o desencadeador, dele depende o início do processo de transição que, aparentemente, não parece necessariamente modificar o fluxo de informações no texto. Ele parece ser um processo digressivo que apenas dá uma informação a mais a respeito do que ali ocorre, sem necessariamente modificar a estrutura da informação. Na verdade, em minha opinião, *flashbacks* são processos de introdução de subepisódios na construção videográfica. Naturalmente eles podem, ou não, como é o caso do vídeo que aqui discuto, modificar padrões do fluxo por se tratarem, muitas vezes, de um "filme dentro do próprio filme", mas acredito que ele tenha uma estrutura própria. No geral, acredito que o *flashback* não modifica o fluxo, apenas introduz uma situação paralela na qual um possível alteração possa ocorrer. Nesse ponto ele seria um recurso muito mais relacionado com a construção do enredo, não modificando a forma de condução de uma interação em específico.

2.3 Resultados parciais

Este capítulo teve por objetivo discutir a Metafunção Organizacional (ou Textual) a partir da possível aplicação de seus conceitos-chave na imagem estática e em movimento. Iniciei a discussão trazendo alguns conceitos abalizares, observando como essa teoria é normalmente aplicada em textos verbais (sejam escritos ou falados), discutindo o conceito de Tema como um elemento funcional e expresso posicionalmente na maioria das línguas indo-europeias (seção 2.1 na página 86). Além disso, observou-se que ele pode ser um elemento importante no processo de coesão textual, sendo que muito da significação do texto se dá por meio dos acúmulos das escolhas temáticas, os diversos estudos da literatura mostram que cada gênero possui um padrão diferente, que está relacionado às características contextuais e discursivas de cada um deles.

Em um segundo momento, busquei teorias que adaptavam esse conceito à imagem estática. Partiu-se do modelo sugerido por Kress e van Leeuwen (1996), que se baseia na observação de três elementos-chave: 1) valor da informação; 2) saliência e 3) emolduramento. Com relação a este modelo, observou-se que ele permite não apenas que a noção posicional seja levada em conta na análise de imagens estáticas, mas também ajuda a inserir alguns critérios básicos de análise de elementos técnicos da imagem estática.

Seguiu-se então para a discussão de como esse modelo poderia ser aplicado à imagem em movimento, com especial atenção ao trabalho de van Leeuwen (2001). Como se pôde concluir, seu trabalho possui raízes muito fortes na teoria desenvolvida por Kress e van Leeuwen (1996), e tenta realizar essa adaptação também apostando em um modelo triádico. Além de considerar os critérios já desenvolvidos para a imagem estática, ele fia sua argumentação em três noções: 1) sequência temporal de ações; 2) simultaneidade e 3) *flashback* (ou lembrança).

Apesar desse modelo de estruturação da informação em imagem em movimento dar conta de uma série de situações, ele carece de alguns refinamentos importantes. Ele parece não levar em conta gestos e movimentos de atores, suas atitudes perante outros atores, locações e os objetos em cena, além de não se ater à questão das escolhas no nível da edição. É relevante também observar que repetições — elementos que parecem constantes entre os vários frames de um filme ou vídeo — também não são exploradas por esse modelo, que parece se debruçar no quadro congelado como uma fotografia.

A imagem em movimento, certamente, necessita de uma abordagem que não veja apenas esse fotograma estático. Foi essa constatação que levou à discussão teórico-metodológica presente no capítulo 6, na página 213, lugar em que me debruço sobre o processo de construção coesiva e sobre o fluxo de informações nos vídeos de "Get a Mac".

Definição do corpus, temática da campanha e elementos de linguagem videográfica

Analysis of visual records of human experience is a search for pattern and meaning, complicated and enriched by our inescapable role as participants in that experience.

M. Collier

Introdução

Este capítulo tem por objetivo apresentar o corpus de estudo deste trabalho, além de trazer algumas considerações importantes sobre a linguagem vídeo-fílmica. Na seção 3.1, Descrição do corpus, na página seguinte, tenho por objetivo discutir a composição do corpus de pesquisa em duas perspectivas. A primeira, seção 3.1.1, Composição do Corpus e procedimentos de coleta, na próxima página, explica a sua composição, os critérios de seleção e recorte na sua escolha, assim como seus procedimentos de coleta; segue-se, então, a seção 3.1.2, Temática da campanha e dos vídeos do corpus principal, na página 114, na qual são

discutidos tanto a temática da campanha publicitária que originou os vídeos que fazem parte desta pesquisa, assim como de cada um dos 12 vídeos que compõem o corpus de análise.

Segue-se então 3.2, Noções básicas de linguagem vídeo-fílmica, na página 118, que traz algumas definições técnicas importantes sobre a área de vídeo, vocabulário que será utilizado no decorrer da análise.

3.1 Descrição do corpus

3.1.1 Composição do Corpus e procedimentos de coleta

Esta pesquisa contou com um corpus de vídeos publicitários da empresa Apple Computer, Inc., parte da campanha "Get a Mac". Esta campanha teve seu andamento entre maio de 2006 e outubro de 2009, para produção de vídeos, (Nudd, 2011) e abril 2010 (para o encerramento da campanha),³⁹ ou seja, ainda em andamento e em processo de lançamento dos vídeos no início desta pesquisa. Os vídeos utilizados neste trabalho se dividem da seguinte forma:

1. doze vídeos da campanha estadunidense, sendo quatro de cada fase da campanha, esses vídeos serão o corpus principal;
2. doze vídeos da campanha japonesa;
3. quinze vídeos da campanha do Reino Unido.

Para viabilizar uma análise que fosse capaz de mapear o maior número possível de modos de linguagem e funções, optei dividir o corpus de vídeos da campanha em diversos grupos, cada um com uma função específica dentro desta tese. No primeiro, estão doze vídeos da campanha estadunidense que foram analisados em profundidade, eles são o corpus de pesquisa que serviu como base para o estudo, estando presentes em todos os aspectos da análise. Uma

³⁹Ver <http://www.brainstorm9.com.br/12767/diversos/um-tributo-a-campanha-get-a-mac/>, acessado em: 10 de outubro de 2011.

questão importante a ser citada é o recorte temporal realizado nesta tese, os vídeos do corpus principal agregaram lançamentos até maio de 2009; no período estudado foram lançadas mais de 42 peças publicitárias, incluindo vídeos e *banners*, para a campanha "Get a Mac". Outros vídeos que foram lançados após o início da análise de dados não foram considerados no corpus devido à necessidade de um recorte, que se justifica pela extensão da campanha; outros tipos de mídia como *banners*, por exemplo, também não foram considerados. A ideia de manter um corpus principal é possibilitar uma análise exaustiva de um grupo menor de vídeos, algo que talvez não fosse possível em todos os sessenta e cinco vídeos e vinte e dois *banners* — total de mídia audiovisual produzida para a campanha —, dada a grande quantidade de dados.

As campanhas do Japão e do Reino Unido foram utilizadas apenas na seção 5.3, na página 199, na qual foram comparadas com a campanha estadunidense. Como veremos no referido capítulo, essas duas campanhas surgiram como desdobramento da campanha dos EUA, sofrendo algumas alterações temáticas e troca de atores para adaptá-la ao público-alvo. Como já colocado na introdução deste trabalho, na página 1, essa comparação visa a observar até que ponto as diferenças de produção podem diferenciar as peças publicitárias nipônicas e britânicas das estadunidenses e, além disso, discutir se há alguma modificação no discurso devido à internacionalização da campanha. Os vídeos da campanha estadunidense foram exibidos na televisão, em revendas especializadas Macintosh nos EUA e, durante a época de veiculação dessa campanha (2006-2010), também disponibilizados em um hotsite (sítio específico para divulgação de uma campanha publicitária).⁴⁰

O quadro 1 traz um resumo dos doze vídeos do corpus principal. Na primeira coluna temos a fase da campanha que cada um dos vídeos representa, os quatro primeiros datam do início da campanha, os quatro em posição central do meio da campanha, e os últimos de sua etapa final; o título dado a cada um não foi aleatório, mas sim proveniente do próprio sítio da Apple. O critério para divisão das fases da campanha foi subjetivo, uma vez que não há indicativo desse tipo de segmentação na exibição das peças publicitárias, seja na TV ou no

⁴⁰<http://www.apple.com/getamac>. Acessado em: 05 de abril de 2010. Agora já está desativado.

Quadro 1 Composição do corpus principal

Vídeos			Sinopse
Fase	Nome	Número	
Início	<i>Viruses</i>	V-01	Computadores Apple são imunes a vírus
	<i>Out of the box</i>	V-02	Computadores Apple funcionam assim que saem da caixa
	<i>Restarting</i>	V-03	Computadores Apple não travam
	<i>Networking</i>	V-04	Mostrar que os computadores Apple são mais simples de se conectar a periféricos
Meio	<i>Better results</i>	V-05	Mostrar que os Macs fazem melhor tarefas multimídias
	<i>Security</i>	V-06	Computadores Apple são mais seguros
	<i>Choose a Vista</i>	V-07	Simplicidade do sistema operacional da Apple
	<i>Surgery</i>	V-08	Simplicidade nas atualizações de software e hardware da Apple
Fim	<i>Boxer</i>	V-09	PC encara a relação entre Mac e PC como uma luta a ser vencida
	<i>Meant for Work</i>	V-10	Mostrar que os PCs só foram feitos para trabalho
	<i>Fat HD</i>	V-11	Mostrar que computadores Apple não vêm cheios de programas inúteis
	<i>Breakthrough</i>	V-12	Mac e PC discutem seus problemas em uma sessão de terapia

hotsite. Usou-se como base a data de lançamento e a data limite de coleta dos dados para corte dos dados, dividiu-se o número de vídeos por três e escolheram-se vídeos aleatoriamente em cada uma das fases.

Cada um dos vídeos do corpus principal recebeu uma numeração própria, de acordo com sua ordem de lançamento em relação aos demais; assim *Viruses* é identificado com V-01 porque ele foi lançado previamente a *Security* (V-06) e a *Meant for Work* (V-10), esses números serão frequentemente utilizados como referências rápidas a estes vídeos no trabalho. Na última coluna temos uma rápida sinopse de cada um desses vídeos, apenas para referência, pois a sinopse será ampliada na seção 3.1.2, Temática da campanha e dos vídeos do corpus principal, na página 114. Os vídeos da campanha japonesa e britânica, quando forem citados,

terão seu enredo discutido, oportunamente, durante a análise. A escolha de vídeos de diferentes fases da campanha tem por objetivo generalizar os resultados dentro de sua extensiva duração.

Os procedimentos de coleta constituem no download do vídeo através de software específico. O download dos vídeos não era incentivado no sítio da campanha — que os distribuía de forma muito similar ao *YouTube* e *Vimeo* — e necessitava do uso de programas específicos para sua realização. Minha opção foi utilizar a extensão complementar *Fast Video Download*⁴¹, instalada e utilizada no *browser* Firefox, para plataforma Macintosh, ela permitiu que os vídeos fossem salvos sem perda de qualidade. Visitas ao sítio da campanha aconteceram periodicamente, cada vídeo salvo recebia o nome oficial dado pela campanha, sendo acrescido de um número que identificava sua posição de publicação. Nenhuma edição foi realizada nos vídeos.

Esses vídeos foram armazenados em disco rígido, onde permanecem para análise posterior. Ainda com relação a coleta, vale a pena ressaltar que esta pesquisa só utilizará vídeos obtidos do próprio sítio da Apple; esta medida visa a evitar que possíveis alterações ao material original sejam desavisadamente adicionadas a esta pesquisa.⁴² Além disso, não foram consideradas versões dubladas em qualquer língua, a pesquisa foi realizada apenas com vídeos originais, o que se justifica pela minha preocupação em evitar que traduções não oficiais dos vídeos causassem algum problema de distorção no conteúdo; um número enorme de paródias, redublagens e modificações terem ocorrido em vídeos distribuídos extraoficialmente por decorrência do sucesso dessa campanha. Apesar da própria Apple oferecer dublagens

⁴¹Disponível em <https://addons.mozilla.org/pt-br/firefox/addon/fast-video-download-with-search/>.

⁴²Apesar do sítio da campanha já ter sido retirado do ar, ver nota de rodapé 40, na página 111, eles são ainda largamente encontrados em sítios de compartilhamento de vídeo, como Youtube (<http://www.youtube.com>) e Vimeo (<http://www.vimeo.com>). O portal sobre comunicação chamado *Brainstorm9* realizou uma compilação de 5 minutos sobre essa campanha, que pode ajudar o leitor na compreensão dos temas gerais do vídeo, basta acessar: <http://www.brainstorm9.com.br/12767/diversos/um-tributo-a-campanha-get-a-mac/>. Acesso em: 10 de janeiro de 2012. Outras possibilidades de recuperação de tais vídeos estão no sítio <http://http://adverlicio.us/>, uma rede social que "salva" anúncios publicitários para referência na área de comunicação, e no sítio Adweek, <http://www.adweek.com/adfreak/apples-get-mac-complete-campaign-130552>, um portal online especializado em crítica e análise de mercado publicitário.

em português da campanha dos EUA, optou-se também por não utilizá-los dado o fato de alguns desses vídeos se basearem seu enredo em uma tradução visual e metafórica de uma expressão linguística; exemplos que talvez ilustrem minha preocupação nesse ponto sejam os vídeos *Viruses* (V-01) e *Out of the box* (V-02), ver resumo dos vídeos no quadro 1, na página 112, e na seção a seguir.

Em termos de suas especificações técnicas, cada um dos vídeos é disponibilizado em alta definição (HD), associado ao plugin Quicktime, de fabricação da Apple e instalável em qualquer browser de internet, sendo que seu CODEC é o H264, padrão de muitos vídeos distribuídos na rede. Sua distribuição se dá em vídeos em cores, com trilha sonora musical que marca o seu encerramento. O hotsite específico da campanha que possuía um player proprietário na parte central da tela, além de um menu para a escolha dos vídeos.

3.1.2 Temática da campanha e dos vídeos do corpus principal

A campanha "Get a Mac" — "Compre um Mac", em uma tradução livre do inglês — foi uma campanha publicitária da empresa fabricante de computadores e programas estadunidense Apple Inc, criada pela agência publicitária TBWA\Media Arts Lab; os vídeos foram produzidos entre 2006 e 2009 e a campanha foi exibida até abril de 2010. Foram no total 66 anúncios em vídeo, mais 15 vídeos em formato de *banner*, os quais foram disponibilizados apenas em portais de tecnologia na web (NUDD, 2011). Os vídeos foram exibidos em diversos países de língua inglesa (lojas Apple e televisão), além de estarem disponíveis no sítio oficial da campanha; algumas versões dubladas oficiais também existiam, como era o caso do Brasil.⁴³ Campanhas específicas foram realizadas para o Reino Unido e para o Japão, como veremos a seguir na seção 5.3, na página 199, essas campanhas possuíam alguns *remakes* (vídeos que imitavam os originais dos EUA) mas também alguns vídeos originais.

Como já colocado anteriormente, o enredo das peças é centrado em um homem de roupa informal que se autointitula Mac, e outro, com trajes corporativos que se autonomeia PC. Os

⁴³<http://www.apple.com.br/getamac>. Acessado em: abril de 2010 e agora já desativado.

dois têm, assim, uma pequena interação na qual os atributos de Mac e PC são comparados. PC é caracterizado como formal e educado, mas deveras preocupado com o trabalho, além de frustrado com as incapacidades de sua plataforma em comparação com Mac, retratado como alguém mais leve, uma plataforma mais simples e menos problemática, que se importa tanto com trabalho como com lazer. A campanha foi premiada em 2007, com o Effie Worldwilde Awards, na categoria principal.⁴⁴

Todos os vídeos ocorrem dentro do mesmo cenário, um fundo branco infinito no qual duas personagens, homens que se automeiam Mac e PC, discutem questões computacionais. Como podemos observar na figura 14, página 116, invariavelmente, Mac ocupa a parte esquerda do vídeo e inicia a interação apresentando-se (*Hello, I'm a Mac*), seguido quase que imediatamente por PC, que realiza sua apresentação (*And I'm a PC*). É a partir desse ponto que a peça audiovisual se desenvolve.

Cada um dos filmes publicitários possui em média 30 segundos, sendo que os artistas e a trilha sonora jamais são substituídos; apenas algumas personagens são acrescentadas dependendo do tema discutido. O padrão de vestimenta também não se modifica, PC está sempre vestido de terno e gravata, ao passo que Mac sempre se veste informalmente, com Jeans, camiseta, tênis e moletom.

Sua temática geral é a superioridade dos produtos Apple em relação aos fabricados pelas demais empresas, computadores e softwares padrão IBM-PC, por conta disso atenção crítica é dada à Microsoft Inc., fabricante do software Windows e principal rival comercial da Apple. O cenário minimalista e a ausência de equipamentos computacionais nos filmes fazem com que boa parte da ação aconteça verbalmente, as duas personagens discutem de forma metafórica questões técnicas que, invariavelmente, acabam por mostrar a inferioridade técnica de PC. Isso faz com que metáforas sejam usadas constantemente, boa parte da discussão ali presente se constrói a partir da metaforização de processos computacionais que passam a ser representados no domínio humano, como um computador com vírus que é representado por

⁴⁴<http://www.effie.org/winners/showcase/2007/1826>. Acesso em: 10 de dezembro de 2010.

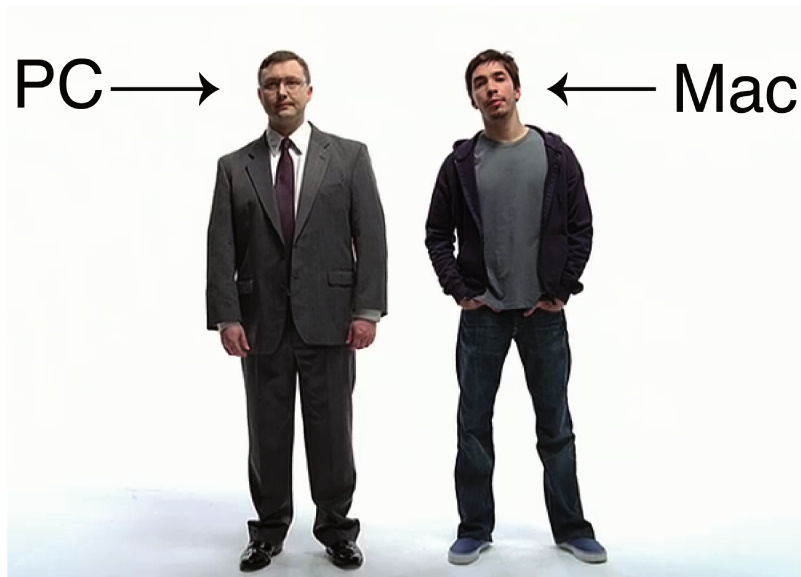


Figura 14: Frame ilustrativo do vídeo *Restart* (v4)

um homem gripado, ou uma máquina fotográfica do Japão que é representada por uma mulher de origem japonesa. Cada um dos vídeos traz um diferente tipo de comparação: nenhum deles repete a temática ou mesmo as metáforas.

Um ponto importante a ser observado é que os vídeos tratam o sistema operacional Windows como sendo sinônimo da plataforma PC. Essa escolha parece ter caráter comercial e faz parte do processo de argumentação dos vídeos, apesar de reducionista, ao igualar a plataforma ao sistema operacional PC está-se, na verdade, retirando da discussão outros sistemas operacionais, como Linux, Unix e OS2.

Como visto no quadro 1, na página 112, são doze os vídeos principais que fazem parte desta campanha. A seguir temos uma sinopse expandida de cada um deles:

***Viruses* (V-01)** PC foi infectado por um novo vírus computacional, aqui representado por um resfriado. Mac oferece ajuda para PC que, inicialmente, a recusa para não infectar Mac. Este, por sua vez, diz que Macs não são infectados por vírus, e PC desmaia representando uma falha de sistema do computador.

Out of the box (V-02) Mac aparece dentro de uma caixa branca (marcando a identidade de cor dos produtos Apple) e PC em uma caixa marrom claro. Eles começam a discutir o que farão quando “saírem da caixa”. Mac diz que já vai logo começar a trabalhar, ao passo que PC tem de passar por todo um processo de montagem antes de se tornar funcional. Mac salta da caixa e se despede de PC, que tem de esperar seus pedaços em outras caixas.

Restarting (V-03) Mac e PC iniciam uma conversa discutindo o que têm em comum. Mas PC não consegue funcionar, fica congelando e travando, o que é representado pela posição estática da pessoa. Mac sai de cena para chamar os profissionais de TI.

Networking (V-04) Mac e PC dão as mãos para representar sua capacidade de formar uma rede. De repente uma mulher japonesa entra em quadro e dá as mãos para Mac, de forma a representar sua capacidade de conexão com qualquer equipamento. PC não se conecta com a câmera porque não fala japonês, mostrando-se confuso.

Better Results (V-05) Mac e PC discutem suas habilidades em realizar vídeos caseiros. A modelo Gisele Bündchen entra em quadro para representar o filme de Mac. Ela insiste para PC mostrar o seu, mesmo resistindo ele o faz. Seu vídeo é representado por um homem de perucas e roupas iguais às de Gisele.

Security (V-06) Um homem de preto fica entre Mac e PC, representando um programa Firewall. A interação se mostra cansativa porque PC tem de autorizar todas as mensagens que entram e saem de seu sistema, algo desnecessário nos firewalls da plataforma Mac.

Vista Choose (V-07) PC tem de escolher uma entre as 6 versões de Windows Vista e, para auxiliar escolhe uma roleta. Mac retruca, dizendo que possui apenas uma versão que serve para tudo. Ao rodar a roleta, PC perde a vez e Mac diz: mas não foi você quem fez essa roleta?

Surgery (V-08) PC aparece com uma camisola de hospital e diz que vai sofrer uma cirurgia de instalação do Windows Vista. Diz que esse é um processo complicado e ele tem medo de não voltar, por isso doa seus periféricos a Mac. Mac retruca dizendo que os upgrades de Mac são simples e não trazem problemas.

Boxer (V-09) PC é apresentado como um lutador de boxe, dizendo que não irá ser derrotado sem lutar. Mac diz que eles não estão competindo; as pessoas simplesmente estão mudando para um computador melhor.

Meant for Work (V-10) PC está coberto de adesivos, reclama que foi comprado para uma família, tendo, portanto, de ouvir música e jogar ao invés de fazer trabalho de escritório.

Fat HD (V-11) PC aparece gordo, lento e com sérias dificuldades de movimentação, representando seu HD cheio de programas inúteis que são instalados em um novo computador PC. Mac diz que só vem com os programas realmente necessários para o trabalho e diversão.

Breakthrough (V-12) Mac e PC vão ao terapeuta. Lá tentam conversar com PC dizendo que os problemas não são culpa dele, mas da estrutura da plataforma que é licenciada para vários fabricantes. PC levanta dizendo que entendeu tudo e que a culpa é de Mac.

3.2 Noções básicas de linguagem vídeo-fílmica

Esta seção tem por objetivo discutir algumas definições técnicas importantes sobre a área de vídeo, vocabulário que será utilizado no decorrer da análise.

Vídeo ou cinema são, na verdade, uma projeção sequencial de imagens, sendo que conseguimos percebê-las como movimento graças a uma incapacidade do olho humano. Cada vez que essas imagens são projetadas uma após a outra, nossa retina guarda resquícios da imagem anterior, sendo que isso que nos permite vê-las da forma como a vemos. Em outras palavras,

ela mal apagou uma imagem e já é exposta a uma nova: nossos olhos, dada a velocidade em que essa exposição acontece, são incapazes de separar cada imagem e as tratam como um elemento único. No caso do cinema, para que isso possa acontecer de forma natural, essa projeção deve ser de 24 imagens por segundo, no vídeo digital esse número sobe para 29,97, sendo que o mínimo de imagens projetadas para que esse fenômeno aconteça em qualquer situação é 15. Cada uma dessas imagens recebe nome técnico de frame ou quadro, essa é uma unidade importante, pois é ela que serve como parâmetro de análise nessa linguagem.

No momento da gravação, uma câmera de vídeo colorida não é uma, mas sim três. Isso porque todas as cores que se formam na tela são o resultado da combinação entre azul, vermelho e verde, que usualmente chamamos de sistema RGB (*Red*, *Green* e *Blue*) (ZETTL, 2004). As câmeras modernas capturam essas cores em chips chamados de CCDs, um para cada cor base. Nossa televisão e nosso computador são capazes de interpretar essas combinações de cores, transmitidas em forma de elétrons, posteriormente retraduzidos em imagens em nossa tela. Zettl (2004, p.35) coloca que uma televisão, um monitor, projetor ou qualquer equipamento que receba e traduza ondas de elétrons em vídeo seguem o mesmo padrão de formação de imagem.

Em termos técnicos, nossa televisão ou monitor é um tradutor de códigos magnéticos, sendo que cada tipo de televisor (analógico ou digital) tem uma taxa diferente de atualização da imagem que vemos na tela (ZETTL, 2004). No caso específico deste trabalho, analiso vídeos cuja distribuição se dá via internet, o que implica utilização de tecnologia digital para distribuição e gravação. A diferença entre os dois tipos de sinal está na forma como o sinal é constituído: no caso do vídeo analógico temos uma gravação contínua sobre uma fita magnética, ao passo que na digital, toda a informação é transformada em um código binário, formado por algarismos 0 e 1. Entretanto, vale ressaltar que as mesmas normas para a combinação RGB são válidas para o sistema digital, a mudança se dá apenas na forma de codificação e exibição.

Como coloca Zettl (2004, p. 42-44), apesar de o sinal analógico ser perfeito em termos de representação da imagem e do som, a escolha do digital como padrão de mercado ocorre devido a algumas características próprias desse padrão. Entre eles estariam:

Qualidade. Mídias digitais são menos propícias a deterioração por agentes naturais. Além disso, sempre há problemas na cópia de mídias analógicas: à medida que são feitas cópias a partir de uma outra cópia a informação se perde. Essa perda pode ser medida em termos de um conceito chamado de geração: uma cópia está a uma geração do original, uma cópia feita a partir de uma cópia está a duas gerações, e assim por diante. No caso do vídeo digital, essa questão é resolvida, uma vez que todas as cópias de um original sempre serão idênticas a ele (MANOVICH, 2001).

Compressão. No caso do vídeo poderia ser definida como um re-empacotamento de dados, porque nem todo o conteúdo de um arquivo digital é necessário para que sua qualidade seja mantida. Em termos bem genéricos a compressão implica “retirar” informação redundante de um arquivo, deixando apenas aquela estritamente necessária, facilitando o armazenamento e a transmissão.

Facilidade de manipulação. Vídeo e áudio digitais são mais facilmente manipuláveis, facilitando os processos para edição.

O vídeo e o cinema têm muito em comum com a linguagem fotográfica, e dela herdam boa parte da sua linguagem de gravação/filmagem e edição. Um dos elementos que caracterizam um vídeo é a distância que a lente objetiva está em relação aos modelos que estão sendo gravados, essa distância pode variar, gerando uma diferença em relação à quantidade de detalhes que somos capazes de apreender de uma cena. Dentro da literatura relacionada a vídeo e cinema (BARRET, 2004; MILLERSON, 1994; SADUN, 2003; WATTS, 1984; ZETTL, 2004) pode-se observar que o padrão de linguagem de cada filme é fruto do uso de uma série de enquadramentos; sua função parece revelar aquilo que o diretor do trabalho pretende. Esses enquadramentos são de tipologia limitada, o que os permite serem utilizados como

guia rápido de referência entre profissionais e críticos da área (MILLERSON, 1994). Apesar de alguma variação terminológica, é possível estabelecer, pelo menos, quatro tipos de planos, representados na figura 15.⁴⁵

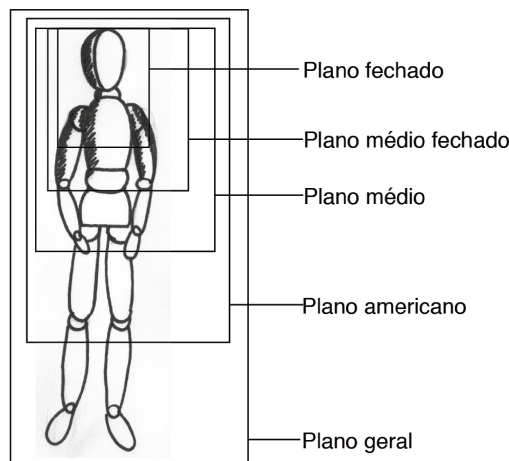


Figura 15: Planos em vídeo, cinema e fotografia

No chamado plano geral a objetiva está completamente aberta e todos os modelos ou elementos gravados estão à mostra, nesse caso temos um grau pequeno de detalhe do modelo e o quadro é composto por todos os elementos que compõem o ambiente gravado. No plano seguinte (americano), temos o joelho da personagem como medida básica de enquadramento, a objetiva está um pouco mais fechada e temos menos detalhes do entorno. No plano médio, o modelo ocupa boa parte do quadro, teríamos ali a parte superior do tórax e, em alguns casos, partes dos membros superiores. Por fim, no plano fechado ou close, temos uma objetiva fechada, mostrando apenas uma parte específica do modelo. A divisão de planos ocorre não apenas em relação à quantidade de informação presente na tela, mas pode também variar de acordo com a profundidade de campo. Esta última pode ser definida como o efeito causado pela proximidade ou distância daquilo que gravamos, ou fotografamos, e que está no foco de nossa lente. Essa variável está relacionada a quão clara e definida uma cena, ou parte dela, está. A distância focal determina se um elemento estará focado (nítido) ou desfocado (embaçado), ou seja, o campo de visão de uma lente. Um quadro pode ser composto de

⁴⁵Todas as figuras que não tiveram sua origem assinalada foram desenhadas especialmente para esta tese.

forma a demonstrar primazia de parte daquilo que se mostra. Na literatura técnica da área (BARRET, 2004; MILLERSON, 1994; SADUN, 2003; WATTS, 1984; ZETTL, 2004), essa variação está condicionada por duas variáveis: 1) a abertura do zoom da lente e 2) a quantidade de ações simultâneas em um quadro. Dessa forma, quanto mais fechada estiver a objetiva, menos informação temos em quadro; o menor número de informações faz com que menos ações ocorram simultaneamente chamando nossa atenção. Além disso, vale ressaltar que a proximidade da objetiva também influencia em outra variável, a distância focal.

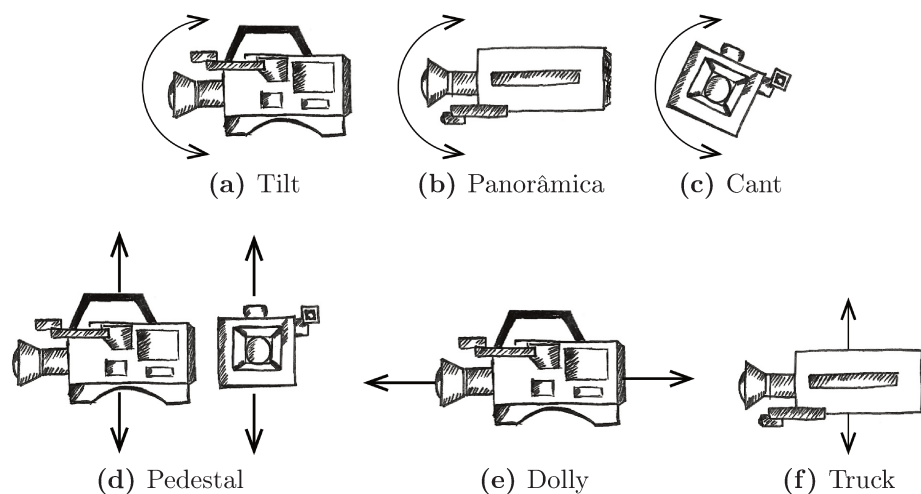


Figura 16: Movimentos de câmera

Vídeo e cinema não são apenas quadros estáticos, são também câmera em movimento. Esses movimentos (ver figura 16, na página 122) são caracterizados por seguir um padrão de linguagem específico, que também funcionam como um elemento de referência para estudiosos e profissionais da área. De acordo com Zettl (2004)⁴⁶ esses movimentos de câmera poderiam ser definidos como:

Tilt (up/down): movimento circular de levantar e abaixar a câmera, imita o girar de uma cabeça para cima (figura 16a).

⁴⁶Apesar de nem todos os movimentos definidos nesta lista estarem presentes nas obras analisadas nos capítulos a seguir, a definição de todo vocabulário me parece importante por mostrar ao leitor a totalidade e complexidade dessa linguagem.

Pan (Panorâmica): movimento horizontal da câmera da direita para a esquerda ou vice-versa, implica um movimento no qual o operador está parado, o que temos é um movimento dentro do eixo, como um girar de coluna (figura 16b).

Cant: derrubar a câmera para a direita ou para a esquerda, imitando o tombar de uma cabeça (figura 16c).

Pedestal: movimento para cima ou para baixo no mesmo eixo da câmera, seguindo a coluna do tripé ou equipamento próprio (figura 16d).

Dolly: mover a câmera para frente ou para trás, aproximando-se ou distanciando-se do modelo (figura 16e).

Truck ou Track: (no Brasil também chamado de *travelling*) consiste em movimentar a câmera para a direita ou esquerda utilizando-se um trilho ou equipamento próprio (figura 16f).

Boa parte desses movimentos não acontecem simultaneamente em todas as obras audiovisuais.⁴⁷ Na verdade, esses elementos de linguagem são escolhidos de acordo com a necessidade de expressão presente em cada uma das produções. Esses conceitos serão importantes para que se possa entender com mais acuidade os processos de expressão utilizados nos vídeos da campanha que constituem o corpus de pesquisa, uma vez que, como veremos no capítulo seguinte, a constituição da linguagem audiovisual é um elemento importante na construção dos significados.

Vídeo gravados com objetivo profissional não são realizados na sequência narrativa em que são exibidos para o seu público. Normalmente a gravação é dividida por locações, locais de ação narrativa, produzidos um por vez. Um exemplo seria um episódio de uma novela que ocorra em três locações: uma casa, um supermercado e um teatro; o processo de gravação ocorreria em cada um dos locais separadamente; gravariam-se todas as cenas na casa,

⁴⁷Neste trabalho fiz um opção por citar obras que definem a linguagem do vídeo e do cinema a partir apenas de uma perspectiva técnica, ao invés de embasar minha análise em estudos como Xavier (1983; 1977) e Betton (1987).

independente de sua posição na história, e assim sucessivamente em cada uma das locações. Finalmente o produto final é montado de forma que as cenas sejam colocadas na ordem narrativa. Toda vez que um trecho é "emendado", temos um corte, perceptível pelo salto na imagem que vemos. Esse corte pode ser mantido sem nenhum tratamento, o que se chama "corte seco", ou sofrer algum tipo de adição de efeito de forma a não percebermos o salto, algo que chamamos de transições. Entre os diversos tipos de transição, poderia destacar:

Mix/Crossdissolve/Crossfade. Quando dois frames se fundem, um anterior se apaga completamente, enquanto o seguinte surge delicadamente.

Fade (in/out). Quando uma imagem inicia um *fade out*, ela passa a se desvanecer até tornar-se totalmente preta; ao passo que *fade in* representaria o desaparecer gradualmente da cor para se formar a imagem.

Liquid. Dá a impressão de uma submersão do quadro na água, enquanto a outra imagem surge lentamente. Utilizado para simbolizar a passagem do tempo.

Wipe. Um quadro "empurra" o outro para fora, tomando seu lugar.

Vale ressaltar que o vídeo é um sistema de imagens técnicas que precisa de um suporte para sua existência, ou seja, para que possamos gravá-lo ou assisti-lo há a necessidade de um aparato técnico especializado. No caso da gravação existem as câmeras que podem ser tanto profissionais como amadoras, as primeiras são maiores e utilizadas apenas por indivíduos que trabalham na indústria televisiva, cinematográfica e publicitária, sendo que sua compra é realizada apenas com empresas especializadas, já as últimas são mais familiares ao público em geral e podem ser encontradas em qualquer loja de eletrônicos. A diferença básica, nos dias de hoje, está principalmente na quantidade de recursos, nas suas opções de operação e na bitola utilizada. As câmeras amadoras, de forma a facilitar sua operação, costumam a ter alguns recursos de pré-edição, como efeitos e transições, além de ajustes técnicos já pré-programados; câmeras amadoras comumente não possuem a possibilidade de ajustar foco

e zoom manualmente. Já as câmeras profissionais costumam a ter todos os ajustes manuais, além de poucas opções de efeitos e pré-edição — pressupõe-se que um profissional realizará a edição *a posteriori* da gravação —, além de possuírem mais recursos para a qualidade do áudio.

A bitola é o suporte físico efetivo da gravação, pode ser uma fita — existem diversos modelos, cada um utilizado em uma época diferente — ou um disco (CDs, DVDs); no caso de câmeras que gravam em arquivos digitais, a diferença está na qualidade do arquivo (melhor nas profissionais) e no CODEC utilizado. A palavra Codec é a abreviatura para Code e Decode e se refere a programas que codificam e decodificam vídeos em sistemas digitais. Cubitt (2008) coloca que a profusão de CODECs disponíveis na internet leva a uma falta de padronização. Esse fato faz com que diferentes distribuidores possuam parâmetros quase proprietários de compressão; como consequência a democratização efetiva do conteúdo fica prejudicada pela não só constante necessidade de atualização, como também pela atualização constante do sistema. Se não bastasse isso, muitos desses CODECs são proprietários e não gratuitos. Mesmo havendo câmeras amadoras capazes de gravar em ótima qualidade, o padrão *Full HD*⁴⁸ ainda não é o padrão, dado o tamanho dos arquivos e as exigências de hardware.

Como podemos ver na figura 16, na página 122, uma câmera de vídeo se divide em duas partes, o corpo e a objetiva. O corpo se define como a unidade de processamento dos conteúdos captados pela lente objetiva. Segundo Zettl (2004, p. 49) a lente é responsável pelo que a câmera pode ver, sendo classificada de acordo com sua distância focal (termo definido a seguir), sendo que a qualidade ótica da lente é responsável direta pela qualidade do vídeo. Uma boa câmera não captará uma boa imagem se a objetiva for de qualidade inferior. A distância focal pode ser definida como aquela entre a parte interior da lente e o objeto a ser

⁴⁸O padrão HD (*High Definition*), traduzido para o português como alta definição, poderia ser caracterizado como uma forma de gravação e exibição de vídeo que permite uma apreensão dos detalhes com uma qualidade muito superior a mídias em SD (*Standard Definition*, ou definição padrão), utilizada nos players (DVDs, B-Rays) e televisores convencionais. O HD é fruto da interação entre diversas tecnologias: do suporte de gravação, da sua forma de codificação, da proporção de tela (4:3 ou 16:9) e do número de linhas por quadro (720, 1080, entre diversas outras).

gravado; ela é ajustada de acordo com o lugar em que o objeto está: quanto mais perto, menor será essa distância e haverá menos "espaço" para que o foco seja realizado.

Reflexões sobre a transcrição multimodal

The central issue for a corpus-based multimodal analysis of texts has to do with the formulation of appropriate units of analysis and the various intermediate levels of analysis that are relevant to the analyst's purposes [...]

A. Baldry & P. Thibault

Introdução

Este capítulo tem por objetivo expor um modelo de análise para textos multimodais, em especial os vídeo-fílmicos, objeto de estudo desta tese. Sua inspiração tem como fonte principal a necessidade de desenvolver critérios e parâmetros para uma análise multimodal em nível microtextual, entendendo como os diferentes modos de linguagem são conjugados de forma a construir o significado em textos multimodais. A necessidade de desenvolvimento deste tipo de modelo se justifica não apenas pela construção de modelos analíticos que nos permitam compreender a construção de tais textos, mas também pela premência no desenvolvimento de modelos de letramento que deem conta de uma realidade midiática que congrega

diversas formas de expressão de significados, todas elas motivadas culturalmente. Acredito que o desenvolvimento de critérios para análise desses textos seja um primeiro passo para a compreensão de seus fatores sociais e culturais; sendo que sua expressão dentro das linguagens múltiplas que utiliza é motivada por eles. Logo, parto do texto audiovisual de forma a criar uma ponte e os níveis (ou *strata*) de instanciação e os que estão no nível do conteúdo, ver figura 1 na página 31, parece-me ser um caminho bastante profícuo.

De forma a trilhar este objetivo, este capítulo se inicia na seção 4.1, Estabelecendo procedimentos de transcrição, na próxima página, que traz à baila uma reflexão sobre como sistematizar dados em um estudo multimodal. Essa seção se divide em duas partes, na primeira, sistematizo o levantamento e transcrição de dados. Baseado nos estudos de Baldry e Thibault (2006), Hellwig, Uytvanck e Hulsbosch (2008) e Norris (2004), defino um procedimento de anotação que ocorre em duas fases, uma manual e outra por software. Não obstante haja redundância nesses procedimentos, ela se mostrou importante por ajudar na transcrição dos diversos modos de linguagem presentes no vídeo. Em 4.2, Constituição: elementos criadores de significado, na página 138, concentro-me na forma como os diferentes modos são agregados aos vídeos para criação de seus significados, tentando discutir como os níveis mais abstratos se manifestam em escolhas dentro de cada modo de linguagem. Essa seção é importante pois esse princípio teórico-metodológico norteará todas as análises desenvolvidas nos capítulos 5, na página 163, e 6, na página 213, além de constituir uma das principais contribuições desta tese para a análise de vídeos multimodais. Em 4.2, observo que as escolhas dos modos de linguagem não são aleatórias em pelo menos dois níveis: o constitucional e o combinatório. O primeiro está relacionado às manifestações do discurso em um modo de linguagem, demonstrando como escolhas ali podem ser justificadas pelo significado social ou situacional a ser expresso; já o segundo é a combinatória entre os diferentes modos de linguagem, sendo responsável por fazer com que os modos de linguagem funcionem em uníssono para a construção do significado da proposição visual. Nessa seção também são discutidos quais são os procedimentos de observação dos diferentes modos de linguagem, ten-

tando desenvolver um caminho sistemático de análise. A seguir, na seção 4.3, na página 149, discute-se de forma empírica como esse modelo de análise poderia ser aplicado no corpus de estudo deste trabalho, analisando-se vídeos da campanha.

Já a seção 4.4, Algoritmo final de análise, na página 158, traz o algoritmo de análise do corpus de estudo; um resumo procedural de como o tratamento dos dados ocorreu. Apesar de ser uma elucidação dos passos realizados, ele é importante por revelar qual a sistemática de análise que ajudou a materializar em um estudo empírico as reflexões colocadas em 4.1 e 4.2. Seguem-se então os Resultados parciais, na página 160.

4.1 Estabelecendo procedimentos de transcrição

Durante o processo de realização e planejamento dessa pesquisa, sentiu-se necessidade de se desenvolver um sistema de transcrição multimodal que permitisse dar conta do maior número possível de modos de linguagem e, ao mesmo tempo, fosse capaz de ser aplicado às particularidades dos textos videográficos aqui estudados. Buscou-se então um modelo misto, que traz elementos de três outros estudos: Baldry e Thibault (2006), que são responsáveis por criarem uma metodologia de análise comparativa entre os diversos modos de linguagem na linguagem videográfica, Hellwig, Uytvanck e Hulsbosch (2008), que desenvolveram o programa ELAN 3.9,⁴⁹ específico para criação de anotações complexas em arquivos digitais de áudio e vídeo, e Norris (2004), que desenvolve uma metodologia para anotação de diversos modos de linguagem, como gestos, olhares e a proximidade entre interactantes.

Textos vídeo-fílmicos se desenvolvem através do tempo e são gravados em algum tipo de suporte, como fitas, arquivos digitais ou película e, a exemplo de uma interação gravada em uma fita de áudio, para que a análise possa ser realizada, é necessária uma transcrição. Não obstante, dada a natureza da análise multimodal, outras variáveis estão em jogo, uma vez que os diversos modos devem ser, de alguma forma, recuperáveis para o analista. Ao decidir por este tema de estudo, a dificuldade em se registrar as diferentes informações em um sistema

⁴⁹Disponível em <http://www.lat-mpi.eu/tools/elan/elan-description>. Acesso em dezembro de 2011.

claro e simples logo se mostrou premente. Baldry e Thibault (2006, p.174-181) discutem essa questão propondo uma metodologia de transcrição baseada em uma tabela composta por seis colunas e cada uma delas traz uma informação relevante para o processo de análise. Dentro do modelo proposto pelos autores, essas colunas seriam: (1) tempo; (2) frame; (3) imagem visual; (4) ações sinestésicas; (5) trilha sonora; e (6) interpretação metafuncional. A coluna 1 seria responsável por especificar o tempo em segundos do vídeo, o que poderia ser determinado pelo player utilizado. Já a coluna 2 traz uma imagem referente ao frame que corresponde ao tempo mostrado na coluna 1, ao passo que a coluna 3 tem por objetivo trazer informações analíticas referentes ao frame disposto na coluna anterior. Essas informações são necessariamente seletivas: nem todos os dados presentes em um frame podem ser transcritos, tampouco representados em linguagem verbal. Na coluna 4, encontraríamos as notações referentes às ações físicas, como olhar, movimentos corporais, sorrisos etc. Já a coluna 5, trilha sonora, é referente a toda a banda sonora, o que incluiria música, fala e os mais diversos tipos de sons que podem estar presentes. Essa opção deve-se ao fato de os autores considerarem todos esses elementos como um único fenômeno, sem necessariamente esbarrar em uma divisão quase matemática desses códigos, além de argumentarem em favor da não supremacia da fala em relação aos demais modos sonoros. Por fim, a coluna 6 responderia por uma tentativa de interpretação a toda essa semiose em termos de significados metafuncionais.

Esse modelo em colunas foi aplicado na análise inicial de todos os vídeos que compõem o corpus. A partir dele iniciou-se o processo de transcrição utilizando o software ELAN 3.9 (HELLWIG; UYTVANCK; HULSBOSCH, 2008). O ELAN (ver figura 17, na página oposta) imita as interfaces de programas de edição e montagem não linear de vídeo digital como Final Cut Pro, desenvolvido pela Apple Inc., e o Adobe Premiere, desenvolvido pela Adobe. A figura 17, na página 131, traz um exemplo desse programa (quando usado para analisar *Out of the box* (V-02)); pode-se perceber que alguns retângulos foram desenhados de forma a isolar alguns elementos da interface, sendo que uma letra identifica cada uma dessas partes. Vemos assim na parte superior uma janela que realiza o *playback* do vídeo a ser analisado

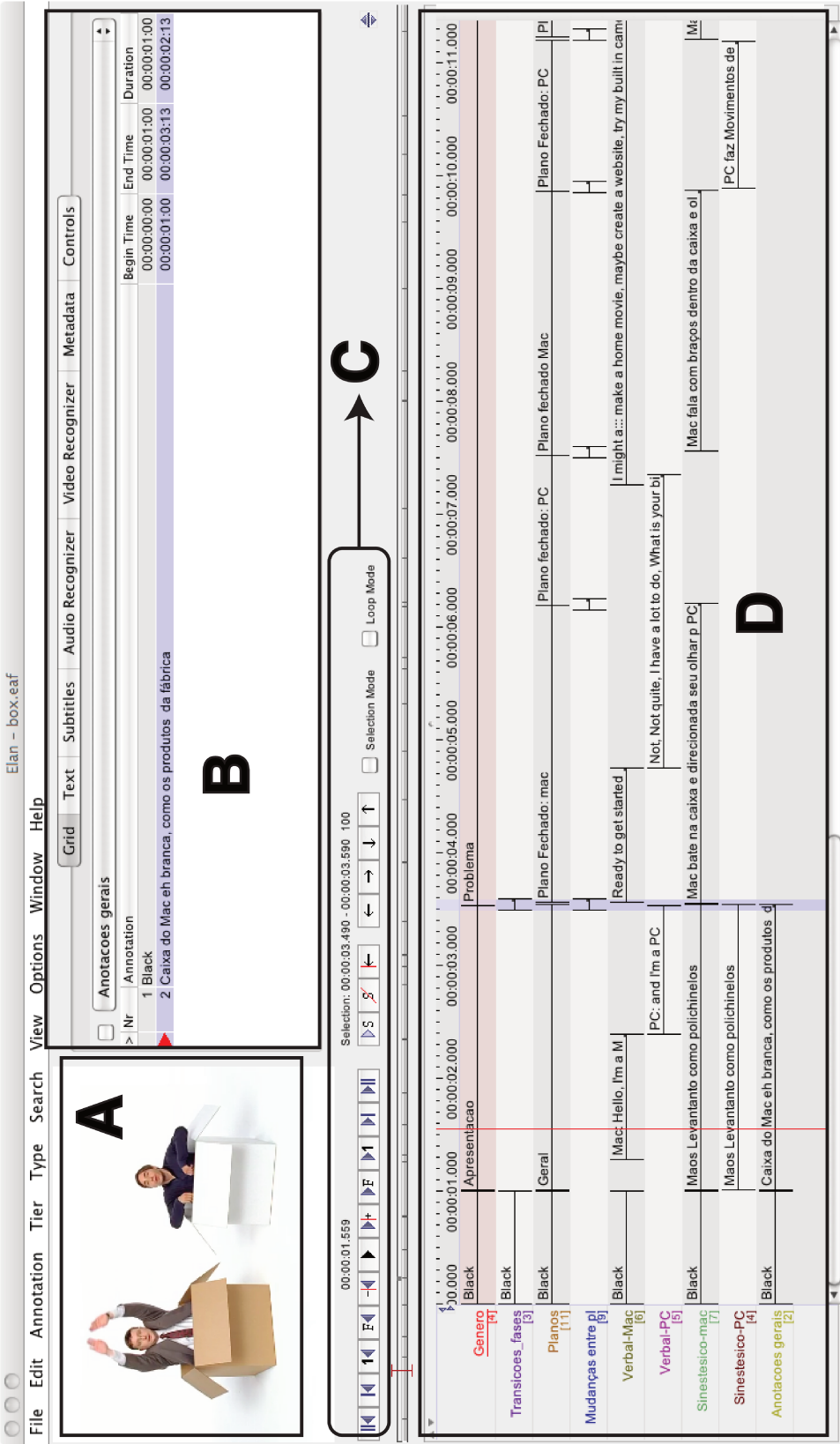


Figura 17: Janela principal do software ELAN 3.9

(parte A) e uma *timeline* na parte inferior, essa janela é destacável e permite a visualização do vídeo para que sua transcrição possa ser realizada. Na parte B da figura 17, observa-se um conjunto de notas que resumem os aspectos apontados na transcrição. Na parte C temos os controles técnicos de avanço e recuo, esses controles permitem o controle do vídeo quadro a quadro, ou em termos de milésimos de segundo. Finalmente na parte D da figura 17, temos as anotações dos modos de linguagem analisados. Uma das vantagens oferecidas por esse formato é que podemos observar cada um dos modos sobrepostos na linha do tempo, facilitando sua comparação. Evidencia-se, assim, quando uma ação em determinado modo de linguagem coincide com outros em termos de sua realização temporal, além de haver uma sincronização entre aquilo que está sendo analisado e sua instanciação no vídeo.

Trabalhos como o de Baldry e Thibault (2006) e o de Hellwig, Uytvanck e Hulsbosch (2008) são importantes por se centrarem em formas de anotação do vídeo. Todavia, eles não resolvem todas as questões a serem levadas em conta durante a transcrição multimodal, uma vez que é necessário sistematizar a forma como esses modos serão anotados. Apesar de estudos na área de Linguística Aplicada terem uma grande tradição na transcrição da linguagem oral, poucos trabalhos pensam a forma como a transcrição e a anotação de outros modos de linguagem podem ser realizadas. Norris (2004) é um exemplo de trabalho que caminha por essa seara por lançar sua atenção à sistematização dos diferentes modos, oferecendo ferramentas metodológicas que podem ajudar o analista nessa tarefa.

Assim, para Norris, um dos grandes desafios para os linguistas é observar as diferentes estruturas dos diversos modos de linguagem. Isso se dá porque cada um deles tem uma sintaxe própria, não tão bem delineada como a linguagem verbal. Adicione-se a isso o fato de cada um possuir uma diferente materialidade de instanciação, sendo que a fronteira entre alguns modos de linguagem é muito tênue. Isso porque a determinação do que seja um modo de linguagem, como vimos no capítulo 1, é um conceito definido apenas para fins analíticos, uma vez que eles são unidades heurísticas que se autocomplementam, trabalhando para construir a complexidade das interações (NORRIS, 2004, p. 3-4).

Em sua abordagem, os modos de linguagem são, inicialmente, separados na medida do possível, sendo que a transcrição se inicia pela linguagem verbal. Isso se dá porque, para a autora, a linguagem verbal possui grande de informatividade (NORRIS, 2004, p. 65-66). Norris tece uma série de estratégias relevantes para meu estudo, em especial no que diz respeito à transcrição de ações corporais e proximidade entre os interactantes, este último definido pela autora como proxêmica. A partir desse ponto, Norris (2004, p. 19-30) estabelece uma longa discussão sobre como analisar os diversos modos de linguagem com os quais está preocupada. Aqui discutirei apenas aqueles que são mais relevantes para este trabalho, entre esses elementos estão:

- Proxêmica (*proxemics*): Está relacionada à forma como os participantes em uma interação organizam seu espaço, ou seja, busca-se observar a distância que indivíduos tomam em relação a outros, bem como a objetos relevantes à interação. Esta unidade está condicionada e integrada às ações complexas, bem como ao ambiente no qual a interação ocorre. Seus *insights* são importantes uma vez que ela é socialmente motivada (2004, p. 19-20).
- Postura: Está relacionada à forma como os participantes posicionam seus corpos em uma dada interação, sua postura pode ser aberta ou fechada, orientando-se através dela. Dois parecem ser os elementos relevantes para observação em relação a ela: 1) a forma da posição do corpo 2) a direção postural. A primeira está relacionada à posição do corpo, a segunda, à postura que um indivíduo toma, sendo que o interesse do estudo da postura é a relação entre ela e as ações complexas.⁵⁰ O sistema de representação da postura também é aprendido através da socialização, podendo também ser diferentes em culturas diversas (2004, p. 24-25).

⁵⁰As ações complexas são definidas na seção 4.2, na página 138.

- Gestos: os gestos estão diretamente relacionados às questões culturais, sendo que dificilmente ocorrem de forma independente, sempre percebidos durante interações cotidianas. Norris categoriza os gestos de acordo com sua função:
 - Gestos icônicos: são gestos que possuem um componente pictórico, sua função é imitar aquilo que está sendo falado, dando à interação mais autenticidade.
 - Gestos Metafóricos: também possuem um componente pictórico, todavia eles representam uma ideia abstrata.
 - Gestos Dêiticos: apontar para um objeto ou pessoas do mundo físico, muitas vezes substituindo um dêitico verbal.
 - Gestos de marcação: relacionados com o marcar do tempo, um exemplo de sua manifestação é o maestro e a orquestra.
- Olhar: Essa unidade é referente a organização direção e intensidade do olhar, podendo ser uma sequencia estruturada e estando ou não subordinado à língua. Entre suas características está o fato de poder variar entre escolhas sistemáticas (que estão estruturadas de forma a construir um significado específico) ou assistemáticas (que não são estruturadas). Participantes em uma interação percebem e determinam a direção de um do olhar e muito do valor interacional do olhar advém de sua interpretação pelos interactantes. A estrutura do olhar depende de outras ações realizadas pelo falante da interação, sendo que essas ações não precisam ser necessariamente de caráter verbal (NORRIS, 2002, p. 28-29).

Cada um desses modos será transcrito de acordo com alguns critérios-chave, também definidos por Norris (2004, p. 66-72):

1. Transcrição Proxêmica e de Postura:

- (a) usa-se uma imagem do frame, mostrando a distância entre os participantes e dos objetos;

- (b) depois oferece-se um novo frame para cada mudança, marcando-se o tempo nos vídeos, o que implica em um novo *snapshot*, a cada novo movimento.

2. Transcrição de gestos:

- (a) oferecem-se *stills* do início do movimento e do seu ponto de realização;
- (b) oferecem-se *stills* separados das fases intermediárias e da fase de pós-realização, se for necessário.

3. Movimentos de cabeça:

- (a) inicia-se com a cabeça em posição de repouso;
- (b) quando necessário, as fases do movimento devem ser documentadas;
- (c) alguns movimentos dêiticos possuem movimentos pós-realização, e devem ser indicados em um frame separado.

4. Olhar

- (a) qualquer modificação no olhar deve ser documentada, trazendo especificações sobre sua trajetória e o ângulo assumido por ele.

Neste estudo, opto por uma forma de trabalho mista, que herda aspectos das três abordagens discutidas acima e é dividida em duas fases. Na primeira, realizei um levantamento em forma de tabela de alguns dos itens sugeridos por Baldry e Thibault (2006); na segunda criei um modelo próprio de análise utilizando o programa ELAN 3.9 (HELLWIG; UYTVANCK; HULSBOSCH, 2008) que visa a melhor traduzir minhas necessidades nesta tese. O trabalho de Norris (2004) serviu como um método ancilar de transcrição de modos relacionados ao corpo humano — ou seja, forneceu os critérios de identificação dos gestos, olhares e proxêmica —, e foi usada nos dois momentos da análise. Apesar do relativo grau de redundância entre os procedimentos realizados, eles são complementares: a primeira fase tem por objetivo

oferecer um mapeamento inicial do vídeo, possibilitando o levantamento das unidades de análise. Já o segundo, torna essa análise mais profunda, por demonstrar como esses modos se manifestam na linha do tempo e por registrar mais precisamente os modos estudados.

Inicia-se, então, a análise partindo de alguns dos procedimentos colocados por Baldry e Thibault (2006), que ajudaram a construir uma análise inicial em forma de tabela: 1) minutagem das cenas e ações realizadas; 2) divisão das sequências/cenas e 3) mapeamento e comparação dos elementos verbais e não verbais. No caso da minutagem, todos os vídeos foram estudados de forma a ser realizada uma marcação precisa e detalhada da ação realizada em cada um deles dentro da linha do tempo. Esse tipo de levantamento foi essencial para divisão das sequências/cenas, por meio do qual defini as unidades orgânicas, tipicamente dentro do mesmo tempo/espço, que caracterizam a ação realizada em um vídeo. Em outras palavras, ao se dividir um vídeo em cenas, ou sequências, está-se determinando unidades de ação que possuam algum significado. Essa divisão é importante porque no processo de composição do roteiro de um vídeo tanto o texto quanto o conteúdo visual são pensados em termos dessas unidades.

O mapeamento e comparação dos elementos verbais e não verbais consiste na construção de um mapa conceitual em forma de tabela em cujas colunas são dispostas as seguintes informações: 1) tempo em que a ação se desenvolve; 2) um *screenshot* com a imagem de início da unidade analisada; 3) descrição da ação; 4) descrição da ação sinestésica; 5) trilha sonora e 6) análise da Metafunção da linguagem representada pela ação (um exemplo de realização desse processo de transcrição manual pode ser visto no Anexo A, na página 265). O tempo de desenvolvimento é responsável por trazer anotações referentes ao *timecode* do vídeo, sendo seu objetivo propiciar referências para a análise a ser realizada. O *screenshot* traz uma referência visual do que está sendo analisado, ao passo que a descrição da ação visa à discussão dos conteúdos presentes nessas ações orgânicas do vídeo, definidas por Zetttl (2004) como unidades orgânicas, ou conjuntos de ações que se caracterizam por uma unidade de significado dentro do texto videográfico. A descrição da ação sinestésica objetiva ser uma

notação detalhada de elementos paralinguísticos, como gestos, expressões faciais e corporais, olhares e outras ações não verbais realizadas pelas personagens dos vídeos, servindo de insumo para a análise de seu significado. É nesse ponto que o trabalho de Norris (2004) traz sua contribuição para esta pesquisa. Aqui serão mapeadas a proxêmica, o olhar, os gestos e os movimentos de cabeça. Na trilha sonora, transcrevi as falas realizadas pelas personagens e qualquer realização sonora, como músicas e efeitos. Finalmente, o último item é relacionado à realização das Metafunções de Linguagem realizadas pelo texto videográfico e imagético.

Depois de compor esse mapa conceitual realizado em forma de tabela, como apoio para refinamento da análise, utilizei o software gratuito ELAN (HELLWIG; UYTVANCK; HULSBOSCH, 2008). Nele, todos os modos analisados nos mapeamentos foram transcritos para o formato *timeline*, com alguns acréscimos de informação, referentes à fase do texto vídeo-filmico e aos planos de gravação utilizados em cada frame dos vídeos. Como podemos ver na figura 17, na página 131, esse trabalho contou com a criação de novas categorias, as quais objetivaram aproximar os dados então levantados das reais particularidades desta pesquisa. Essas categorias são: 1) *timecode*; 2) fase do anúncio; 3) transição entre fases; 4) mudança entre planos; 5) modo verbal Mac; 6) modo verbal PC; 7) sinestésico Mac e 8) sinestésico PC. O *timecode* é automaticamente fornecido pelo programa e se sobrepõe a todas os modos dispostos na *timeline*, possibilitando a verificação da duração de cada um dos modos de análise levantados nessa fase. Na descrição das fases do anúncio, procurou-se dividir os anúncios em blocos informativos.⁵¹ Na transcrição entre fases, busquei levantar qual recurso foi utilizado para a transição entre os blocos informativos da campanha: por exemplo se houve alguma mudança de enquadramento, algum efeito visual etc., ao passo que a mudança de planos registra quais planos estão sendo utilizados em cada momento do vídeo. O modo verbal (PC e MAC) registra a fala de Mac e PC, associando-a ao momento em que os demais modos de linguagem ocorrem nos vídeos. Por fim a transcrição sinestésica (PC e MAC) traz

⁵¹Como já discutido na introdução deste trabalho uma discussão detalhada sobre a noção de gênero dentro da GSF, e na teoria multimodal, ocorre no capítulo 1.

informações sobre gestos e olhares, bem como proxêmica; cada um dos modos não verbais foi analisado de acordo com os critérios estabelecidos por Norris e discutidos anteriormente.

Entre as vantagens do uso desse software, como vemos na figura 17, na página 131 (parte D), está a comparação entre as anotações de modo por ele permitida; a função dos modos de linguagem na peça audiovisual estudada é facilmente observável, criando uma transcrição frame a frame daquilo que se busca pesquisar. Observa-se, assim, a sincronia (ou não sincronia) entre os modos de linguagem estudados. Entre as possíveis desvantagens desse sistema está a necessidade dos vídeos serem salvos em um dos diversos CODECs aceitos pelo ELAN, exigindo algum conhecimento em TI para sua instalação e funcionamento.

4.2 Constituição: elementos criadores de significado

Nesta seção discuto como o conceito de constituição⁵² pode ser utilizado na análise de textos multimodais, desenvolvendo uma reflexão sobre a forma como os constituintes dos diversos modos de linguagem podem ser analisados em conjunto. Acredito que, a exemplo dos textos ditos monomodais, textos multimodais também são constituídos por diversos *strata* (ver figura 18, na página 145), realizados por uma gramática própria.⁵³

A observação conjunta, como já sugerida por autores como Baldry e Thibault (2006), não é uma novidade para análise multimodal. Contudo, a maioria dos trabalhos (c.f. NORRIS, 2004; IEDEMA, 2001; BALDRY; THIBAUD, 2006) busca desenvolver elementos procedurais que não refletem diretamente sobre a integração dos modos de linguagem e suas gramáticas. Esse é justamente o ponto que esta seção busca discutir, observando como escolhas estão interligadas; minha hipótese é que, ao escolher construir um significado qualquer, autores de textos videográficos têm disponível para si uma complexa rede de significados e escolhas que perpassa todos os modos de linguagem. Isso, de certa forma, permitiria uma delimitação metodológica mais clara, fornecendo alguns critérios objetivos para análise de

⁵²Ver página 33 para discussão do conceito de constituição.

⁵³Gostaria de agradecer a Geoff Thompson e a Peter Fries pelas profícuas discussões sobre esse tema.

textos videográficos: sugiro que o significado seja o ponto de partida para a análise, pois será ele que permitirá um mapeamento mais claro de como essa rede de escolhas se tece. Em outras palavras, acredito que os níveis mais abstratos, como o do conteúdo, como o semântico, selecionam os modos de linguagem a serem utilizados; esses modos de linguagem, por sua vez, têm selecionado os constituintes que devem estar presentes nos vídeos. Os constituintes de cada modo de linguagem se relacionam, instanciando o significado e criando a rede de escolhas.

Uma questão importante a ser levada em conta nesse trabalho é que o vídeo é um modo de linguagem que pode abarcar uma série de outros. Se pensarmos no vídeo como um simples registro do cotidiano (como já discutido no capítulo 1, seção 1.1, na página 25), ao gravarmos uma simples interação familiar, por exemplo, temos uma peça audiovisual composta pela fala dos interactantes, a música ambiente (se houver), seu gestos, sua expressão facial, conversas paralelas registradas pelo microfone, elementos próprios da linguagem videográfica, iluminação e infinitas outras possibilidades. Agora, se o vídeo é pensado como peça artístico-profissional, como é o caso dos que compõem o corpus desta tese, os mesmos modos de linguagem estariam presentes, mas em uma situação na qual o planejamento prévio levaria a um produto final totalmente controlado. Em ambos os casos, o significado não reside apenas em um ou outro modo de linguagem, mas sim na somatória de todos eles; cada um contribui de forma única para construção do significado; não haverá dois vídeos idênticos, pois qualquer transformação em alguma variável de linguagem levará a um produto novo.

No caso específico do corpus de análise, dadas suas características — cenário minimalista, sem móveis, poucos atores na tela, sem discussões paralelas, trilha sonora simples⁵⁴—, são poucos os modos de linguagem empregados na construção dos significados, entre eles talvez seja importante destacar: 1) elementos de linguagem videográfica; 2) fala; 3) proxêmica e 4) gestos. No caso de 1), refiro-me ao cabedal de possibilidades combinatórias que envolvem elementos básicos de linguagem videográfica, já discutido na seção 3.2, na página 118, no

⁵⁴Ver capítulo 3 na página 109, seção 3.1 na página 110, para uma definição mais detalhada do corpus de pesquisa.

capítulo 3, que podem ser combinados em unidades discursivas. Essa combinação levaria à construção de um texto que se caracteriza por uma intenção de significar bem definida, como veremos a seguir; nesse caso o frame ou quadro, unidade objetiva mínima do vídeo, pode ser expandido em unidades subjetivas que constroem algo muito similar a uma proposição. A fala (2) é o modo de linguagem mais comumente estudado pela linguística e já possui metodologias de transcrição desenvolvidas. Já a proxêmica (3) e os gestos (4) estão relacionados com as escolhas significativas realizadas durante a interação, cada um deles possui diferentes características e possibilidades de escolha. Eles são importantes por não apenas agregar significado mas também por serem relevantes para o fluxo de informação, como veremos a seguir no capítulo 6, na página 213.

É importante observar que tanto os estudos que discutem a questão da significação das escolhas da linguagem videográfica, como é o caso de Iedema (2001), que analisa um documentário sobre as condições de trabalho em um hospital em Sydney, Austrália; como os trabalhos de Norris (2002; 2004; 2006) teorizam sobre a forma como os diferentes modos de linguagem presentes em nossas ações cotidianas, em termos de suas ações constituintes, são tímidas tentativas de estudar de forma conjunta a linguagem videográfica e os modos de linguagem corporais e comportamentais, como são a fala, os gestos e proxêmica. Entre eles talvez o trabalho mais profundo seja o de Baldry e Thibault (2006), já citado anteriormente, mas com uma ressalva: os autores oferecem uma extensa reflexão sobre um método de transcrição, mas com poucas incursões na forma como esses modos se interligam na criação do significado. Esta seção caminhará por esta seara, ao tentar relacioná-los e aplicar o construto desenvolvido no corpus de pesquisa deste trabalho. O caminho a ser percorrido será uma discussão dos trabalhos de Norris (2004) e Iedema (2001) para posterior exemplificação desse princípio por meio da análise dos vídeos que compõem o corpus.

Norris (2004) analisa como nossas interações cotidianas se organizam na criação de significados. Sua preocupação é entender como o potencial significativo de uma ação não linguística se constrói, voltando seus olhos para a forma como cada elemento de uma interação pode

contribuir para o seu todo significativo. Na opinião da autora, entre os modos de linguagem utilizados pelo homem, a língua escrita é de longe o mais estudado, além de ser considerado por muitos o que possui o maior grau de informatividade e expressão, fazendo com que outros modos sejam subordinados a ele. Ela coloca justamente o contrário dessa visão tradicional: a linguagem verbal pode, também, ter um papel de igualdade em relação a outros modos de linguagem (NORRIS, 2004, p. X). Isso se dá principalmente porque eles também são elementos agregadores de significado, e sendo igualmente relevantes para a interpretação da interação.

Uma vez que Norris está preocupada com o processo de interação, ela parte da ação como unidade mínima significativa, buscando observar quais são seus elementos típicos e constituintes. Uma ação é definida como uma unidade responsável pela confluência de diversos modos que realizam e mediam seus significados (NORRIS, 2004, p. 11-13). Para tanto, ela cria um modelo de análise hierárquico e de traços paradigmáticos, muitas vezes também dicotômico, que busca entender como essa constituição se dá em termos estruturais e discursivos. De uma maneira geral, essas escolhas são originárias da filiação teórico-epistemológica que Norris realiza: a análise multimodal proposta pela autora tem suas bases na GSF, uma abordagem altamente hierarquizante e que possui ecos teóricos advindos do modelo de descrição de Bloomfield (1933), o que justifica sua tendência à análise paradigmática.

Dentro do colocado por Norris, qualquer texto, oral, escrito, impresso, estático ou dinâmico pode ser considerado como uma ação. Dessa forma uma conversa ao telefone, um livro ou mesmo um vídeo seriam ações, pois eles representam um conjunto de modos de linguagem unificados na construção do significado. Norris coloca que existem ações compostas por elementos mais complexos que outros: falar ao telefone, em termos estruturais, é algo muito mais longo e que encerra mais ações do que acenar; um programa de televisão seria uma ação ainda mais intrincada, ao passo que uma conversa cotidiana se perde no tempo/espço. Para dar conta dessa diferença Norris subdivide o conceito de ação em diversos subtipos:

ações simples, ações complexas, ações congeladas, corporificadas e não corporificadas, todas definidas a seguir.

No caso das ações complexas (*Higher-Level Actions*), teríamos ações constituídas por uma série de outras ações menores, como seria o ato de pedir horas a um estranho: olharíamos para a pessoa, perguntaríamos as horas, essa pessoa olharia para o relógio, interpretaria seus signos e nos responderia. Nós agradeceríamos e sairíamos; se houvesse recusa, o estranho nos ignoraria. Obtendo ou não a resposta desejada, temos uma ação que se define pela confluência de ações menores, “pedir as horas” se constitui enquanto significado por ser um conjunto e não por ser um ato isolado em si. Na verdade, Norris coloca que as ações complexas são marcadas por um início e um fim, com um significado social definido em uma comunidade e pela constituição em atos mais simples. Já as ações simples (*Lower-level actions*) são responsáveis por compor uma ação maior, com uma função discursiva.

Segundo Norris, a maioria das ações que realizamos em nosso cotidiano, e a conversa é um exemplo disso, se perdem no espaço/tempo, não podendo ser recuperadas após sua instanciamento. Todavia, algumas dessas interações se dão dentro de um suporte físico que permite que sejam revisitadas, mesmo que não em tempo real; exemplos seriam o livro ou uma gravação musical. Em ambos os casos, ao lermos o livro ou ouvirmos a música, estamos retomando a informação em um tempo diferente de sua realização efetiva. A essas ações, Norris dá nome de Ações Congeladas (*Frozen Actions*) (NORRIS, 2004, p. 14).

Além disso, tais modos de linguagem variariam de acordo com a relação que possuem com o corpo humano, sendo classificáveis a partir de sua relação com ele. Nesse sentido, modos como a fala, olhar e gestos seriam corporificados (*embodied*),⁵⁵ ao passo que outros modos como trilha sonora, tipografia e layout seriam não corporificados ou impendentes (*disembodied*) (NORRIS, 2004, p. xi). A diferença está na forma como eles são instanciados: modos de linguagem corporificados dependem do corpo humano e normalmente não são congelados, à medida que os modos não corporificados são instanciados em suportes externos.

⁵⁵ Agradeço à Carolina Siqueira Muniz Ventura pela discussão terminológica.

Vídeos gravados — não apenas os que compõem este trabalho, mas todos de uma forma geral — se caracterizam por serem ações congeladas e complexas. No caso do vídeo, esse congelamento da ação pode ou não ser algo artificial a ele: eventos sociais simplesmente gravados, como é o caso de festas, casamentos e batizados, eternizam algo naturalmente efêmero, muitas vezes sem planejamento⁵⁶ ou preocupação estética; já uma reportagem, um documentário ou um comercial, *e.g.*, só existem graças a algum meio de gravação. Isso porque essas peças precisam do suporte para que sua existência se dê efetivamente: são produtos de edição, montagem e correções técnicas; ações que dependem de algum suporte, que pode ser desde uma fita até um disco rígido de computador ou uma memória flash. O vídeo, assim como a fotografia, representa ações que se dão no universo exterior a ele. Isso faz com que ele seja uma ação eminentemente complexa e comumente congelada, sendo composto não apenas pelos elementos de sua própria linguagem, mas assimilando os conteúdos de todos os modos de linguagem nele representados.

Iedema (2001) analisa um documentário sobre as condições de trabalho em um hospital em Sydney, Austrália, sendo seu objetivo entender como os processos de produção e edição de um filme podem contribuir para a criação de um estado discursivo. Para o autor, textos videográficos são socialmente reconhecidos e se caracterizam por não ocorrerem de forma natural: eles são representações da realidade, estabelecendo um tempo/espço próprio. Vídeo e filmes obedecem a lógica interna da sua construção e de sua veiculação, por meio de técnicas especializadas. De forma a entender como essas construções efetivamente ocorrem, Iedema (2001, p. 188-190) estabelece seis níveis de constituição de textos vídeo-fílmicos. Esses níveis seriam: 1) Frame; 2) *Shot*; 3) Cena; 4) Sequência; 5) Estágio genérico e 6) Obra completa. Definidos a seguir, a partir do menor nível de constituição em direção ao maior:

1. **Frame:** Além de ser a unidade mínima de formação da imagem em um vídeo ou filme, um frame (ou quadro) se define como o aspecto saliente de um *shot* (ou movimento de câmera). Caracteriza-se por ser uma imagem estática que, em conjunto com outras,

⁵⁶Ver discussão em torno da figura 2 na página 32, para uma explicação mais detalhada dessa diferença.

forma a imagem em movimento. Isso faz do frame (ou quadro) a unidade mínima composicional.

2. **Shot:** Um *shot* seria um movimento de câmera sem cortes ou transições. A modificação da posição da câmera (panorâmicas, zoom etc.) deve ser contínua sem que haja edição de qualquer tipo.
3. **Cena:** Uma cena pode ser definida como um conjunto de *shots* (movimentos), mas com uma característica básica: todos eles estão em um mesmo tempo/espaço. Isso determina uma unidade muito próxima ao que haveria na linguagem verbal em que esse conceito também é o definidor de um período.
4. **Sequência:** Em uma sequência, a câmera se movimenta através de diferentes tempos X espaços em um conjunto de fronteiras menos movediças, diferentemente de uma cena que teria fronteiras menos claras, as sequências seriam mais claramente determinadas.
5. **Estágio:** Para o autor cada gênero dentro da linguagem vídeo-filmica teria um conjunto específico de estágios. Ele exemplifica essa questão colocando que filmes narrativos costumam ser compostos por orientação, complicação, resolução e coda, ao passo que documentários e filmes de caráter documental (ou fatural) parecem se caracterizar por uma introdução, uma série de argumentos e fatos e uma conclusão. Cada um desses estágios seria reconhecido pelos espectadores e produtores desses filmes.
6. **Obra completa:** Dependendo dos níveis mais baixos de análise, o trabalho pode ser caracterizado como pertencente a um determinado gênero, aqui a noção utilizada é a de gêneros cinematográficos, que tem como principal determinante suas relações socioculturais, econômicas e simbólico-míticas (PUDOVKIN, 2007).⁵⁷

⁵⁷Aqui o autor não está discutindo os gêneros a partir de uma perspectiva discursiva, mas sim utilizando o termo no sentido dos tradicionais "Gêneros Cinematográficos". Que incluam uma divisão inicial entre ficção — que incluiria comédias, aventura, romance, drama, entre outros — e não ficção, responsável pelos diversos gêneros de documentário. Claro que gêneros híbridos nascem, existindo documentários ficcionais como é o caso de *Zelig* de Woody Allen.

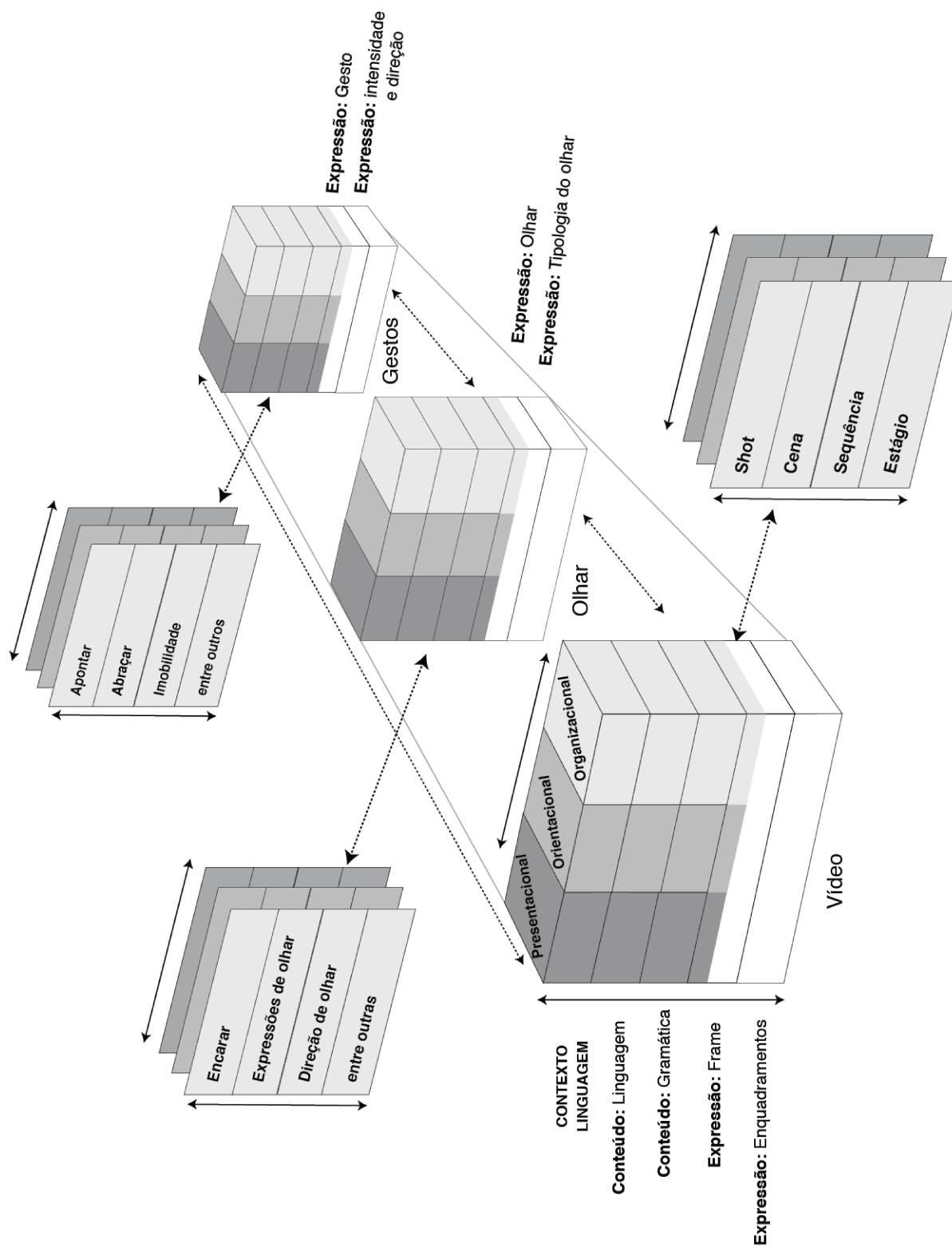


Figura 18: Sistemas de significado e linguagem videográfica (inspirado em Halliday; Matthiessen, 2004, capa)

Efetivamente, o que proponho aqui é que, durante a concepção desses produtos videográficos, os indivíduos se servem de dois conjuntos de modos de linguagem: tanto os que constituem as diferentes ações (fala, olhar, toque, entre outros), que aqui passo a chamar de componente atitudinal, como os relacionados à linguagem videográfica (enquadramentos, movimentos de câmera, entre outros), que aqui passo a chamar de componente estruturador (ou estrutural).

Acredito que esses dois componentes são igualmente importantes na criação dos significados presentes nos textos videográficos que estudo: os diferentes sistemas podem influenciar uns aos outros, gerando uma rede de escolhas bastante complexa. Um exemplo de como isso poderia ocorrer está na figura 18, na página precedente. Ali, à guisa de exemplo, temos a relação entre três modos de linguagem: cada um deles possui um próprio sistema de escolhas e, logicamente, de construção de significados, que são auto-influenciados no processo de significação. Cada um dos cubos centrais representa um modo de linguagem; nesse caso escolhi retratar o vídeo, os gestos e o olhar, dada sua importância especial para o corpus de minha pesquisa. As metafunções de linguagem são representadas pelos tons diversos, mas se mesclam nos níveis (ou *strata*) que estão mais relacionados à expressão, uma vez que todas as escolhas se realizam de forma simultânea. Os níveis (ou *strata*) mais abstratos, como conceito e linguagem, são comuns a todos os modos de linguagem — mesmo sabendo-se que a linguagem de cada um deles deve ser específica —, sendo que o processo de expressão acaba sendo particular e específico de cada um. Os retângulos que se abrem à direita e à esquerda representam as unidades no nível da expressão, com as formas articulatórias de cada um dos modos de linguagem. Naturalmente, perceber-se-á que o modo de linguagem videográfico tem unidades melhor articuladas e definidas que o olhar e os gestos; isso se dá porque o filme, o vídeo e a fotografia, diferentemente de outros modos de linguagem, já possuem uma longa tradição de análise, já tendo sido expostos a discussões sobre suas características estruturais desde o início do século XX. Além do mais, sua segmentação física em fotogramas, quadros

ou frames, já discutida na seção 3.2, na página 118, no capítulo 3, é propícia a uma análise semântico-estrutural.

A figura 18, página 145, nos mostra por meio de suas linhas longitudinais que cada um dos modos de linguagem influencia e é influenciado pelos demais, formando um sistema de significados coeso. A ideia por trás das linhas é demonstrar que cada vez que se realiza a escolha em um modo de linguagem, ela influenciará os demais fazendo com que as escolhas no mesmo sistema busquem instanciar o mesmo componente semântico, contribuindo de diferentes formas para sua existência. Em um vídeo, ao escolher, por exemplo, um significado de solidariedade, pode-se utilizar alguns modos de linguagem corporificados como olhar e toque, que seriam responsáveis pelo componente atitudinal, e um plano fechado, que levaria o espectador a estar mais próximo das feições e lhe possibilitaria perceber essa proximidade. Tais elementos constituintes da linguagem videográfica que contribuiriam para nossa percepção tanto dos demais modos de linguagem como da relação estabelecida por eles constituiriam o componente estrutural desse vídeo. Como esses dois componentes são motivados pela construção de um significado, que não é a simples soma de suas partes, alguma modificação em qualquer um deles levaria a uma transformação daquilo que é dito, mesmo que isso se dê em termos da percepção desse significado: um plano geral, por exemplo, não nos permitiria perceber a intimidade tão claramente, visto que apenas o enquadramento fechado também não garante esse significado, dado que esse plano pode ser utilizado com um número enorme de funções. É um arranjo entre os diversos modos de linguagem e seus constituintes que são responsáveis pela construção dos significados multimodais.

De certa forma, analisar textos multimodais é ter consciência de que eles são naturalmente híbridos, por ecoar significados e funções vindas de várias frentes. Ao construirmos um texto similar aos vídeos que compõem o corpus desta tese, temos que desenvolver métodos de compreensão de uma série de modos de linguagem que não são naturais para nossa formação de linguistas. O que proponho nesta tese é que todos eles devem ser observados simultaneamente, entendidos como um elemento significativo único. A opção pelo modelo

multimodal advém especificamente de seu parentesco com a GSF, permitindo que esse tipo de análise ocorra: todos os níveis em cada um dos modos de linguagem podem ser considerados em um patamar de igualdade.

Como já discutido, a figura 18, na página 145, tenta representar como esses sistemas se relacionam. Pode-se observar que um dos modos é representado por um retângulo tridimensional possuindo uma constituição e gramática próprias. Todos os modos de linguagem parecem estar em consonância nos níveis de conteúdo mais abstratos. Quando chegamos a níveis mais próximos da instanciación, cada um desses modos de linguagem parece tomar um rumo distinto, mas sempre orientado de forma a realizar o discurso. Em outras palavras, a relação direta entre a estrutura de cada modo de linguagem e o discurso fazem com que escolhas específicas sejam realizadas; tais escolhas são responsáveis por uma unicidade discursiva garantida pela conjugação de sua variedade de recurso, razão pela qual a análise de textos multimodais não pode prescindir de:

1. mapear cada um dos modos de linguagem presentes no texto em questão;
2. determinar quais constituintes de cada modo de linguagem estão interligados;
3. relacionar o significado de cada constituinte com o discurso do texto.

O mapeamento está relacionado com o levantamento de quais modos de linguagem constituem um texto multimodal. Muitas vezes nossa percepção de determinados textos multimodais é prejudicada pela naturalização de alguns modos de linguagem, o que pode ser resultado da sua constante presença em nosso contexto social. Um exemplo poderia ser os próprios vídeos que compõem o corpus desta tese: modos de linguagem como iluminação, interpretação dos atores, cortes de câmera e planos são tão comuns em todas as produções audiovisuais a que estamos expostos — além de comerciais, incluem-se os filmes, novelas, telejornais, *siticom*s e inúmeros outros — que muitas vezes nos esquecemos de que eles foram construídos de forma a instanciar algum significado. A determinação dos constituintes de cada modo de linguagem está relacionada à observação detalhada de quais elementos estão

constituindo os modos de linguagem no momento de instanciação de um dado significado; um mapeamento mais específico de forma a se observar quais ações estão atreladas à construção de determinado significado. Por fim, ao relacionarmos o significado de cada constituinte com o texto, estamos tentando isolar a forma como cada constituinte individual contribui no contexto de análise.

É importante ressaltar que alguns problemas podem surgir nesse processo de análise, principalmente no que tange à separação e ao isolamento dos constituintes no modelo de análise proposto. Efetivamente, ao observarmos a constituição do significado em nossas interações cotidianas, percebemos que elas não dividem os significados como o faz uma proposta analítica qualquer, as ações no mundo formam um amálgama de ações conjuntas. Separar suas ações constitutivas é apenas uma forma de possibilitar a observação mais detalhada de cada modo de linguagem de forma a entender como se dá a criação de uma gramática local. Minha proposta aqui é ajudar em uma reflexão a respeito dos possíveis parâmetros para estratégias de letramentos múltiplos. Textos multimodais não podem prescindir desse tipo de realidade, sua construção eminentemente multimodal faz com que a expressão dependa do conhecimento de diversos modos de linguagem, além de uma constante reflexão sobre suas possibilidades expressivas. Meu objetivo nesta tese é tentar transpor parte das dificuldades em construir um modelo como esse, partindo de uma análise em nível textual a fim de explorar seus significados textuais. Como veremos a seguir nos capítulos 5 e 6, essa proposta se tornou bastante profícua para o levantamento de dados e compreensão de como esses textos são construídos.

4.3 Construção do significado: operando vários modos de linguagem

Nesta seção, discutirei a construção do significado nos vídeos que compõem o corpus, partindo do pressuposto de que cada modo de linguagem contribui de forma própria e conjunta

para a construção do significado nos vídeos de "Get a Mac". Os diversos modos de linguagem passariam por escolhas dentro das possibilidades de sua gramática, de forma a constituir o significado a ser veiculado; esses elementos constituintes parecem ser recombinaados em uma gramática do vídeo, que forma proposições multimodais capazes de significar. Isso teria resultado em diferentes níveis (ou *strata*): o discurso a ser veiculado ao vídeo cria necessidades de design que levam a uma produção que instancie, por meio dessas proposições, os significados pretendidos. Não se pode esquecer de que o vídeo é um modo de linguagem socialmente reconhecido e, logo, está sujeito ao reconhecimento do seu uso pelo grupo: certos usos podem não estar entre os aceitáveis por uma determinada comunidade, o que poderia levar a uma não compreensão da mensagem, ou à refutação de um determinado texto vídeo-fílmico.

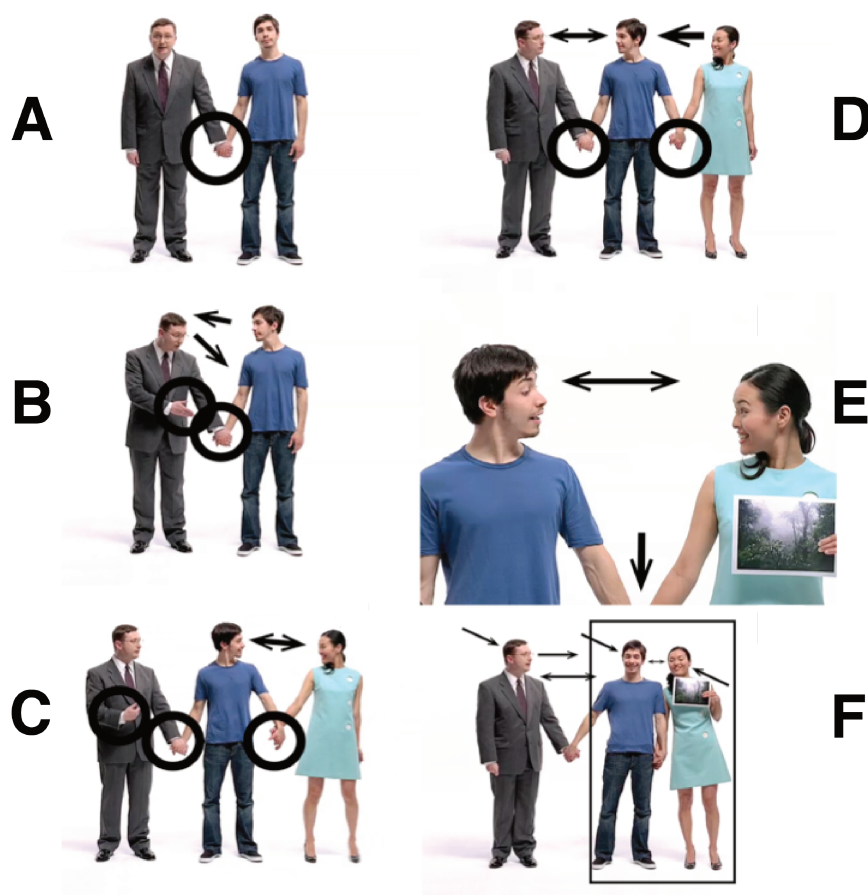


Figura 19: Elementos atitudinal e composicional em *Networking* (V-04)

Um exemplo de como esse tipo de relação ocorre no corpus pode ser observado em *Networking* (V-04) — que trata da relação entre as plataformas e seus periféricos:⁵⁸ Mac interage com PC e uma Câmera Digital, vinda do Japão e representada por uma mulher de traços orientais —, como podemos observar na figura 19, na página oposta. No frame A (superior esquerda), temos Mac (de jeans e camiseta) e PC (de costume), durante a apresentação das personagens; ali se enfatiza o fato deles estarem de mãos dadas, simbolizando uma conexão em rede. O plano é geral, totalmente aberto, caso a objetiva estivesse em um plano mais fechado em algum detalhe essa noção do todo se perderia. O mesmo enquadramento é mantido nos frames B, C, D; cada um com um objetivo próprio: em B a ênfase é dada aos olhares das personagens: Mac direciona seu olhar a PC, mostrando que ele é quem está com o turno e fonte da informação — ver análise do fluxo de informações no capítulo 6, na página 213 —, ao mesmo tempo PC dirige seu olhar e seus gestos para as mãos dadas e seu significado. No Frame C, há a entrada da câmera digital, que acontece da direita para esquerda; ela não apenas amplia a rede, como atrai o olhar de ambos, sendo que apenas Mac recebe uma resposta. Na verdade, a conexão da câmera se dá apenas com Mac, ela não se conecta a PC a não ser por seu intermédio, como podemos ver no frame D. Apesar de PC e ela se observarem com uma certa constância, é Mac quem traduz suas falas para PC. O frame E parece marcar o início de uma nova sequência, na qual Mac e a Câmera Digital passam a interagir, ignorando PC, falando uma língua que ele não consegue entender. Nesse frame, temos o corte para um plano médio-fechado, que parece enfatizar o olhar de cumplicidade recíproca entre as personagens; neste mesmo frame a Câmera Digital oferece uma fotografia para Mac. Ali o plano médio-fechado preenche o quadro apenas com uma fração da ação, cabendo ao espectador preencher um vazio: ele deve imaginar que a ação que não está presente no quadro seja a mesma que ele vira em plano geral.

Um resultado é a posição que as personagens se encontram no frame F, no qual uma análise proxêmica (figura 19, página 150) demonstra que Mac e a Câmera Digital formam um bloco de

⁵⁸Em caso de dúvida, gostaria de direcionar o leitor para o quadro 1, na página 112, e para seção 3.1.2, na página 114, responsáveis por trazer sinopses dos vídeos que compõem o corpus de estudo.

solidariedade, eles estão muito mais próximos com os braços entrelaçados. PC aparece mais distante e excluído da interação, o que parece ser ainda mais enfatizado pelos olhares e sorrisos que são ali estampados. Isso pode refletir um discurso que enfatiza a suposta incapacidade de outras plataformas em interagir de forma fácil e direta com esses equipamentos. Esse componente atitudinal não é aleatório, busca refletir a construção de significados discursivos específicos: esse plano seria o único capaz de mostrar esses componentes atitudinais em sua totalidade.

Quadro 2 Relação entre os elementos atitudinais e composicionais em *Networking* (V-04)

Significado	Componente	
	Atitudinal	Estrutural
Rede	Mãos dadas	Plano geral
Não cumplicidade	Olhar para o público	Plano geral
Cumplicidade	Olhar entre personagens	Plano médio-fechado
Cumplicidade e Conexão	Mãos dadas,	Plano Geral
	olhares e risadas	Plano médio-fechado
Desconexão e não cumplicidade	Distância	Plano Geral

No quadro 2, página 152, há uma correlação entre os dois tipos de componentes nos frames analisados de *Networking* (V-04). Como pode-se observar, os significados condicionam os componentes atitudinais e estruturais, criando alguns padrões de instanciação. As mãos dadas, simbolizando a rede, são responsáveis por estabelecer a dinâmica de interpretação do vídeo, inclusive nos ajuda a interpretar a noção de exclusão sentida por PC. Esses resultados são importantes porque ajudam a criar um parâmetro de interpretação para constituição dos textos videográficos, os resultados parecem mostrar que o nível semântico (conteúdo) e os modos de linguagem presentes nos vídeos estão integrados, formando proposições visuais que funcionam como instanciadores de significado dentro do vídeo. Assim, cada vez que um plano geral é utilizado em *Networking* (V-04), entendemos que o objetivo é mostrar qual o status das relações entre Mac, PC e a Câmera Digital, ao passo que os planos mais fechados ganham tônica de cumplicidade (quando aplicados a Mac e à Câmera Digital) e de solidão, quando aplicados a PC. Ao mesmo tempo, essa relação de intimidade é também significada

por escolhas dentro de modos corporificados: os gestos e olhares traduzem para algo tangível os conceitos de compatibilidade computacional, linguagem de programação, rede e vias de troca de dados. Todos eles essenciais para compreensão da mensagem principal do vídeo: que Mac é inteiramente compatível com dispositivos de imagem, ao passo que PC enfrenta graves problemas de linguagem e de conexão.

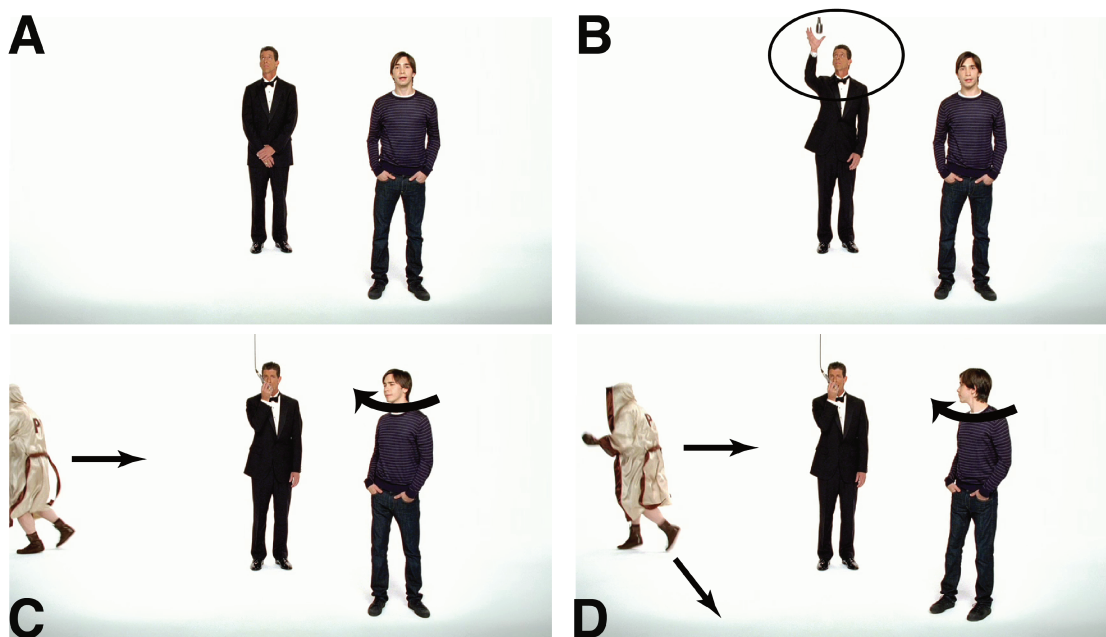


Figura 20: Componentes atitudinal e composicional em *Boxer* (V-09)

Outro exemplo seria o vídeo *Boxer* (V-09) — que tem por objetivo argumentar que os fabricantes da plataforma PC encaram sua inferioridade em relação à plataforma da Apple como uma luta a ser vencida a qualquer custo.⁵⁹ Apesar de usar os mesmos modos de linguagem de *Networking* (V-04), tem um arranjo bastante diferente (ver figura 20, na página 153). O frame A mostra Mac à direita, mas sem PC ao seu lado, na região central da tela há um homem magro e alto vestido como um apresentador de lutas. No frame B, ele olha para cima, seu olhar nos ajuda a perceber que algo virá daquela direção, temos um microfone de locução descendo, o apresentador levanta a mão e começa a anunciar a entrada de PC no palco, à guisa do que é feito em uma luta de boxe, como vemos no Frame C. Simultaneamente,

⁵⁹Ver seção 3.1.2, na página 114.

PC entra pelo lado esquerdo enquanto Mac vira seu rosto, compondo um indicador de qual é a origem da informação. Nesse anúncio o apresentador troca as informações tradicionalmente indicadas em anúncios de boxe, o peso de PC é substituído pelo tamanho de seu disco duro (... *80 gigs*...) e a cidade de origem pelo uso comum (... *from pops*...), como mostra a transcrição 5:

- (5) **Apresentador:** In the left corner, coming in, at 80 gigs, him, from pops' desktop everywhere, P(3)::::::::::C! *Boxer*
(V-09)

Quadro 3 Relação entre os elementos atitudinais e composicionais em *Boxer* (V-09)

Significado	Componente	
	Atitudinal	Estrutural
Luta de boxe	Gestos de aquecimento	Plano Geral
	Saltitar de lutador	
	Presença do apresentador	
	Fala do apresentador	
	Roupas de PC	
Direção da informação	Olhares	
Indiferença de Mac	Mão no bolso e postura não modificada	

Nessa sequência de *Boxer* (V-09) (ver figura 20, página 153) há apenas o plano geral, sem nenhuma variação, que parece ser uma consequência de uma necessidade na construção do significado: não seria possível mostrar a entrada ao estilo boxeador de PC se o plano estivesse mais fechado. A presença do locutor (na posição central em todos os frames) também serve a esse propósito de representação, uma vez que essa é a forma típica de anúncio dos competidores em uma luta. As roupas de PC não são o usual terno e gravata; as modificações em PC não estão relacionadas apenas a sua roupa, elas também estão em seus gestos, que agora lembram o aquecimento de um atleta em sua entrada para luta, e em seu olhar, agora baixo e escondido por um gorro sem encarar o oponente, postura comum na entrada de lutadores no ringue, ver frames C e D. PC se aproxima andando em linha reta e posteriormente na diagonal do vídeo, de forma a se abeirar de Mac e perfilar-se diante ele.



Figura 21: Frame em *Security* (V-06)

O quadro 3, na página oposta, traz uma relação entre os componentes estruturais e comportamentais no segmento analisado de *Boxer* (V-09). Ali se pode observar que não há muitas variações de planos, esses momentos iniciais do vídeo podem ser classificados em apenas uma sequência, que vai desde a apresentação de Mac até a chegada de PC, que virá concomitantemente ao término da apresentação. Esse significado se subdivide em diversos modos de linguagem constituintes, como saltitar, gestos, roupas, além do modo verbal, todos parecendo contribuir igualmente para a criação de uma ambientação de luta de boxe, que servirá como metáfora da guerra de PC contra Mac. A indiferença de Mac parece se aplicar aos esforços de PC para alcançá-lo: as mãos no bolso se qualificam como um gesto típico de desprezo; esse mesmo desprezo também é retratado pela forma fria com que ele não muda sua postura corporal ao perceber PC ali presente.

Outro exemplo bastante relevante é *Security* (V-06), no qual é discutida a não-funcionalidade dos sistemas firewall da plataforma PC por meio do contraste com a simplicidade do sistema da Apple. Como podemos observar na figura 21, na página corrente, o Firewall é representado por um homem utilizando costume preto e óculos escuros, com uma

postura típica que lembra agentes federais americanos. Ele fica entre Mac e PC, filtra as mensagens que são enviadas de um lado para o outro, pedindo que PC autorize toda a comunicação que entra ou sai de seu computador. Como se pode observar (figura 21, na página anterior), Mac dirige seu olhar para Firewall por saber que é dele a origem da mensagem, este mesmo Firewall, por impedir a comunicação direta entre PC e Mac, parece motivar PC a não olhar para seu real interlocutor, ficando cabisbaixo.

Toda a comunicação é necessariamente seguida de uma frase de autorização de PC para que Firewall possa permitir o fluxo de informações, como mostra a transcrição 6. Esse tipo de situação faz com que a interação ocorra de forma ritmada e tediosa. Tédio que também é refletido no mesmo olhar de PC.

- (6) **Mac:** Hello I'm a Mac...
- Firewall:** Mac is greeting salutation cancel or allow?
- PC:** Allow.
- PC:** And I'm a PC...

Firewall, por sua vez, parece não exprimir nenhum tipo de emoção ou mesmo se importar com os fatos que estão acontecendo. Ele se mantém imóvel, à exceção de sua cabeça, que parece se mover no sentido da origem e do destino da informação. Ao se observar novamente a figura 21, na página precedente, notar-se-á que Firewall está mais próximo de PC do que de Mac, indicando a quem ele presta seus serviços. Novamente, a exemplo dos vídeos anteriores, temos um sistema de construção de significados que parecem depender da associação de componentes em diversos modos de linguagem.

Como podemos observar no quadro 4, na página oposta, o fato de o vídeo estar sendo instanciado apenas em um plano geral, não o faz menos complexo: há uma série de significados ocorrendo simultaneamente, todos eles em associação com algum modo de linguagem corporificado. Em alguns casos, como o olhar de PC, mais de um significado: tanto a depressão como a incapacidade de comunicação são expressas por ele. Certos modos, como é o caso da proxêmica podem ser interpretados dependendo do ponto de referência adotado.

Quadro 4 Relação entre os elementos atitudinais e composicionais em *Security* (V-06)

Significado	Componente		Participante Realizador
	Atitudinal	Estrutural	
Fonte de informação	Olhar em direção à firewall	Plano Geral	Mac
	Apontar para firewall		
	Olhar em direção de quem fala		
Destino da informação	Olhar em direção a quem fala		Firewall
Função fiscalizatória (agente federal)	Roupas e óculos		
Prestação de serviço para PC	Proximidade de PC		
Autorização de entrada e saída de informação	Fala com escolha específica		PC
Chateação, incapacidade e tédio	Não olhar para interlocutores		
	Olhar para baixo		
Tédio e dificuldade de comunicação	Fala ritmada pela intermediação		Todos

A construção do significado nesses vídeos parece ser, logo, resultado de uma complexa rede de escolhas que interrelaciona modos de linguagem com constituintes bastante diversos, criando algo que parece ser maior do que a soma de cada gesto, olhar ou plano em especial. Um elemento importante a ser observado é que todos parecem igualmente importantes nesse processo, se modificarmos a roupa de PC em *Boxer* (V-09), por exemplo, o significado da luta de boxe contra Mac seria difícil de ser estabelecido. Ao mesmo tempo, os planos mais fechados que eliminam PC do quadro e unem Mac e a Câmera digital em *Networking* (V-04) (ver figura19, na página 150) ajudam a inferir um relacionamento mais solidário entre essas duas personagens, excluindo PC.

É claro que cada tipo de produção audiovisual deve ter uma linguagem própria, mas os componentes estruturais são os mesmos, o número de constituintes é limitado, não se criam novos planos, movimentos de câmera ou enquadramentos a todo o momento, a grande maioria dos diretores e fotógrafos tende a usar um número limitado deles,⁶⁰ disponíveis no sistema de significados que a linguagem videográfica representa. A interligação entre os diferentes modos

⁶⁰Ver seção 3.2, na página 118, no capítulo 3, para uma discussão desses planos.

de linguagem se caracteriza por ser uma espécie de catalogação de quais recursos que são empregados por cada modo de linguagem na construção de um significado. Se observarmos novamente a discussão realizada sobre *Networking* (V-04) (ver figura 19, na página 150, e quadro 2, na página 152), notaremos que diversos significados são construídos pela ação de mais de um modo de linguagem simultaneamente. Isso seria o caso da intimidade entre Mac e a Câmera Digital: ali há o olhar, a fala e o plano médio-fechado, todos contribuindo para a construção de um significado único. Já a relação entre os constituintes de cada modo de linguagem e o significado do texto pode ser definido como uma ponte entre aquele significado, em um nível mais concreto, e nível mais abstrato do texto. Se retomarmos *Networking* (V-04) e *Boxer* (V-09), (para o último ver figura 20, na página 153, e quadro 3, na página 154) observaremos que, no primeiro, a intimidade entre Mac e a Câmera Digital não parece estar lá à toa, ela serve ao propósito geral do comercial, que é mostrar a incapacidade de PC em se comunicar com o novo hardware; o mesmo pode se dizer a respeito do segundo, no qual a roupa de boxeador de PC representa sua postura de encarar as vantagens de Mac como uma guerra a ser vencida. No caso de *Security* (V-06), (ver figura 21, na página 155 e transcrição 4, na página precedente), temos uma constituição de modos que se caracteriza por uma grande variedade de significados simultaneamente. Ele estabelece uma relação de superioridade por parte do sistema Mac, que não sofre os problemas de um Firewall tão restritivo como o portado por PC.

4.4 Algoritmo final de análise

A função do algoritmo de análise é permitir que os mesmos procedimentos de análise sejam realizados em todos os vídeos de forma semelhante. Logo, a partir do discutido e exposto acima, o seguinte algoritmo de análise foi desenvolvido:

1. Acesso ao sítio oficial da Apple Inc.

- (a) Acesso a subpágina referente à campanha estudada.

2. Verificação de novos vídeos para download.
 - (a) Se houver novos vídeos, efetuar o download.
3. Nomear o vídeo e salvá-lo em diretório específico.
4. Assistir ao(s) vídeo(s) no player do computador.
 - (a) Escolher o(s) vídeo(s) para análise.
5. Assistir ao vídeo escolhido transcrevendo-o em língua original.
 - (a) Assistir pelo menos duas vezes para conferência.
 - (b) Assistir ao vídeo pelo menos duas vezes para checar a transcrição.
6. Assistir ao vídeo escolhido, transcrevendo-o no sistema de tabelas.
 - (a) Assistir pelo menos uma vez para o preenchimento de cada coluna.
 - (b) Assistir uma vez para a conferência geral.
7. Abrir o vídeo dentro do software ELAN.
 - (a) Realizar a transcrição dos modos um por vez.
 - i. Assistir ao vídeo pelo menos uma vez para a transcrição de cada modo.
 - ii. Reassistir pelo menos 1 vez para verificar a transcrição (1 verificação separada para cada modo de linguagem).
8. Comparar resultado de cada vídeo da campanha à procura de elementos comuns.

4.5 Resultados parciais

De maneira geral, este trabalho procurou compreender a integração entre os diversos modos de linguagem na construção de sentidos nos textos audiovisuais analisados. Como pudemos observar, isso se dá principalmente devido ao número de escolhas complementares nos diversos modos de linguagem, cada uma sendo responsável por uma parte dessa significação. De forma a entender qual a contribuição dos diferentes modos de linguagem na construção de sentidos, desenvolvi no capítulo 4 alguns critérios metodológicos e princípios de análise para textos multimodais que permitissem realizar tal objetivo. A primeira reflexão (ver seção 4.1, na página 129) estava relacionada aos procedimentos sistemáticos de transcrição e anotação dos diferentes modos de linguagem. Nesse momento, centrei o desenvolvimento de uma abordagem que utiliza tanto um processo de transcrição e anotação totalmente manual como uma ferramenta computacional de análise. A sugestão de um procedimento híbrido se justifica pela necessidade de dar conta de forma sistemática de todos os modos de linguagem que se combinam de forma a constituir o significado que interpretamos ao assistir aos vídeos que fazem parte do corpus desta tese. O levantamento manual é baseado no estudo de Baldry e Thibault (2006) e tem por função estruturar as impressões iniciais que temos ao tomar contato com um texto multimodal. Entretanto, a transcrição baseada na construção de uma tabela comparativa, como se dá nesse levantamento manual, parece não dar conta de um importante fator da análise: a simultaneidade entre as diferentes ações (NORRIS, 2004) dos diversos modos de linguagem. Por conta disso, discutiu-se a aplicação de outra abordagem que parece complementar: o software ELAN 3.9 (HELLWIG; UYTVANCK; HULSBOSCH, 2008), ver figura 17, na página 131, supre a deficiência citada, além de aprofundar a transcrição à medida que permite a recuperação das anotações ao mesmo tempo que assistimos ao vídeo. Aqui também foram discutidos procedimentos de transcrição de gestos, postura e proxêmica. Esses critérios já presentes na literatura (NORRIS, 2004) são importantes por oferecer critérios mais objetivos para análise de modos de linguagem relacionados às ações do corpo humano.

Seguiu-se, então, uma reflexão sobre como o processo de constituição pode ser traduzido para textos multimodais (ver seção 4.2, na página 138). Meu objetivo foi observar como a constituição de textos multimodais é naturalmente múltipla: diferentes modos de linguagem com gramáticas bastante específicas se combinam de forma a criar um significado que parece ser diferente da simples somatória de todos eles. No caso específico do vídeo, articulei um conceito de constituição que leva em conta dois elementos básicos, no primeiro estariam elementos próprios da linguagem videográfica, como enquadramentos e movimentos de câmera, e, no segundo, os modos de linguagem instanciados pelos atores presentes no vídeo. O primeiro foi nomeado como componente estrutural — uma vez que é o processo de edição e gravação do vídeo que dá sua estrutura dimensional no tempo e no espaço — e o segundo de componente atitudinal — pois é dependente de atitudes humanas para ser instanciado. Defendeu-se, então, a ideia de que esses componentes são combinados de forma a constituir o significado dos vídeos analisados nesta tese.

Na seção 4.3, na página 149, realizei uma aplicação do princípio que desenvolvi em 4.2, analisando como os componentes atitudinais e estruturais se combinam em dois dos vídeos estudados nesta tese. Como resultado pude observar uma grande variação entre os vídeos, apesar de serem pertencentes a mesma campanha, o que parece ser fruto do tipo de significado instanciado em cada um deles; é esse significado que parece ser responsável pela seleção e combinação dos constituintes de cada modo de linguagem nos vídeos estudados.

Esses resultados são importantes por alguns motivos. Em primeiro lugar, eles corroboram as reflexões teóricas observadas nos capítulos anteriores; em especial a conjunção de diversos modos de linguagem no processo de constituição de significados. Diferentemente de textos ditos monomodais, esses resultados parecem mostrar que uma abordagem multimodal leva a uma percepção do texto como algo que se constitui em uma múltipla articulação. Isso porque o significado é dependente de uma série de signos que são articulados em conjunto; o significado não é apenas solitariamente verbal, ele é o resultado do agrupamento de signos advindos dos mais diversos modos de linguagem. Esse processo leva a uma relação significa-

tiva que é, necessariamente, complexa e baseada em uma infinidade de recursos multimodais que estão disponíveis para o emissor. De fato, a análise realizada no capítulo 6 mostra exatamente isso, é necessário que haja um constante levantamento dos signos constituintes na construção de cada um dos significados presentes em um texto multimodal. Cada um desses elementos parece ser responsável pela realização de uma função específica na instanciación, como mostraram os vídeos analisados.

Esse construto funcional é importante para a análise que se segue nos capítulos 6 e 5, uma vez que ambos discutem instanciaciones dependentes desse processo de articulação entre escolhas realizadas simultaneamente em diferentes modos de linguagem. É correto, por conseguinte, afirmar que os capítulos que se seguiram foram aplicações desse método em duas esferas: uma local, que se caracteriza pela compreensão da tecitura textual dos vídeos, e outra mais relacionada ao nível discursivo e da adaptação desse discurso, e desses textos, a outras realidades culturais.

Capítulo 5

O vídeo enquanto sistema de construção de significados: explorando modos de linguagem

One outcome for research into learning with such multimodal resources is the need for methods and theories that attend to a range of modes of communication beyond language.

C. Jewitt

Introdução

Este capítulo tem por objetivo discutir os vídeos que compõem o corpus de estudo como sistema de constituição de significados. Aqui, procurarei observar quais são os constituintes básicos das ações nele realizadas, de forma a refletir sobre o seu padrão de escolhas desde *strata* mais abstratos, como discurso e a produção. De forma a cumprir esses objetivos, esse capítulo se divide nas seguintes seções: 5.1, Constituição do discurso: tradução e metáfora; 5.2, Organização e padrões em "Get a Mac"; 5.3, Escolhas e manutenção do significado em diferentes ambientes de produção e veiculação e 5.4, Resultados parciais.

Na primeira, amplio a discussão sobre os constituintes iniciada na seção 4.3 na página 149, refletindo sobre os processos de tradução e de metáfora assinalados por Kress (2010), utilizados na construção de uma rede de significados com um claro objetivo discursivo: facilitar o processo de compreensão das plataformas computacionais pelo público-alvo da campanha "Get a Mac". Já na segunda, tenho por objetivo discutir quais são as estratégias de organização visual e discursiva desses comerciais; observando também como eles estabelecem um diálogo intertextual com obras cinematográficas, além de outras campanhas da Apple, numa clara estratégia de manutenção de um discurso tecnológico e da identidade da campanha.

Em 5.3, Escolhas e manutenção do significado em diferentes ambientes de produção e veiculação, na página 199, comparo a campanha estadunidense, foco desta tese, com suas variantes veiculadas no Reino Unido e no Japão, de forma a observar até que ponto elementos de produção local modificam o discurso e os padrões até então estabelecidos. Nesse ponto, observo que os elementos constituintes gerais da campanha são mantidos, havendo algumas modificações nos modos de linguagem corporificados, na escolha dos atores e em alguns temas gerais dos comerciais. Isso pode ser o resultado de uma adaptação necessária para a produção da campanha nesses novos contextos que não o estadunidense.

Seguem-se, então, os Resultados parciais, na página 209.

5.1 Constituição do discurso: tradução e metáfora

Como já discutido na seção 1.2.2.1, na página 57, o discurso dos comerciais da campanha "Get a Mac", analisada nesta tese, parece ter por objetivo criar um espaço de significação no qual o consumidor seja motivado a comprar graças a uma sensação de exclusividade que a compra de produtos Apple talvez transmita. Como colado por Kahney (2004), ver página 60 para maiores detalhes, a Apple cria uma relação ideológica com seus consumidores; ao investir em design parece construir um discurso relacionado à quebra dos padrões comerciais da indústria de TI. No caso dos vídeos da campanha que estudo, esse tipo de discurso parece

se manifestar por meio de processos de tradução⁶¹ (KRESS, 2010) que parecem se manifestar de duas formas: a primeira estaria em uma comparação entre o universo corporativo vs. criativo e, o segundo, na facilitação de conceitos computacionais. Em ambos os casos, há a constituição de metáforas que possibilitam a criação de um discurso mais compreensível ao leitor e dele mais próximo.



Figura 22: Comparação roupas-Mac x PC em alguns vídeos do corpus

⁶¹Aqui utilizo a definição de tradução de acordo com o colocado por Kress (2010, p. 124) — já citado na seção 1.2.1.2, na página 49 —, para quem a tradução é uma ação pela qual o significado é transportado de um contexto ou de um modo de linguagem para outro. Similaridades são vistas com o conceito de ressemiotização, já introduzido por Iedema (2003), que também versa sobre a representação entre modos de linguagem. Entretanto, como também já citado, Kress (2010) cria uma subcategoria, a transdução, termo que se aplica exclusivamente quando a rearticulação de significados envolver mais de um modo de linguagem.

O que observaremos na análise dos dados a seguir é o fato da forma como essas metáforas fogem do conceito de metáfora tradicional, que está, normalmente, relacionado a uma figura de pensamento, sendo definida como uma comparação elíptica (TAVARES, 1974, p. 368). Nesse conceito tradicional, o significado metafórico de uma palavra ou expressão se contrapõe a seu significado não metafórico, ou literal, que é a base de sua interpretação. Já Lakoff e Johnson (1980)⁶² dão um importante passo para a análise da metáfora ao determinarem que ela vai muito além do simples embelezamento do discurso, ela é parte integrante de nossa construção cognitiva da cultura, da linguagem e da própria sociedade humana. Para os autores, nosso sistema de expressão simbólica, por mais simples e cotidiano que seja, é recheado de metáforas (LAKOFF; JOHNSON, 1980, p. 3-4), que, de uma maneira geral, conceituam cada domínio de nossa experiência a partir de outra. Essa teoria é não universalizante por pregar que a criação das metáforas será resultante de diferentes mapeamentos — definidos como relações intradomínios —, baseados em elementos ideológicos e culturais de cada grupo social. Isso contrasta com as abordagens tradicionais nas quais a metáfora é o resultado da capacidade individual de criação de ações de linguagens que primam pelo belo e pelo inusitado. Para Lakoff e Johnson (1980), a metáfora é um construto do coletivo legitimado pela nossa necessidade de conceitualização cognitiva, povoando nossa compreensão do mundo e a forma como nos expressamos.

No caso desta tese, a noção de metáfora conceitual pode trazer importantes contribuições para a compreensão tanto do processo de simplificação como o de valorização de produtos anunciados no corpus de estudo desta tese. Nesse sentido, a metáfora inicial é discursivamente importante, pois é ela que estabelece os parâmetros interpretativos necessários para a compreensão do comercial. Em um primeiro momento, ao associar Mac e PC aos estereó-

⁶²Como já colocado na nota de rodapé 27, na página 53, Kress (2010) não deixa claro em sua obra qual o conceito de metáfora que utiliza. No caso específico deste trabalho, utilizo o conceito definido por Lakoff e Johnson (1980) de forma a suprir a necessidade de uma definição mais precisa. Não obstante gostaria de ressaltar que não é objetivo deste trabalho refletir sobre esse conceito, tomando-o como objeto de estudo. Dado o espaço exíguo desta tese, fiz uma opção por tomar o conceito de Lakoff e Johnson como uma teoria ancilar, consciente de que uma discussão mais profunda sobre a aplicabilidade desses conceitos em textos multimodais deve ser contemplada em estudos futuros.

tipos de seus usuários, está-se definindo o domínio da informática em termos do humano, traduzindo as questões técnicas em um nível palatável ao usuário comum.

Se tomarmos os vídeos da campanha estudada nesta tese como exemplo (ver figura 22, página 165), perceberemos que o indivíduo representando os computadores Macintosh (sempre o mais magro à direita), utiliza roupas invariavelmente menos formais que PC. Podemos perceber que, na maioria dos vídeos, Mac está vestido com jeans e camiseta, com o uso de um moletom modelo *jumping* ou de malha *skinny* — em *Viruses* (V-01), *Out of the box* (V-02), *Restarting* (V-03), *Networking* (V-04), *Meant for Work* (V-10) e *Fat HD* (V-11) —; mesmo quando sua roupa é mais formal — *Better Results* (V-05), *Security* (V-06) e *Vista Choose* (V-07), elas são camisas em tom azul ou chumbo, não próprias para gravata, e usadas por fora da calça. PC, por seu turno, veste-se de forma bem mais formal em praticamente todos os vídeos, estando quase invariavelmente de gravata. Mac usa tênis na maioria dos vídeos; ao passo que PC compõe seu visual formal apenas com sapatos.

Essa análise inicial das roupas é importante por nos ajudar a entender o perfil retratado para ambos os computadores ali presentes. Mac é informal, despojado, um exemplo típico do *Mac person* como colocado por Wozniak (apud KAHNEY, 2004, p. contracapa).⁶³ O objetivo é, num primeiro momento, opor esse discurso mais moderno da Apple ao corporativo e empresarial, representado pelo PC; as roupas são uma forma de criar essa impressão inicial, tornando o Macintosh imediatamente mais jovem que PC. Essa relação discursiva também se manifesta nos enredos gerais de todos os vídeos: como já citado no capítulo de metodologia, todos tratam da soluções de problemas vividos apenas no universo PC, inexistentes dentro do universo Mac. A ideia central é retratar um mundo de estereótipos baseado em oposições dicotômicas muito claras: PC é difícil de usar, cheio de problemas e pertence ao mundo corporativo, enquanto Mac é moderno, fácil de usar e se ajusta perfeitamente em seu estilo de vida. Um fator importante a ser lembrado é a caracterização dos atores que participam dos vídeos, PC traz uma associação visual clara com Bill Gates, fundador aposentado da

⁶³Ver discussão na seção 1.2.2.1, na página 57.

Microsoft, ao passo que Mac imita a forma de se vestir utilizada por Steve Jobs, cofundador da Apple. Isso pode ser um resultado do processo de sacralização já citado acima, e de um processo de espezinamento de Gates, que passa a ter sua imagem inevitavelmente associada aos fracassos dessa plataforma, claro que dentro da perspectiva discursiva ali oferecida.

Efetivamente poucos vídeos do corpus principal narram problemas típicos da plataforma PC, como é o caso de *Out of the box* (V-02), *Networking* (V-04) e *Security* (V-06); os demais vídeos do corpus, sem exceção, tratam de uma temática mais voltada para questões aplicáveis ao sistema operacional Windows, fabricado pela Microsoft e não à plataforma em si. Há, por conseguinte, uma forte tendência de generalizar esses problemas para toda a plataforma indiscriminadamente. Ao meu ver isso objetiva construir um discurso que prima por algumas características, entre elas: 1) a facilitação dos conceitos de informática para o público leigo e 2) utilização de estratégias discursivas para valorizar os produtos Apple. Com relação ao primeiro, vê-se uma postura discursiva cujo objetivo é de tentar exemplificar de forma mais simples e palpável para o usuário comum as questões técnicas envolvidas em uma discussão tão complexa como a comparação entre diferentes plataformas computacionais. Afinal, muitos dos usuários nem sequer são conscientes da existência dessas diferenças. A segunda cria uma atmosfera nos comerciais que nitidamente dá a Mac a superioridade em termos de facilidade de uso, tanto em termos de software como de hardware; mantendo um processo de oposição constante, presente desde campanhas mais antigas da marca.

No caso das duas primeiras, temos a utilização de metáforas que auxiliam na transposição de conceitos computacionais e na construção de estratégias discursivas. Elas são responsáveis por criar uma parte importante do ambiente interpretativo e temático no qual os vídeos se desenvolvem. Na figura 22, na página 165, temos o frame inicial de nove dos doze vídeos que compõem o corpus central de análise. Todos, invariavelmente, iniciam com a mesma sequência de fala e de imagens, na qual Mac está posicionado do lado esquerdo do vídeo e PC, do lado direito. Em todos os anúncios esse momento inicial serve para a apresentação das personagens; eles estão de pé, olhando para a objetiva e encarando o espectador, dizendo:

(7) **Mac:** Hello, I'm a Mac...

PC: ...and I'm a PC

É nesse momento que se estabelece a primeira metáfora, ela é um processo de associação entre a imagem e a fala, ambos se declaram representantes humanizados daquela plataforma de computadores, o que é reconhecível e aceitável graças ao seu parentesco visual com o estereótipo dos usuários e, principalmente, à similaridade física e à maneira de vestir com Bill Gates e Steve Jobs. Não é meu objetivo de pesquisa centrar-me exclusivamente na construção da metáfora enquanto objeto de análise, mas sua importância para a construção do discurso demanda um olhar mais detalhado. No caso das metáforas estabelecidas nesses textos, elas parecem estar relacionadas com dois domínios: o verbal e o visual, sendo igualmente dependente de ambos. Na figura 23, na página corrente, observa-se a interação entre dois domínios: o domínio visual, representado pela imagem de Mac e PC, e o verbal, representado pela fala. Efetivamente, esses processos metafóricos são dependentes de ambos os modos de linguagem para se constituir; a fala e os estereótipos visuais fazem existir uma metáfora bastante complexa.

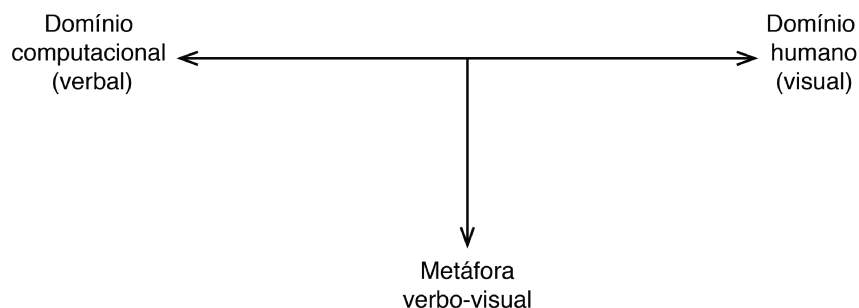


Figura 23: Constituição da metáfora de domínio computacional

Naturalmente esse relacionamento pode ser expandido, se pensarmos que diversos desses vídeos constroem sua metáfora, observaremos que, muitas delas, partem de uma metáfora definidora, ou balisar, que já é de conhecimento do espectador. Nesses casos, a metáfora verbo-visual parte de uma metáfora linguística já existente: essa metáfora seria mantida como elemento abasilar de constituição. Essa tradução, contudo, parece ser o fruto de um

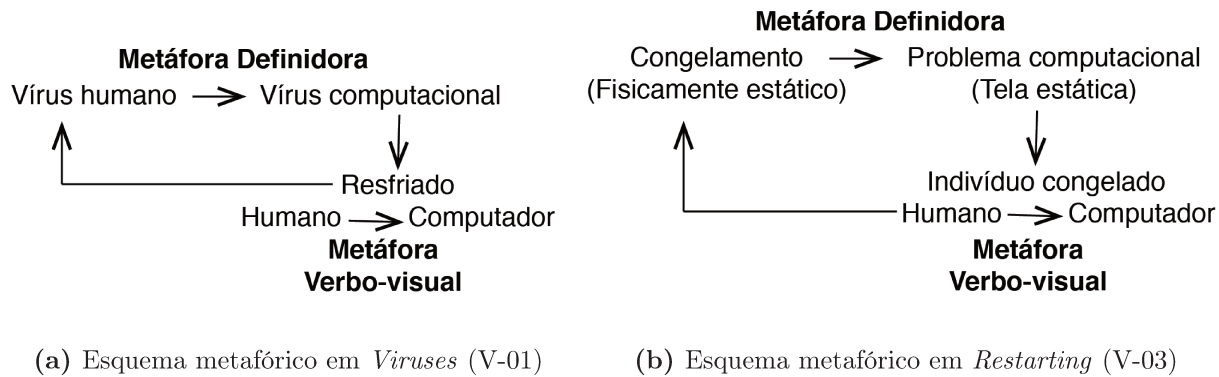


Figura 24: Metáforas em dois vídeos do corpus

complexo sistema de sobreposição de metáforas, algo que possibilita a constituição de um ponto de partida para quem assiste. Isso parece ocorrer em especial em alguns vídeos, como é o caso de *Viruses* (V-01) e *Restarting* (V-03).

No caso de *Viruses* (V-01), temos a infecção de Mac por um resfriado — ver transcrições de 12 a a 12 g, na página 176 e discussão detalhada do enredo mais à frente —, que simbolizaria um vírus computacional. Nesse caso há uma cadeia de significados que parte do senso comum: muitos sabem que computadores contraem vírus. Entretanto, a metáfora vírus de computador vs. vírus humano, em muitos casos, me parece naturalizada e pouco percebida pelos usuários, a relação entre esses domínios pode parecer quase invisível em alguns contextos. Talvez por conta disso que ela seja usada como ponto de sustentação dess processo de facilição representado pela metáfora. Já em *Restarting* (V-03) — ver transcrições de 11 a até 11 l, na página 175, e discussão mais detalhada do enredo logo a seguir — PC é um computador que trava e congela sua tela. Esse congelamento é motivado por falhas sistêmicas e é representado pelo ator imóvel. Em ambos os casos, essa metáfora definidora, que parece dar base para a argumentação desenvolvida se mostra como o ponto de partida e chegada da metáfora verbo-visual. Esta última a ela se remete, como ponto de partida da sua construção e de apoio para sua representação, como podemos ver na figura 24.

A figura 24, página 170, mostra que a relação dessas metáforas é de constante referência, um sistema de construção discursiva que me parece complexo por lidar com uma lógica tradutória que permite que diversos modos interajam na construção do significado. Em outras palavras, ao utilizar a imagem como parte constituinte da mensagem, essa metáfora verbo-visual engloba todos os modos de linguagem que são representados pela visualidade do vídeo — como gestos, roupas, postura, proxêmica, para citar alguns —, além de seus conteúdos verbais — que neste caso considero sendo não apenas a fala, mas também ruídos não articulados mas com valor comunicativos, como é o caso dos espirros e do assoar de nariz em *Viruses* (V-01). Ao mesmo tempo que a metáfora definidora acontece constantemente, servindo como base para a construção de uma nova metáfora, a metáfora verbo-visual, apesar de sua complexidade, parece ser uma interface entre o espectador e as questões tratadas nos vídeos. Tal metáfora definidora me parece ser menos óbvia para o espectador, que pode entender os conceitos computacionais colocados no vídeo sem perceber essa operação de resignificação. Essa resignificação parece se dar, por conseguinte, por camadas de metáforas. Como se pode observar na figura 24, a metáfora definidora cria uma camada mais abstrata e pressuposta que permite o desenvolvimento da metáfora verbo-visual. Seria como se ela estivesse sempre presente pairando sobre a interpretação da metáfora que facilita a compreensão do universo computacional, o que só é possível pelo fato dessa metáfora definidora estar totalmente absorvida no cotidiano desse espectador.

A partir dessa relação inicial, novas metáforas são estabelecidas, sendo dela dependentes; dois exemplos poderiam ser *Out of the box* (V-02) e *Networking* (V-04). No caso do primeiro (figura 22, na página 165), Mac e PC estão dentro de duas caixas, metaforizando a abertura dos computadores logo após a sua entrega. Partindo da expressão em inglês *out-of-the-box* — que pode significar tanto “funcionar logo que tirado da caixa” como “facilmente encontrável” —, o comercial explora a modularização do hardware PC — ele é comumente dividido em vários pedaços (monitor, CPU, alto-falantes etc.) que devem ser montados para funcionamento — vs. a produção em monobloco — monitor, CPU, alto-falantes e *drivers*, todos já

conjugados em um único volume —, comum na maioria dos computadores Macintosh. De acordo com o discurso veiculado no vídeo, isso levaria maior facilidade de montagem e operação. A constituição dessa metáfora se torna mais transparente ao espectador, novamente, em um processo de relação entre o código visual (ao ver os atores dentro das caixas) e o verbal, no momento em que Mac explica que funcionará tão logo seja ligado, como lemos em:

(8) **PC:** Oh no, not quite... have a lot to do. What's your big plan?

Mac: I might a:: make a home movie, maybe create a website, try my built in camera, I do it already out of the box... So, how about you (...)

Mac: You know it sounds that you have a lot of stuff to do before you do any stuff, I gonna get started because I am a kind of excited... Let me know when you're ready!!

PC: Actually, the rest of me is in some other boxes, so. — *Networking* (V-04)

Em *Networking* (V-04), discute-se a facilidade que computadores Macintosh têm de se conectar a hardwares periféricos (como máquinas fotográficas, impressoras e *scanners*), processo muito mais difícil de ser concretizado em qualquer que seja o Sistema Operacional (doravante SO) utilizado na plataforma PC. Ali, a metáfora se dá no momento em que Mac e PC dão as mãos, facilitando a compreensão do que significaria estabelecer uma rede de computadores e das possíveis diferenças de linguagem entre ambas as máquinas, como podemos observar no diálogo a seguir, que acontece no momento em que PC e Mac explicam o fato de estarem realizando essa ação em quadro:

(9) **PC:** We got a little network going on here

PC: a::nd it was very easy to set up

Mac: we speak each other's language — *Networking* (V-04)

Em seguida, temos a entrada de uma menina com traços orientais, aparentemente japonesa, que dá a mão para Mac e não para PC, sendo anunciada como uma câmera digital (transcrições 10 b e 10 c). Como a metáfora da rede já está estabelecida, o espectador já deduz esse fato, não obstante a percepção de que esta rede está ligada a Mac, e não a PC, só se dá efetivamente tanto no momento em que se percebe a incapacidade de PC de entender a linguagem da Câmera (transcrições 10 c e 10 d) como quando isso é explicitado verbalmente

para o público (ver transcrições 10e e 10f, ambos na página 173). Assim temos uma explicação de forma simplificada de que nem todo o computador pode operar com qualquer periférico.

- (10) (a) **PC**: who now::: Who's this now?
(b) **Mac**: This is a new digital camera from Japan, just came out...
(c) **Mac to Dcam**: em japonês incompreensível
(d) **DCam to Mac**: em japonês incompreensível
(e) **PC**: Ohohohoh PC: You speak her language?
(f) **Mac**: yeah! Absolutely! Everything kinda works on with the Mac!
(g) **Mac to Dcam**: em japonês incompreensível
(h) **DCam to Mac**: em japonês incompreensível
(i) **PC**: Bon giorno, Hello? — *Networking* (V-04)

Na verdade, temos, no estabelecimento dessa nova metáfora da câmera digital, a construção de dois elementos-chave do discurso desses comerciais, eles não só facilitam a compreensão do que seja uma rede entre computadores — assim como da conexão entre o computador e seu periférico, como excluem PC da interação com a máquina fotográfica. Tal exclusão pode se dar tanto pela linguagem verbal como na visual, ao retomarmos a figura 22d, na página 165, e o frame F da figura 19, na página 150, pode-se perceber que, simultaneamente a um dos momentos que Mac e Dcam estão conversando em japonês, simbolizando uma linguagem entendida apenas por Mac, PC está isolado à direita da composição; apesar de estar de mãos dadas com Mac (estabelecendo a rede), este tem toda a sua atenção voltada para Dcam, que se aproxima fisicamente dele. Uma das formas de interpretar essa composição é retomando o conceito de proxêmica e de olhar, como colocados por Norris (2004). Efetivamente, perceberemos que Mac e Dcam se encontram muito mais próximos formando um bloco visualmente destacado, a conexão entre eles não se dá apenas pelas mãos dadas, mas também pela troca de sorrisos e olhares, além dos braços quase que totalmente encostados e da angulação do corpo da menina, que vai claramente em direção a Mac.

Aqui a utilização da metáfora se coloca como algo complexo. Primeiro, porque a simplificação e a declaração de superioridade de Mac estão muito próximas, co-ocorrendo na

mesma construção discursiva: a capacidade de um e a incapacidade de outro se manifestam de forma linguística e visual. Apesar das diferenças de gramática nos códigos ali presentes, todos parecem direcionados para a construção de uma espécie de idiomatismo multimodal; não haveria o mesmo significado se estivessem separados de forma desarticulada. Parece-me que a interação entre esses diversos códigos geraria a criação de uma espécie de *cluster* intra-códigos, que parece permitir que esse tipo de significado articulado e complexo se instancie. Baldry e Thibault (2006, p. 31) fazem uma discussão sobre a importância da utilização de *clusters* (já comumente utilizada em estudos linguísticos) para a análise visual de imagens estáticas e sítios de internet. Para os autores, eles seriam blocos funcionais de imagens e texto que agregam significados ao serem combinados na composição visual das diversas mídias impressas e digitais. Naturalmente a questão colocada aqui é um pouco mais ampla: seria possível afirmar que os diferentes códigos são agrupados em forma de um cluster abstrato de significados que juntos criam o discurso veiculado nos textos multimodais estudados. Diferentemente do exposto por Baldry e Thibault (2006), o agrupamento considerado aqui é representado por códigos visuais e verbais, um sendo responsável por fornecer os parâmetros de interpretação para o outro, em um processo contínuo de influência.

Se momentaneamente retomarmos os conceitos básicos das três Metafunções de linguagem definidos por Halliday (HALLIDAY, 1973; 1985; 1994; HALLIDAY; MATTHIESSEN, 2004) e revistos por Lemke (2006c), observaremos que a orientacional (ou interpessoal) é responsável pela relação entre os indivíduos participantes de um ato interacional. No caso específico do discurso presente nos vídeos aqui estudados, essa diferença entre as capacidades de Mac e PC influencia a relação estabelecida entre as duas personagens no vídeo; no caso anteriormente citado — *Networking* (V-04) —, percebemos que a postura de Mac perante PC é não solidária, o que se revela tanto na postura corpórea, no olhar e na expressão facial do ator, como na proxêmica (ver figura 22d, na página 165, e o frame F da figura 19, na página 150) e na fala, no tom de espanto que Mac coloca em sua fala (ver transcrição 10f, página 173). Na verdade, essa relação de solidariedade e não solidariedade varia de acordo com o propósito

específico do texto vídeo-fílmico: no caso de *Networking* (V-04) procurava-se argumentar sobre o isolamento de PC em relação a equipamentos mais novos, de alta tecnologia, como câmeras fotográficas importadas.

Exemplos desse discurso não solidário também estão presentes em *Restarting* (V-03), que tem por objetivo discutir o congelamento e reinício, fatos mais comuns, segundo o comercial, na plataforma PC. Aqui, o problema é estabelecido quando PC, após a habitual apresentação das personagens, reinicia sua fala, causando estranheza em Mac.

- (11) (a) **Mac:** Hello, I'm a Mac
 (b) **PC:** Hello, I'm a PC
 (c) **Mac:** We have a lot in common these days
 (d) **Mac e PC:** We both run microsoft office
 (e) **PC:** We share files, it is great we just get along.
 (f) **Mac:** PC, oup, OK!
 (g) **PC:** I'm a PC
 (h) **Mac:** OK, move, pass it. We moved beyond that.
 (i) **PC:** Yeah, I had to restart that::: you know how this is
 (j) **Mac:** No, actually I don't...
 (k) **PC:** Oh What, Macs don't have to...
 (l) **Mac:** Had him and lost him... I'm gotta go get IT, keep an eye on him... — *Restarting* (V-03)

Como podemos observar na transcrição acima — que traz o texto integral de *Restarting* (V-03) —, o estranhamento de Mac em relação à situação de PC é manifestado por Mac em 11 h, na página 175, em resposta ao congelamento que ocorre em 11 f e o reinício que acontece em 11 g. Essa expressão em modo verbal é acompanhada por uma imagem de PC parado de forma estática e reiniciando a fala (repetindo parte de 11 b em 11 g). Como podemos observar na figura 25, página 176, a postura de Mac é fria e distante, com tom irônico, em nenhum momento ele se aproxima de PC ou tenta ajudá-lo, quando o problema se coloca insolúvel, ele anuncia sua intenção de chamar profissionais da área de informática, sem demonstrar nenhum tipo de solidariedade. O Frame A, ver figura 25, na página seguinte, mostra o momento do segundo congelamento de PC, que seria representado pela transcrição em 11 k e 11 l, a postura de PC é de não se aproximar (como podemos observar nos frames B e C),



Figura 25: Solidariedade em *Restarting* (V-03)

ele mal olha para PC, prefere continuar encarando o público. Sua expressão é de desdém, ver frame C; na sua saída de quadro, que coincide com o final de sua fala na transcrição 11 l, ele apenas aponta o polegar para PC, com ar de cansado da situação. Na verdade, a postura de não solidariedade de PC em relação ao congelamento de PC, parece refletir uma postura comum aos usuários da plataforma que perdem a paciência com essa situação.

No caso de *Restarting* (V-03) a metáfora do congelamento e reinício do computador se constrói na parada repentina que PC dá em quadro, ao ficar estático e começar a repetir as mesmas palavras do início do vídeo. Nesse caso, acredito que o objetivo dessa representação é demonstrar a não necessidade e perda de tempo que esse problema, que segundo o comercial é exclusivo da plataforma PC, causa para seus usuários.

- (12) (a) **PC:** Prolongados espirros
 (b) **Mac:** Sounds tight. Are you OK?
 (c) **PC:** No... I'm not OK ... I have that virus that is going around.
 (d) **Mac:** Oh Yeah...
 (e) **PC:** [Assoa o nariz] have that virus that is going around.
 (f) **Mac:** Oh Yeah...
 (g) **PC:** Assoa o nariz — *Viruses* (V-01)

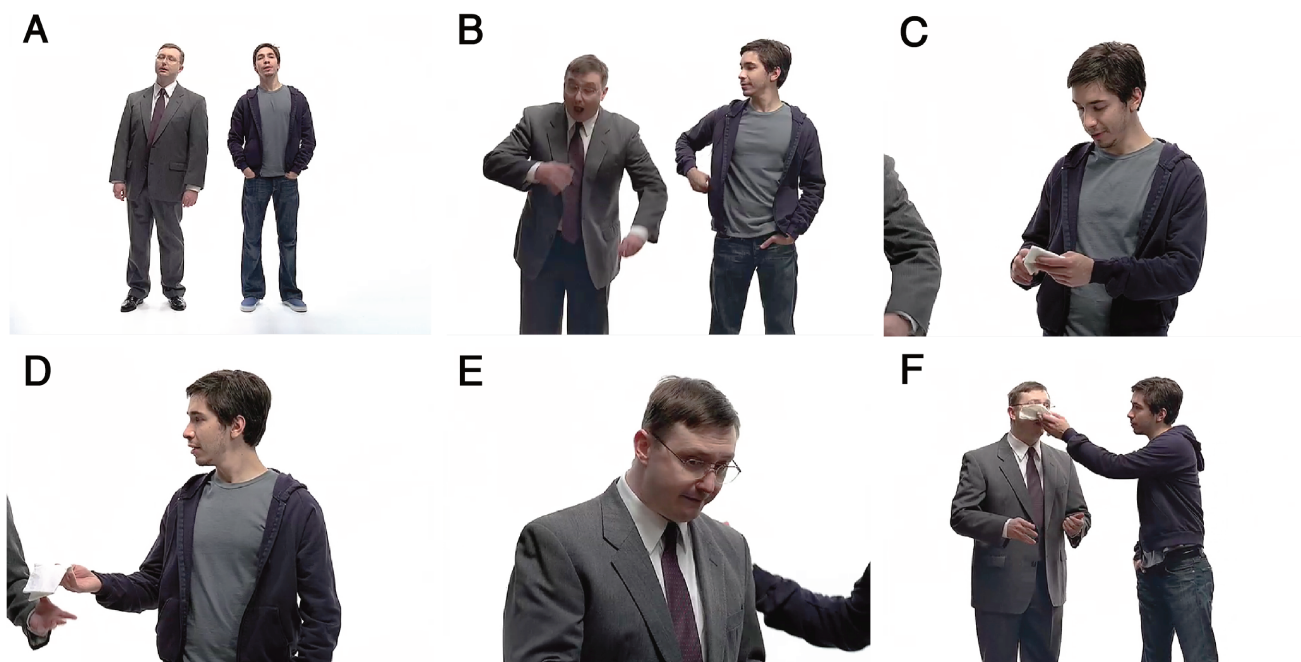


Figura 26: Solidariedade em *Viruses* (V-01)

Vale ressaltar, entretanto, que nem todas as interações entre as plataformas seguem esse caminho, em outros vídeos, há a demonstração de compaixão aos problemas sofridos por PC, como é o exemplo de *Viruses* (V-01), que tem por objetivo discutir a não infecção de Mac por vírus computacionais. Nesse texto videográfico, PC parece acometido de uma gripe, assoando seu nariz e espirando continuamente, essa gripe representa a infecção por um vírus.

A transcrição 12a, na página anterior, reflete os primeiros turnos de *Viruses* (V-01) que ocorrem durante os frames que são vistos na figura 26, nesta página. Mac demonstra solidariedade a PC pelo seu problema, o que se nota não só pelo seu tom de voz mais complacente e triste, mas também modos de linguagem. Na figura 26, página 177, temos uma série de atitudes físicas que, dentro do nosso sistema cultural de valores poderiam simbolizar a preocupação de Mac em relação ao estado de saúde de PC.

No frame A, há os primeiros sintomas da doença de PC, sua voz em estado quase letárgico e nasalizada, ao passo que em B a sequência de espirros que coincide com o primeiro turno da transcrição (ver transcrição 12a, na página anterior). A partir desse momento Mac expressa

sua solidariedade com a pergunta presente na figura 12 b, na página 176, que acontece entre os frames B e C. Nesses mesmos frames, Mac inicia a retirada de um lenço de papel do bolso e o entrega para PC em D, momento em que demonstra estar preocupado com ele, dando-lhe um pequeno tapa nas costas e afagando-o, como vemos em F. O frame F é um momento decisivo para esse processo, uma vez que Mac toma a frente, ajudando PC doente e infectado por vírus. Esses atos de ajuda e cuidado, refletindo uma relação amigável entre as duas plataformas está, na verdade, servindo como um dos instrumentos instanciadores da superioridade de Mac, dado que ele pode realizar tais ações porque a diferença entre plataformas não permite que os vírus computacionais sejam prejudiciais a ele. Essa é a razão pela qual Mac não se afasta quando lhe é solicitado, como vemos em:

- (13) (a) **PC**: You'd better (2x) to stay back::... you...
(b) **Mac**: That's OK, I will be fine
(c) **PC**: No, No, No, Do not be here, at least there are 140000 known viruses for PCs
(d) **Mac**: PCs, not Macs:: — *Viruses* (V-01)

Seu ato de solidariedade incondicional, representado pela aproximação de Mac de PC, não é uma atitude gratuita. Ao declarar sua imunidade para PC, Mac está, na verdade, salientando uma das características de seu sistema: a não infecção por *Malware*, algo tão típico na plataforma PC, parece contar decisivamente para a proximidade de ambos. A busca por uma representação mais palatável para as diferenças entre as plataformas faz com que essa simplificação busque, dentro dos textos audiovisuais que estão sendo estudados, uma forma também mais próxima de representar a diferença e superioridade de Mac. É lógico que o universo discursivo representado por esses comerciais prega invariavelmente a superioridade de Mac sobre PC, dado que esse é o objetivo comercial da campanha.

Em *Viruses* (V-01), a abordagem utilizada foi, novamente, humanizar as questões computacionais por meio de metáforas. Muitas dessas metáforas, como é o caso aqui, entram no mundo da informática já por meio de metáforas definidoras, um exemplo seria a noção de vírus, que em termos computacionais não são organismos vivos, mas programas maliciosos que visam causar danos ao sistema operacional em questão. Ao tornar os vírus computacionais

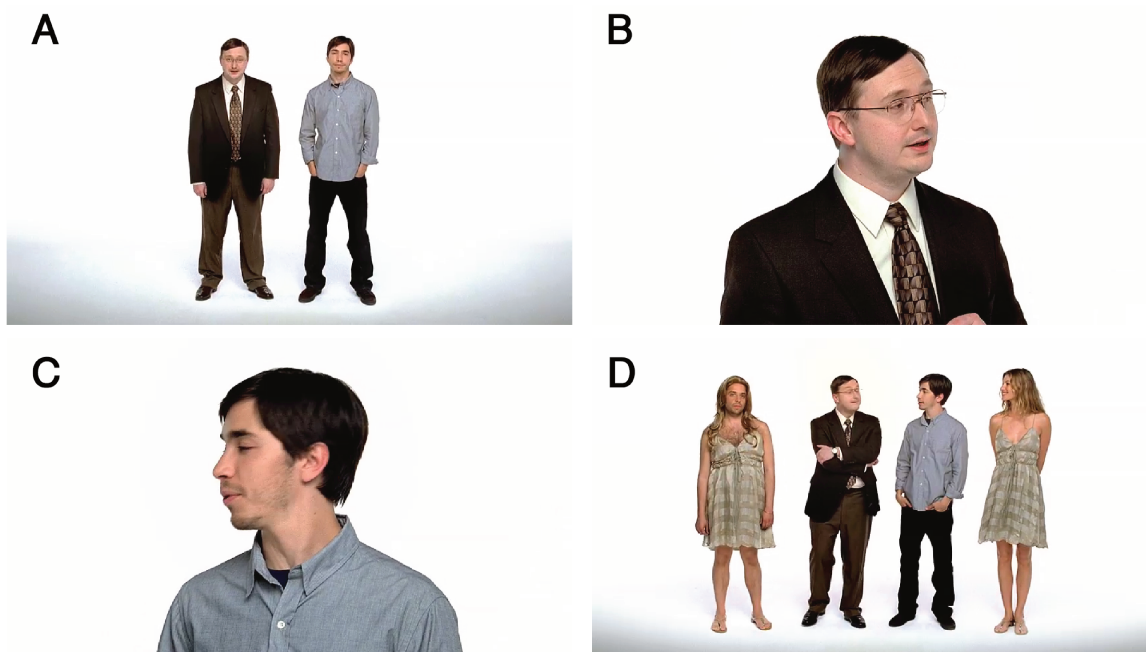


Figura 27: Planos e composição em *Better Results* (V-05)

em humanos novamente, os autores do vídeo estão criando uma instância de explicação de um conceito que, apesar de já ser uma metáfora, parece não totalmente transparente para quem não está bem próximo do universo da informática. Naturalmente que esse processo de simplificação é orientado de forma a criar simpatia ao produto vendido pela Apple, que pelo menos é imune aos problemas.

5.2 Organização e padrões em "Get a Mac"

Um atento olhar ao corpus de estudo permitiu a observação de diversas regularidades e padrões. Tais regularidades parecem ter um objetivo claro: ao se deparar com qualquer vídeo da campanha, qualquer que seja o meio de distribuição — vale lembrar que esta campanha contou com a distribuição dos vídeos via internet e pela lojas Apple, além de um menor número ser veiculado em TVs dos EUA —, o leitor desses textos multimodais seria capaz de identificar o vídeo como pertencente à campanha. Essas regularidades parecem ser de, pelo

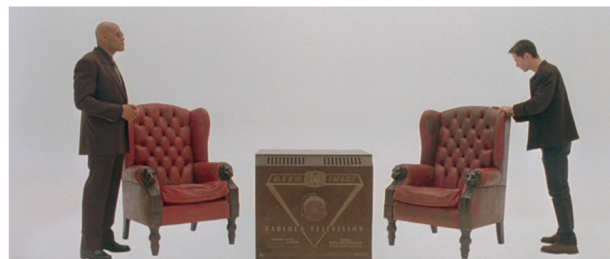
menos, dois tipos: 1) estrutura composicional do quadro e 2) fases que compõem a progressão textual.

A estrutura composicional do quadro pode ser visualizada na figura 27, na página precedente. Ali temos diversos frames de *Better Results* (V-05), que tem por objetivo demonstrar que Mac realiza tarefas multimídia mais competentemente que computadores PC. A abordagem discursiva presente nesse comercial não é diferente das discutidas na seção anterior: PC declara ser capaz de realizar vídeos caseiros e é interrompido por Mac, que também se diz capaz de fazê-los. Ao mostrar seu resultado, interpretado por uma modelo real, metaforizando a boa qualidade da produção, PC quer retirar-se de quadro. Pela insistência de Mac, PC decide mostrar seu vídeo, trata-se de um homem vestido com as mesmas roupas da modelo, imitando-a e simbolizando o amadorismo e fracasso de PC.

Como vemos na figura 27, na página anterior, a composição do quadro nestes comerciais é muito simples, temos um fundo infinito na cor branca, com apenas algumas luzes marcando o chão (visíveis em A e D). Os planos também não sofrem muita alteração, variam entre o plano geral e o plano médio-fechado,⁶⁴ focalizando especificamente a personagem que está com o turno de fala.



(a) *THX 1138* (Lucas, 1971)



(b) *Matrix* (WACHOWSKI; WACHOWSKI, 1999)

Figura 28: Intertextualidade entre "Get a Mac" e o cinema

Apesar da aparente neutralidade do cenário — uma estratégia comum para valorização do produto em vídeos publicitários —, é possível o estabelecimento de uma referência extra-textual a dois filmes de ficção científica. O primeiro é chamado *THX 1138*, de George

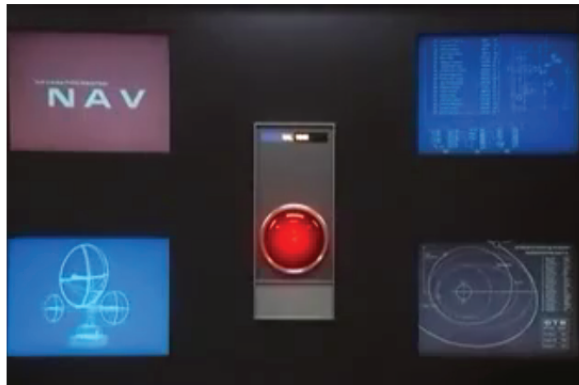
⁶⁴Ver figura 15, na página 121, para uma discussão sobre os planos em vídeo.

Lucas (LUCAS, 1971), que reflete sobre a sociedade de forma apocalíptica, representando os humanos como seres incapazes de pensar e sentir, vivendo em um mundo higienizado e sem personalidade; o segundo filme é chamado *Matrix* (WACHOWSKI; WACHOWSKI, 1999). Nesse último, Neo (o messias anunciado) descobre que o mundo em que vivemos é uma projeção de realidade aumentada por meio de uma matriz computacional que nos aprisiona, numa clara analogia ao Mito da Caverna, de Platão. O momento exato que essa relação intertextual se estabelece é trazida pela figura 28a, no caso de *THX 1138*, e pela figura 28b, na página 180, no caso de *Matrix*. Ali, podemos observar que há algumas diferenças visuais entre o cenário em fundo neutro da campanha analisada e os frames retirados dos filmes: *Matrix* possui um tom mais escuro no pano de fundo, ao passo que *THX 1138* e os comerciais da Apple são mais claros. Apesar disso, a forma de composição visual nos três textos vídeo-filmicos é bastante similar: os atores estão na parte central do quadro sem uma grande quantidade de móveis ou cenários. A Apple inova em sua produção à medida que divide o quadro geograficamente: tudo que está relacionado a Mac ocorre do lado direito do vídeo ao passo que o lado esquerdo é dedicado a PC.

Ao se assistir aos filmes de “Get a Mac” a função dessa relação intertextual não é imediatamente clara, todavia algumas inferências podem ser realizadas. *Matrix* foi um sucesso de bilheteria, atingindo em especial os admiradores de ficção científica e os aficionados por computadores, os chamados Geeks; ao passo que *THX 1138* se transformou em um clássico da estética *cyberpunk*.⁶⁵ Realizar uma citação direta de uma obra como essa é um fator convalidativo para esses grupos que podem se sentir prestigiados por conta dessa citação. Além disso, *Matrix* foi responsável por uma revolução na forma como computadores foram usados

⁶⁵ *Cyberpunk* pode ser definido como um gênero dentro da ficção científica que se define pelo seu enfoque distrópico e apocalíptico, baseando-se no conceito-chave “Alta tecnologia, baixo nível de vida” ou “*High tech, Low life*”. Todos os filmes citados nesta tese e referenciados como constituindo alguma relação intertextual com os comerciais analisados (*2001 — Uma Odisséia no Espaço*; *THX 1138*; *Matrix* e *Mil Novecentos e Oitenta e Quatro*) são expressões desse gênero, que também está ligado à literatura e, atualmente, também à indústria de games. Bons locais para discussão e maiores informações sobre o estilo seriam: <http://project.cyberpunk.ru/idb/index.html> e <http://www.cyberpunkreview.com/>.

na produção de efeitos especiais, tendo seus efeitos imitados por dezenas de filmes, criando o que Sdjic (2008)⁶⁶ chamaria de um arquétipo do design filmográfico.



(a) Bug do Milênio



(b) Ninety eighty four

Figura 29: Intertextualidade audiovisual em duas campanhas Apple

Além disso, há uma tradição da Apple de relacionar suas campanhas a outras produções audiovisuais de importância histórica e estética. Dois exemplos dessa tendência são trazidos pela figura 29, na página corrente. Ali temos em 29a um comercial produzido pela empresa por ocasião do chamado Bug do Milênio (por volta de 1999) para explicar que computadores Macintosh estariam imunes ao problema; nesse anúncio temos como única personagem: Hall 9000, um computador que se tornou célebre no filme de Stanley Kubrick (1968), *2001-Uma odisseia no espaço*. No filme, Hall 9000 é um computador que chega a demonstrar uma personalidade forte e um grande senso de autopreservação, com traços de humanidade muito mais fortes e aparentes que os técnicos que o operam. Nesse comercial, Hall 9000 discute a maior capacidade dos computadores Macintosh, rendendo-se a essa evidência e se enciumando deles, como podemos perceber pelo último trecho de sua fala:

(14) **Hall 9000:** You like your Macintosh better than me, don't you Dave? Dave? Dave? ... Can you hear me?

Já na figura 29b, temos o primeiro comercial dos computadores Macintosh, exibido em 1984 por ocasião do lançamento da plataforma. Nele, há uma citação visual nítida do filme

⁶⁶Ver discussão sobre arquétipos de design em 1.2.2.2, na página 60.

dirigido por Michael Radford (1984) e chamado de *Mil Novecentos e Oitenta e Quatro*, baseado livro homônimo de George Orwell. Esse comercial simula o momento ao final de cada dia em que os moradores da Oceania (nome do regime totalitário da Inglaterra no livro e filme) ouvem as mensagens do Grande Irmão (líder do regime). Tanto no livro quanto no filme, todos ouvem as palavras pacientemente, mas no comercial uma mulher carregando um martelo de arremesso invade a sala e o arremessa contra a grande tela, espatifando-a. Nesse momento vê-se escrito o seguinte letreiro:

(15) On January 24th
Apple Computer will introduce Macintosh
And you'll see why 1984
won't be like "1984"



Figura 30: Vídeo da campanha *Switch*

A ideia de usar referências a filmes de cinema em vídeos publicitários não é uma exclusividade da Apple, dezenas de comerciais são pensados e realizados dessa forma todo ano. A diferença é na natureza dos filmes que a Apple decide citar como base para o design de suas propagandas: invariavelmente são clássicos de ficção científica ou filmes distrópicos. Esse talvez seja um aspecto discursivo que recaia sobre o design e concepção das peças: ao escolher realizar referências a esses filmes, a Apple também está buscando atingir pessoas capazes de compreender essas referências, trazendo-as para o consumo de suas peças ou as fidelizando.

Não pode se esquecer que a Apple mantém uma relativa unidade entre sua campanha anterior chamada "Switch" (do inglês mudar, trocar), realizada entre 2002 e 2004, e "Get a Mac". Nessa campanha, usuários comuns davam seu depoimento sobre o que os motivou a trocar de plataforma, migrando de PCs para Macs. Em ambas, temos o mesmo fundo branco infinito, mas com diferenças na temática e constituição dos planos; em "Switch", como podemos ver na figura 30, na página 183, o plano é exclusivamente médio-fechado, sem movimentos ou variações de câmera. O vídeo representado na figura 30, em especial, foi um sucesso na TV e um viral de internet, dado o aspecto alterado de sua atriz, especulava-se que ela estaria sob efeito de psicoativos.

Na verdade, a manutenção de identidade entre "Switch" e "Get a Mac" é algo importante por ajudar o espectador a reconhecer uma coerência visual entre os anúncios da Apple, além de estabelecer uma importante relação com os produtos produzidos pela Apple nesta época, a maioria na cor branca.

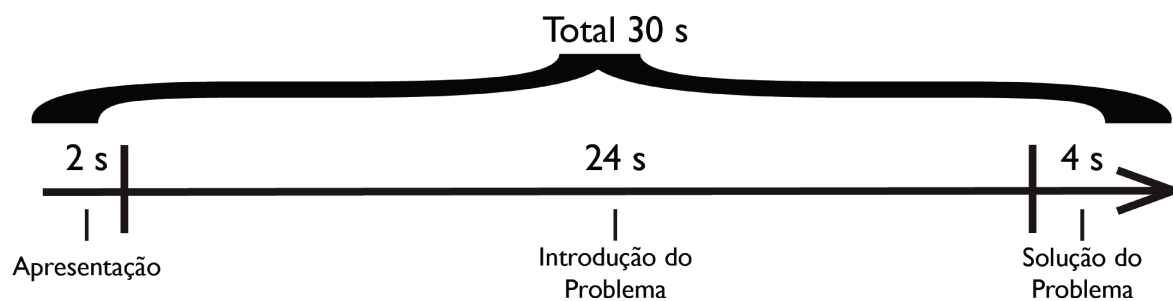


Figura 31: Divisão temporal nos vídeos

Com relação à constituição textual desses filmes, percebe-se que todos eles são formados por três fases distintas, cada uma delas com uma função específica dentro dos comerciais. A primeira, Apresentação, Mac e PC se apresentam para o espectador, na segunda, Introdução do Problema, PC e Mac discutem algum problema sofrido pela plataforma PC, mas não pela plataforma Mac. Já na última fase, Solução do Problema, PC e Mac atestam a superioridade do último em relação ao primeiro. Como já colocado no capítulo de metodologia, cada um dos vídeos tem a duração exata de 30 segundos, algo padrão para peças comerciais em vídeo, sendo que cada uma dessas fases possui uma duração padrão. Como podemos observar na

figura 31, na página 184, esses 30 segundos são divididos de forma igual entre todos os vídeos, sempre com maior tempo para a segunda fase, central no processo de argumentação discursiva para Mac. Isso não significa que as demais devam ser esquecidas, ou mesmo que não agreguem significados ao texto, sendo meramente formais. Na verdade, o maior tempo para Introdução do Problema advém do fato dela estabelecer uma discussão mais detalhada sobre a natureza da superioridade de Mac. A figura 31, na página oposta, mostra como essa divisão temporal acontece nos vídeos, ajudando a entender como ela ocorre proporcionalmente.

Na Introdução (ver figura 31, na página anterior), Mac e PC se colocam lado a lado no primeiro frame do comercial e se apresentam para o público. Um exemplo de como essa apresentação se dá visualmente pode ser visto na figura 22, na página 165. Ali Mac e PC estão postados à esquerda e à direita do vídeo, respectivamente. Em todos os vídeos do corpus essa configuração visual se aplica. Além do padrão visual de apresentação, essa fase inicial oferece também padrões verbais de realização e de duração, entre 01 e 02 segundos iniciais sempre no começo da peça. Com relação à linguagem verbal, temos o estabelecimento de um padrão de autoenunciação que aparece regularmente nesses vídeos:

- (16) **Mac:** Hello, I'm a Mac
PC: . . . and I'm a PC

Como já colocado anteriormente, essa fase é importante por estabelecer a metáfora inicial que será necessária para a estruturação de todo o filme publicitário. Entre as 3 fases que compõem os vídeos da campanha, ela é a única que oferece algum tipo de regularidade linguística e visual, isso porque as demais variam bastante em seu conteúdo e composição devido à grande gama de temas e de recursos de linguagem. É importante lembrar que a composição básica do quadro, inspirado em *Matrix* e *THX 1138*, não se altera em momento algum. Exatamente por ser a única fase que possui uma maior padronização nos usos de linguagem, é que é possível perceber variações em algumas das peças publicitárias que compõem tanto o corpus de pesquisa como os vídeos da campanha como um todo. Se observarmos em termos quantitativos, em torno de 16 deles, $\frac{1}{4}$ do total de vídeos, possui algum tipo de quebra da

Quadro 5 Quebras verbais na apresentação

Vídeos		Transcrição
A	<i>Security</i> (V-06)	Mac: Hello, I'm a Mac...
		Firewall: Mac is greeting salutation cancel or allow?
		PC: Allow.
		PC: And I'm PC...
B	<i>Vista Choose</i> (V-07)	Firewall: You are returning Macs salutation, cancel or allow?
		Mac: Hello, I'm a Mac...
C	<i>Boxer</i> (V-09)	PC: And I'm PC... and its time to play cho:::::se a vista!
		Mac: Hello, I'm a Mac...
D	<i>Fat HD</i> (V-11)	Apresentador: In the left corer, coming in, at 80 gigs, him, from pops desktop everywhere, P(3):::::::::::C!
		Mac: Hello, I'm a Mac...
		PC: Hold up, here I come...

estrutura linguística da apresentação. Normalmente essas exceções se dão de forma a servir ao argumento do vídeo.

Normalmente essa quebra no padrão verbal e visual leva a uma maior duração dessa fase, ultrapassando seus 2 segundos habituais; esse tipo de quebra no padrão verbal, vêm, na maioria das vezes, acompanhada por quebras no padrão visual da fase de apresentação. Na maioria das vezes, essas quebras são realizadas pela introdução de um elemento estranho no quadro ainda na fase de apresentação. A figura 32, na próxima página, traz exemplos de como a quebra de padrão verbal pode ocorrer de forma conjunta com a visual. Em *Security* (V-06), figura 32a, na página 187, Mac e PC têm sua interação interrompida por um homem de preto, que representa um firewall responsável por intermediar toda a comunicação entre eles; em *Vista Choose* (V-07), na figura 32b, na página 187, ao invés de se apresentar, PC introduz um jogo de roleta que tem a função de ajudar a escolher a melhor distribuição de Windows Vista; em *Boxer* (V-09) (figura 32c), logo após Mac terminar seu turno, o narrador de Boxe inicia sua apresentação de PC, que adentra o quadro com movimentos pugilísticos. Por fim, *Fat*

HD (V-11) (ver figura 32d, nesta página) discute como os *trial softwares*⁶⁷ disponibilizados na maioria dos computadores PC levam a problemas de utilização do computador, ocupando espaço e o deixando lento. Nesse vídeo a imagem de PC gordo representa exatamente esse tipo de realidade: o domínio humano representaria o estado do novo computador e seu HD.



Figura 32: Exemplos de quebra de padrões na apresentação

O quadro 5, página 186, traz a transcrição de "Introdução do problema" nos quatro vídeos representados na figura 32, página 187; podemos, então, observar que a motivação da quebra do padrão visual e do verbal parece ser a mesma, fazendo com que ambos caminhem conjuntamente. Em *Security* (V-06), a presença de Firewall como figura intermediária faz com que a quebra da fluidez do discurso se faça necessária de forma a atribuir seu valor na interação. No caso de *Vista Choose* (V-07), a necessidade de mostrar grande número de

⁶⁷São softwares proprietários que possibilitam ao usuário um período de testes, no qual ele pode ser utilizado sem o pagamento das licenças, mas com limitação funcional. Esse tipo de programa, muitas vezes, vem disponibilizado em novos computadores graças a acordos comerciais de distribuição, à revelia do comprador, ou podem ser baixados da internet.

opções disponíveis, opções que não são necessariamente avaliadas positivamente no vídeo, e a dificuldade de escolha gerada por elas, fazem com que o anúncio da roleta impeça a natural devolução do turno a Mac. Já *Boxer* (V-09), também discutido anteriormente na seção 4.3, ver figura 20, na página 153 e quadro 3, na página 154, tem sua motivação na metáfora de que o sucesso de Mac é algo a ser derrubado e vencido por PC em uma luta; o que explica o anúncio realizado pelo apresentador de lutas. Por fim, *Fat HD* (V-11) tem uma fala que pede calma devido à dificuldade de locomoção de PC, numa clara associação ao seu peso.



(a) *Out of the box* (V-02)



(b) *Breakthrough* (V-12)

Figura 33: Vídeos apenas com quebra do padrão visual

Vale ressaltar que nem toda a quebra no padrão visual leva a uma quebra no padrão verbal e vice-versa; aqui exemplos nos quais tais quebras se dão em separado: *Out of the box* (V-02), ver figura 33a e *Breakthrough* (V-12), representado na figura 33b, ambas na página 188. No caso de *Out of the box* (V-02) a apresentação de Mac e PC dentro de caixas vem do uso metafórico de *out-of-the-box*.⁶⁸ Mac aparece dentro de uma caixa branca (marcando a identidade de cor dos produtos Apple) e PC em uma caixa marrom claro. Eles começam a discutir o que farão quando “saírem da caixa”; Mac diz que já vai logo começar a trabalhar, ao passo que PC tem de passar por todo um processo de montagem antes de se tornar funcional. Mac salta da caixa e se despede de PC, que tem de esperar seus pedaços em outras caixas. Em *Breakthrough* (V-12), Mac e PC aparecem lado a lado na frente de alguém que aparenta

⁶⁸Expressão em inglês que pode significar tanto algo que funciona "assim que é retirado da caixa" como também algo que é "facilmente encontrado"; aqui Mac representa a primeira.

ser um terapeuta; o mote dessa peça é uma espécie de tentativa de superação dos traumas que PC sofrera por ser uma plataforma computacional inferior a Mac. Em ambos temos uma quebra no padrão visual dessa fase sem que exista uma quebra no padrão verbal: a apresentação é instanciada da mesma forma que vemos na transcrição 16, na página 185.

Essas quebras de padrão parecem ser resultado da construção do argumento nesses vídeos: procura-se representar algo que necessita de uma configuração especial dos modos de linguagem, algo que não pode ser atendido pela padronização habitual. As quebras parciais parecem ir pelo mesmo caminho; todavia, ao respeitarem o padrão verbal, parecem colocar no visual toda a responsabilidade pela adaptação necessária.



(a) *Security* (V-06)

(b) *Vista Choose* (V-07)



(c) *Meant for Work* (V-10)



(d) *Fat HD* (V-11)

Figura 34: Composições visuais em “Introdução do Problema”

A segunda fase dos textos audiovisuais estudados nesta tese foi nomeada de "Introdução do Problema". Ela tem tamanho estável (24 segundos em média), mas seu conteúdo informativo é bastante variado por depender inteiramente do argumento específico tratado em cada vídeo. Via de regra, ela vem logo em seguida da abertura e tem a função de expor algum problema da plataforma PC que é facilmente resolvido, ou mesmo não é uma questão problemática, na plataforma Macintosh. Nesta fase, a tradução em metáforas — que serve ao

processo de facilitação e de afirmação de superioridade da plataforma Mac, já anteriormente discutido — é a responsável por determinar o tipo de interação, além de ter uma grande influência na composição visual. A figura 34, na página 189, traz o frame inicial dessa fase em 4 vídeos: *Security* (V-06), figura 34a, *Vista Choose* (V-07), figura 34b, *Meant for Work* (V-10), figura 34c, *Fat HD* (V-11) figura 34d.

Quadro 6 Comparação em Introdução Problema-1

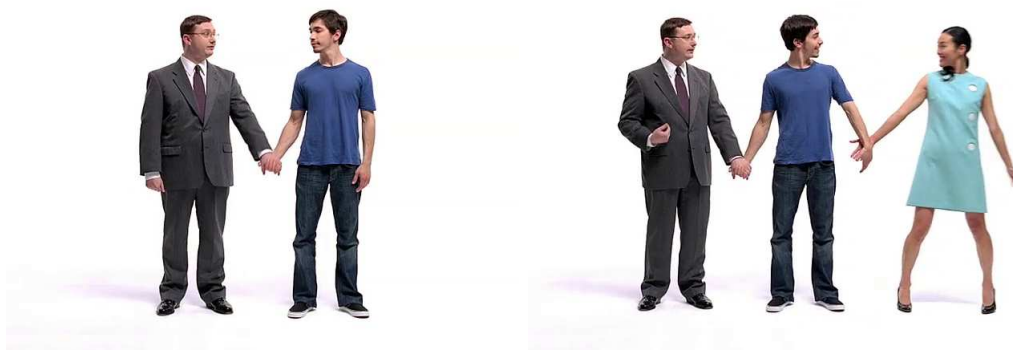
Vídeos		Transcrição
A	<i>Security</i> (V-06)	Mac: Ok, What is this...?
		Firewall: Mac is asking question, cancel or allow?
		PC: Allow! He is part of Vista, my new operating system.
B	<i>Vista Choose</i> (V-07)	Mac: Oh! what's going on?
		PC: Well Vista comes on six different versions and I don't know which to choose.
C	<i>Meant for Work</i> (V-10)	Mac: ... Oh! Oh! What happened to you PC?
		PC: Kids happened!!
D	<i>Fat HD</i> (V-11)	Mac: Mac: Oh! PC!
		PC: It's all these trial software... they pack my hard drive full...

Em *Security* (V-06), um segurança faz o papel de firewall, filtrando as falas entre PC e Mac, e mostrando as dificuldades operacionais que a segurança exigida por PC traz, ao passo que em *Vista Choose* (V-07) há uma roleta situada entre PC e Mac, com o objetivo de ridicularizar as diversas escolhas de sistema operacional disponibilizadas pela Microsoft, gerando dificuldades técnicas de escolha por parte do cliente. Já em *Meant for Work* (V-10), PC aparece cheio de adesivos e com cabelo todo desarrumado. Visualmente isso se justifica pelo problema apresentado na peça publicitária: PC fora comprado para ser usado por crianças em tarefas criativas, elementos que, segundo a representação ali veiculada, seriam mais próprios para Mac, uma vez que PC é um computador corporativo e para adultos. Por fim, em *Fat HD* (V-11) a representação física de PC é de uma pessoa extremamente obesa, o que é fruto de uma metáfora de facilitação: o disco rígido do computador estaria cheio

devido à grande quantidade de *trial software* (programas de computador em versão teste) que se acumulam em PCs novos.



(a) Elemento externo em *Viruses* (V-01)



(b) Elemento externo em *Networking* (V-04)



(c) Elemento externo em *Better Results* (V-05)

Figura 35: Elementos externos em "Introdução do Problema"

Na maioria dos vídeos, o problema da plataforma PC é sempre inicialmente detectado por Mac, que pergunta para ele qual sua dificuldade nesse momento, como podemos observar no quadro 6, na página oposta, que compara as diferentes formas de início de “Introdução do Problema” nos 4 vídeos retratados na figura 34, na página 189. De fato, o padrão linguístico dessa fase é mais dilatado que a fase anterior; muito do que é dito depende do conteúdo

tratado em cada comercial, sendo um padrão mais funcional do que estrutural. Nele há duas formas de iniciação e percepção do problema; na primeira, ela se inicia com uma interjeição de espanto e uma pergunta, respondida por PC.

No caso de *Security* (V-06), indivíduos metaforizados por questões computacionais também passam a integrar a composição visual, assim como o desenvolvimento linguístico do texto vídeo-filmico. Ali há um firewall, cuja função é bloquear acesso de outros computadores a PC quando conectado à internet, por conta disso seu turno é constantemente intercalado entre PC e Mac. Já os vídeos *Vista Choose* (V-07), *Meant for Work* (V-10) e *Fat HD* (V-11) não introduzem indivíduos, mas elementos que podem interferir na constituição visual, logicamente esses elementos são referenciados no texto falado nesses vídeos. Neles, os elementos introduzidos já estão presentes na "apresentação" e já são conhecidos do leitor nessa parte do vídeo. Entretanto existem diversos outros casos nos quais a introdução de outros conceitos ou indivíduos estranhos a Mac e PC ocorre em "Introdução do Problema", como seria o caso de *Viruses* (V-01), *Networking* (V-04) e *Better Results* (V-05), trazidos pela figura 35 na página anterior.

No caso de *Viruses* (V-01) (figura 35a, na página 191), esses elementos externos não são um indivíduo ou objeto, mas um conceito abstrato associado a um comportamento específico de PC: seus espirros simbolizam um processo de metaforização do domínio computacional. O conceito de vírus computacional já é um processo metafórico que traduz para o ambiente da máquina um conceito do mundo natural biológico; ao mesmo tempo, representar esse vírus como uma doença humana é devolver o conceito para o domínio de que foi retirado. Apesar da complexibilidade dessa operação, a tradução realizada parece ser de simples entendimento.

Já em *Networking* (V-04), (figura 35b), a máquina fotográfica japonesa, menina de vestido que entra em quadro, parece instanciar, como já discutido anteriormente na página 150, um estereótipo que serve para traduzir na noção de conexão e do que significa conectar uma câmera estrangeira em seu sistema. Por fim, *Better Results* (V-05) (figura 35b, página 35b) possui dois elementos complicadores, inseridos em momentos diferentes do vídeo. Essa peça

Quadro 7 Comparação em Introdução Problema-2

<i>Out of the box (V-02)</i>	Indagação inicial	Mac: Ready to get started?
	Real introdução do problema	PC: Oh no, not quite... have a lot to do. What's your big plan?
	Sistematização do problema (em turnos)	Mac: I might a::: make a home movie, maybe create a website, try my built in camera, I do it already out of the box... So, how about you? PC: Well, first I have to download some new drivers, I have to erase the trial software that comes on my hard drive, and I got a lot of manuals to read Mac: [sobrepondo PC]: Sweet!! You know it sounds that you have a lot of stuff to do before you do any stuff, I gonna get started because I am a kind of excited.. Let me know when you're ready!!
	Indagação inicial	PC: We got a little network going on here... a:::nd it was very easy to set up...
<i>Networking (V-04)</i>		Mac: we speak each other's language...
	Real introdução do problema	PC: who now::: Who's this now?
		Mac: This is a new digital camera from Japan, just came out
	Sistematização do problema (em turnos)	Mac to Dcam: [em japonês incompreensível] DCam to Mac: [em japonês incompreensível] PC: You speak her language? Mac: yeah! Absolutely! Everything kinda works on with the Mac!

publicitária tem por objetivo discutir as habilidades que os SOs têm ao realizar vídeos caseiros. A modelo Gisele Bündchen entra em quadro para representar o filme de Mac; ela insiste para PC mostrar o seu, que mesmo resistindo ele o faz. Seu vídeo é representado por um homem de perucas e roupas iguais as de Gisele. Novamente, há o trabalho com estereótipos e metáforas: a qualidade do vídeo parece ser traduzida na capacidade da plataforma de mimetizar algo que parece "belo como o real", quesito que a modelo parece ser mais capaz de representar; a falha de PC parece estar em sua incapacidade de produzir uma "Gisele".

Esse elemento externo é essencial para criação de um complicador que é responsável pelo desenvolvimento do argumento entre Mac e PC, é sobre esse complicador que as duas personagens desenvolvem o conteúdo e as comparações entre eles, criando o conflito necessário para o desenvolvimento dos textos.

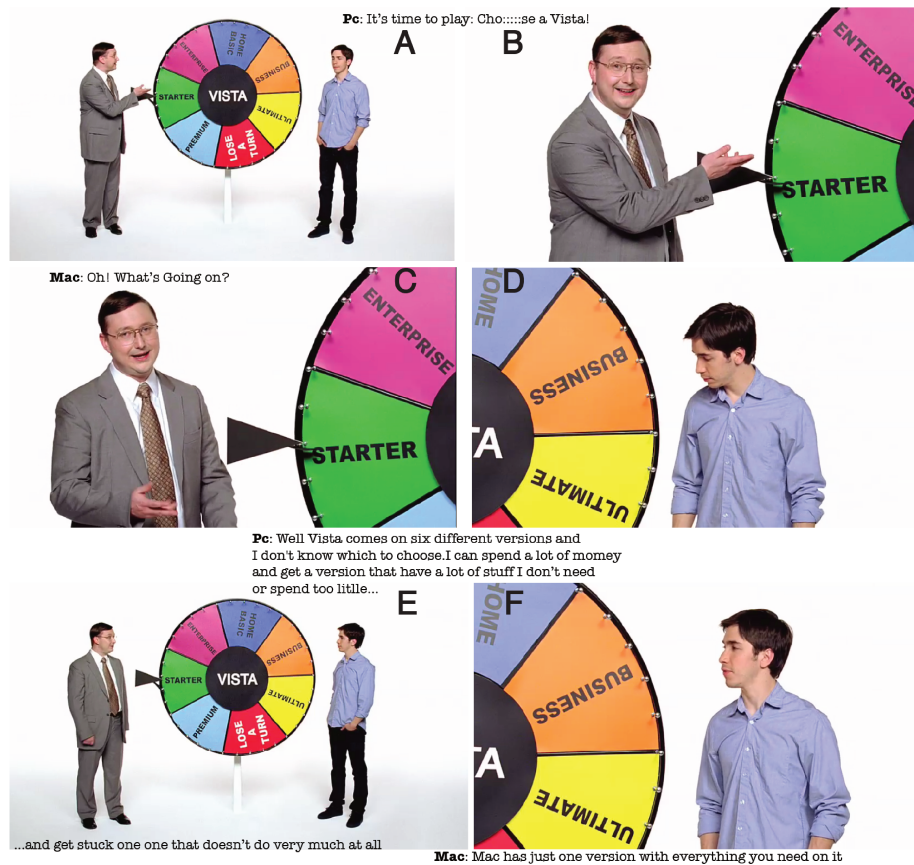


Figura 36: Cortes e olhar em *Vista Choose* (V-07)

A esses turnos iniciais, segue-se uma explicação mais detalhada do problema por parte de PC, com eventuais interrupções de Mac. Uma transcrição completa dessa fase pode ser observada no quadro 7, na página precedente.

Como podemos observar no quadro 7, na página anterior, essa fase pode ser subdividida em, pelo menos, três outras: indagação do problema, seria referente à questão inicial, realizada por Mac ou PC, que desencadeia a discussão sobre o problema. Ela pode se dar em um ou em diversos turnos, dependendo da especificidade de cada problema anunciado, sendo seguido por um turno único, colocado por Mac ou PC, chamado de Real introdução do

problema, cuja função é revelar o problema de forma direta para o espectador a partir das metáforas computacionais já discutidas. Em sistematização do problema temos uma troca de turnos entre Mac e PC, na qual eles discutem o problema, explicando-o.



Figura 37: Frames de Solução do Problema em *Out of the box* (V-02)

É nesse momento que a discussão da diferença entre as plataformas parece se estabelecer de forma efetiva, aqui o problema é efetivamente explicado para o espectador que passa a compreender porque a plataforma Mac parece ser melhor que a PC. Como podemos observar no quadro 7, na página 193, essa fase ocorre em tom de conversa, com Mac e PC intercalando turnos de fala. Não há um padrão que estabeleça qual personagem deve falar primeiro tampouco quanto tempo de fala é dado a cada um; como podemos observar no quadro 7, o número de linhas dedicado à fala de cada personagem também pode variar bastante. A quantidade de turnos nesta subfase depende do tipo de metáfora computacional de simplifi-

cação escolhida. Em minha opinião essa grande variação está relacionada à necessidade de constante criatividade que os autores de textos como este parecem sofrer. Como já discutido no capítulo 1, apesar de apresentarem certas regularidades, anúncios publicitários são conhecidos por inovar constantemente em sua estrutura.

As escolhas visuais também parecem estar relacionadas com a temática discutida em cada um dos vídeos, os enquadramentos parecem alternar de acordo com a necessidade de fazer com que o espectador vivencie o processo de interação ali presente. Como percebemos pela comparação entre alguns frames da figura 36, página 194, um plano geral (frame A) estabelece a existência da roleta a ser usada para sortear a versão do software Windows é colocado bem no centro do quadro, dividindo-o em *Vista Choose* (V-07). Enquanto PC olha para a roleta e para o público o tempo todo, Mac olha incessantemente para a roleta e PC (frames de D a F), de forma a entender o que está acontecendo; mesmo quando fala, com eventuais desvios para a câmera. A sequência de enquadramentos é iniciada com um plano geral e uma série de planos fechados, intercalando entre PC e Mac (frames B, C e D). São esses planos que permitem que observemos a expressão de ambos; PC parece confortável, apresentando sua realidade para o público, enquanto Mac parece perplexo, com as mãos no bolso. Retornamos então para plano geral (frame F), no qual Mac apresenta uma solução mais simples, a escolha de um computador como ele.

A última fase, chamada de “Solução do problema”, tem por objetivo demonstrar a definitiva superioridade de Mac sobre PC, com o desfecho da discussão iniciada na fase anterior, como podemos observar na figura 37, na página precedente, que traz sua representação em *Out of the box* (V-02). Normalmente ela é caracterizada pela inserção de uma proposição por Mac ou PC, que marca uma espécie de irrefutabilidade da superioridade do primeiro, seguindo-se os últimos turnos, e/ou ações físicas que encerram a peça publicitária. No caso de *Out of the box* (V-02), entre os frames A e D, Mac percebe que PC não poderá funcionar diretamente "da caixa" como ele, pedindo para avisá-lo quando estiver pronto. No frame E, PC finaliza a questão, reconhecendo que está desmontado em outras caixas. Como já

dito anteriormente, esse vídeo parece discutir dois padrões de produção de computadores, o sistema monobloco, "all-in-one" da Apple vs. o sistema de licenciamento de peças, como é o caso de PC, no qual o próprio usuário é obrigado a montar o computador.

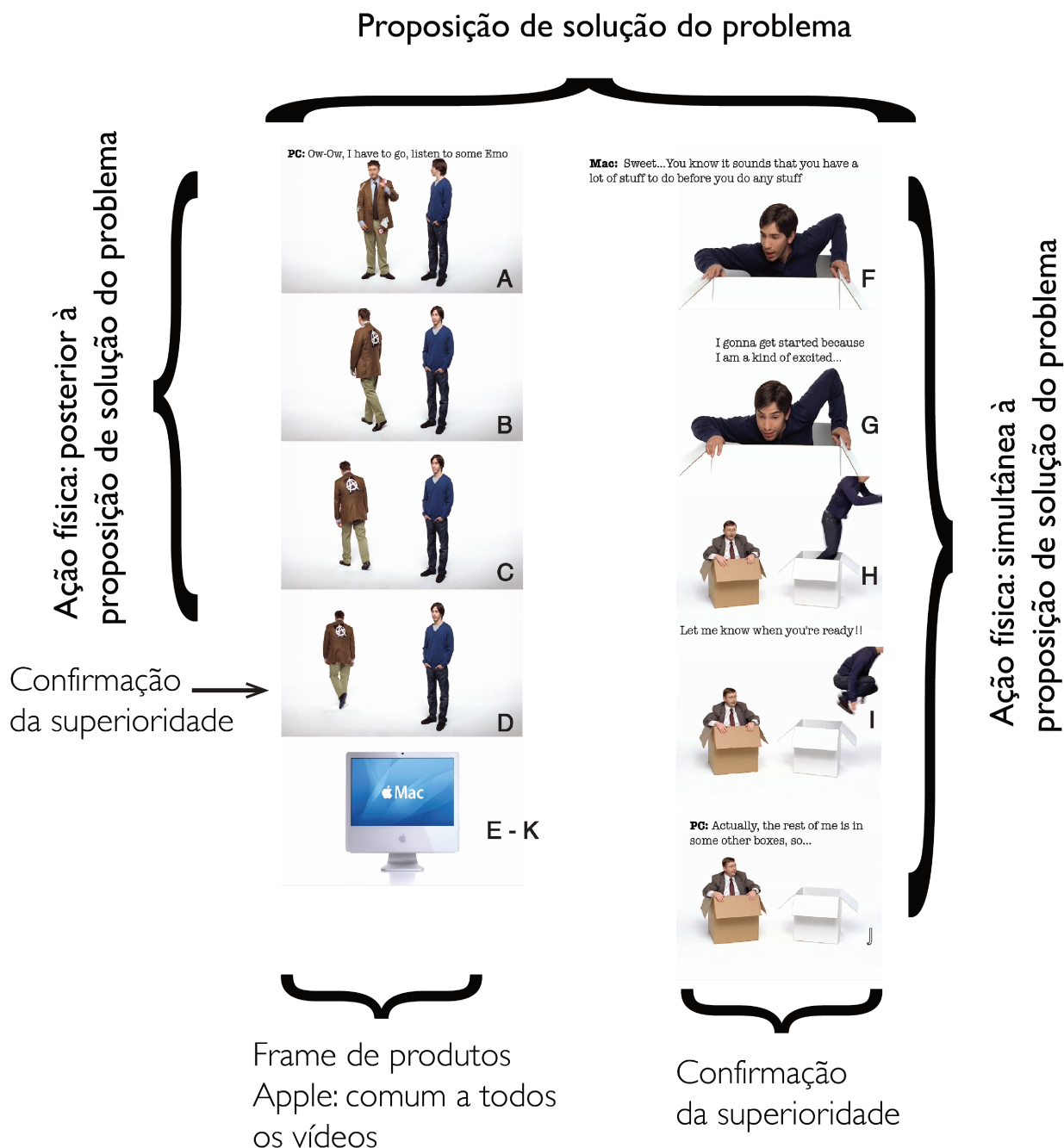


Figura 38: Solução do problema em *Out of the box* (V-02) e *Meant for Work* (V-10)

Esse fechamento em termos gerais é importante por permitir ao texto vídeo-fílmico encerrar o processo de argumentação e facilitação iniciados na primeira fase com uma inequívoca conclusão: Mac é melhor do PC. Igualmente à fase anterior, esta também pode ser dividida em dois passos, que podem ser sequenciais ou simultâneos: um seria uma resposta verbal, chamada de confirmação da superioridade, a outra uma ação física de abandono da interação, ambas realizadas por uma das personagens principais, Mac ou PC.

A figura 38, na página 197, traz a transcrição multimodal dessa fase em dois vídeos que compõem o corpus: *Out of the box* (V-02) e *Meant for Work* (V-10). Podemos observar duas formas de realização: na primeira representada pelos frames de A até D, *Meant for Work* (V-10), a proposição responsável pela introdução da solução do problema é instanciada por PC, que sinaliza o chamado. A essa sequência, instaura-se o silêncio na peça publicitária, inicia-se então um conjunto de ações físicas que marcam a saída de PC do quadro, cabisbaixa e com uma postura derrotista. Já em *Out of the box* (V-02), frames F até J, há uma ação multimodal bem mais complexa: ao mesmo tempo que Mac instancia a proposição responsável por iniciar a última fase do vídeo, ele já inicia ação de saída do quadro, ambas terminando simultaneamente: o salto para fora da caixa — e conseqüentemente para fora do campo de visão do espectador — tem seu ápice no momento em que Mac termina sua fala. Após isso, PC ainda instancia uma nova proposição, reafirmando a superioridade de Mac. Todas as peças publicitárias se encerram com um frame exibindo um dos produtos (hardware somente) da Apple pela época do anúncio, o mesmo em ambos os comerciais e representado pelos frames E – K.

Os dois padrões de realização trazem características importantes a serem pensadas em relação a textos multimodais. A primeira estaria relacionada ao fato de que, nesses textos, um estágio realizado linguisticamente pode desencadear outro que seja instanciado por outro modo de linguagem, como seria o caso da saída de PC no primeiro vídeo. Essa observação é relevante por mostrar que as diferentes semioses são usadas em conjunto na construção da significação, sem que necessariamente haja um processo hierárquico entre elas: diferentes

modos de linguagem interagem no processo de construção de textos multimodais e multimidiáticos, levando à conclusão de que a simples análise do componente linguístico pode causar a inobservância de padrões de linguagem significativos para análise. Fica claro que ações linguísticas e não linguísticas possuem valores similares no processo de instanciação dos significados presentes no vídeo; eles parecem agir complementarmente reforçando os resultados já obtidos na seção 4.3, na página 149.

A segunda estaria relacionada à possibilidade de simultaneidade na realização entre passos dentro de uma mesma fase. Como os vídeos são um texto multimodal por natureza, nada parece impedir que diferentes ações de linguagem sejam realizadas de forma simultânea, sem que haja uma linearidade temporal das ações dentro desses textos. Em outras palavras, o fato de a percepção da superioridade de Mac desencadear uma ação física de retirada não impede que essa ação ocorra ao mesmo tempo que a referida ação oral: a retirada se inicia no momento da percepção dessa diferença, sem esperar que a sequência de turnos que caracterizam essa percepção se encerre. Isso é importante por mostrar que uma realização multimodal não pressupõe uma linearidade textual; mesmo que um texto ocorra no tempo e espaço — como é o caso dos textos videográficos —, as fases ou estágios não precisam ser apenas sequenciais.

5.3 Escolhas e manutenção do significado em diferentes ambientes de produção e veiculação

Um elemento importante a ser considerado em “Get a Mac” é um processo de produção que visa a alcançar um espectador capaz de decodificar as citações cinematográficas e metáforas facilitadoras, como já discutido anteriormente nas seções 5.1, Constituição do discurso: tradução e metáfora, na página 164, e 5.2, Organização e padrões em “Get a Mac”, na página 179. Os comerciais buscam a construção dessas referências por meio da criação de

estereótipos que sejam visualmente conhecidos pelos seus espectadores em potencial, como podemos ver pelas escolhas das personagens.

Todavia, há de se pensar que a Apple, uma companhia multinacional, distribuiu esses comerciais para públicos que não o estadunidense. De fato essa distribuição ocorreu sem que houvesse modificações nos comerciais: a versão original em inglês foi exibida em quase todos os países de língua inglesa e versões dubladas foram disponibilizadas em alguns sítios locais, como foi o caso do Brasil. Em apenas duas situações a Apple modificou a produção de seus comerciais de forma substancial. Entre 2006 e 2008 foram produzidos comerciais não apenas em vídeo, mas também *banners* animados que eram disponibilizados em sítios da área de TI, além de duas campanhas em vídeo específicas, uma para o Japão e outra para o Reino Unido. Os banners fazem uso dos mesmos atores, com diferentes textos e adaptados para interação com o hipertexto, ao passo que as campanhas do Reino Unido e do Japão utilizam atores locais, que não têm inglês americano como variante linguística.

A Figura 39, na página 200, traz uma comparação com frames de peças publicitárias das três campanhas, em 39a vê-se a campanha estadunidense, em 39b há a campanha japonesa e em 39c vê-se a campanha britânica. Como podemos observar boa parte do design presente na campanha dos EUA se mantém comum às demais: há a mesma citação de *THX 1138* e *Matrix* na composição; Mac e PC são representados por pares de atores, dispostos na mesma posição; e mantém-se o contraste moderno X corporativo da campanha original.

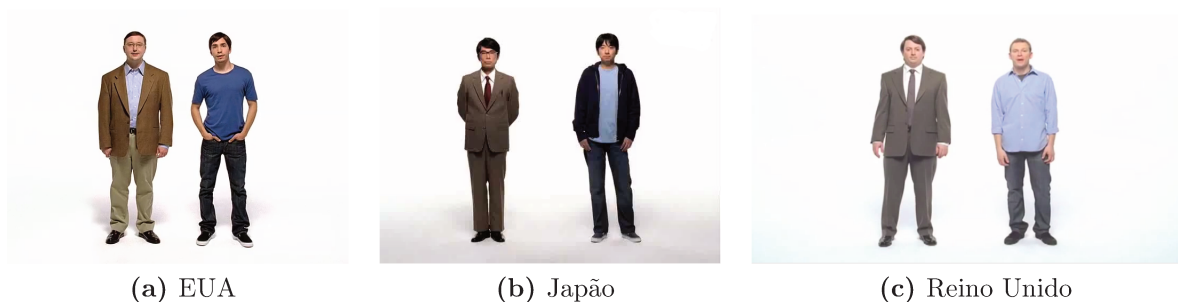


Figura 39: Comparação entre os quadros: EUA, Japão e Reino Unido



Figura 40: Office vs. Home — campanha Reino Unido

Diferenças estão, em primeiro lugar, nos atores que representam cada uma das plataformas computacionais: enquanto os comerciais dos EUA são estrelados por Justin Long (no papel de Mac) e John Hodgman (no papel de PC), dois atores estadunidenses sem nenhuma relação profissional prévia, as peças japonesas e britânicas são estreladas por dois grupos de comédia já conhecidos da população local: os nipônicos pelos *Rahmens*, um grupo de *Owarai Kombi* — tradicional comédia televisiva japonesa —, composto por Jin Katagiri e Kentaro Kobayashi, e por Mitchell e Webb, um duo de comédia britânico, composto por David Mitchell e Robert Webb, conhecidos por sua participação em uma série de TV chamada *Peep Show*, não exibida no Brasil. Do ponto de vista da produção, essa modificação de atores pode ter como objetivo a adaptação dos comerciais para o público local — tanto em sítios de internet como em canais locais de TV —, que passaria a se reconhecer nessas produções, que possuem atores conhecidos e um biótipo mais próximo do próprio telespectador. Além disso, deve-se levar em conta que a mudança de atores trouxe outra transformação importante para esses comerciais, eles não são falados em inglês americano: a série britânica usa a variedade local, enquanto os comerciais nipônicos são falados em japonês.

É possível observar que as peças publicitárias mantêm diversas semelhanças em termos de design, sendo que os atores e linguagem me parecem adaptações que não afetam o discurso estabelecido para a campanha. Ao mesmo tempo que tanto a campanha britânica como a japonesa possuem diversos vídeos adaptados a partir daqueles exibidos nos EUA, novos foram criados exclusivamente para essa localidade. Iniciarei essa discussão analisando um vídeo exclusivo de cada campanha além-mares para, depois, debruçar-me sobre um vídeo presente nas três campanhas.

Entre os vídeos britânicos, destaca-se *Office vs. Home*, uma peça que tenta mostrar como computadores PC são chatos dentro da ambiência doméstica, ao contrário de Macs. Nesse vídeo — como podemos ver na figura 40, página 201 —, PC tenta convencer Mac que a administração de uma casa deve ser feita como a de um escritório, uma vez que esse seria o problema das famílias, e obviamente dos computadores domésticos. A partir da apresentação,

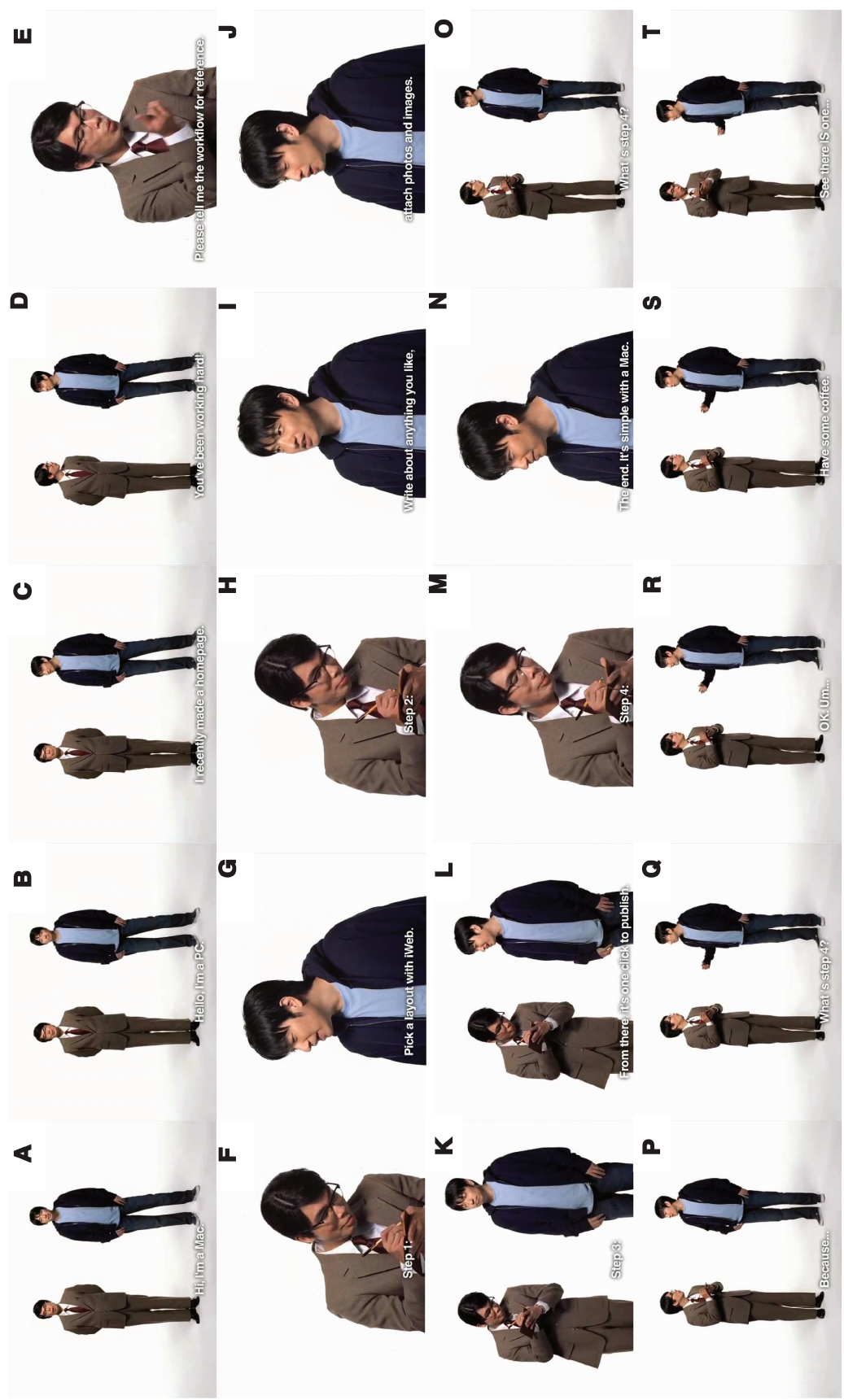


Figura 41: Steps — campanha japonesa

PC inicia um longo discurso, expondo como essa forma de gerenciar a casa pode ser aplicada, como vemos dos frames G até P, e sugerindo um sistema de recompensas e punições muito similares ao mundo corporativo. Apesar de PC expor suas ideias com um sorriso irônico — como podemos observar nos frames I, L, H e G —, Mac parece não dividir a mesma felicidade, uma vez que suas interrupções — como vê-se nos frames D, J e M — demonstram uma clara insatisfação com as ideias que PC defende em seu discurso. O vídeo encerra-se com Mac utilizando ironia — algo que parece típico da cultura britânica e jamais utilizada na campanha estadunidense —, ao dizer que está ansioso para “marcar o ponto do jantar” — o que acontece simultaneamente com o movimento de marcação de ponto feito pela mão esquerda de Mac — algo que PC interpreta literalmente, perguntando se Mac efetivamente aparecerá. Com relação à gesticulação, nesse vídeo, PC gesticula com mais frequência, mostrando-se superexcitado sobre a possibilidade de administrar um espaço doméstico de forma similar à de um escritório, enquanto Mac se faz constantemente desinteressado, com as mãos no bolso.

Nesse texto vídeo-filmico, diferente de todos os da campanha dos EUA, temos a construção do argumento baseada não na questão apenas tecnológica — que é mantida graças ao fato do estabelecimento da metáfora Mac e PC inicial —, mas também em uma discussão tecnofilosófica de qual o melhor modelo de computador deve ser utilizado: o corporativo ou o não corporativo. Um elemento importante que não pode ser deixado de ser comentado é que, tanto nos vídeos britânicos como nos nipônicos, as metáforas que se referiam a Bill Gates (PC) e Steve Jobs (Mac) se perdem em favor da ideia corporativo vs. criativo. A perda acontece principalmente pelo fato de os atores não mais representarem uma referência direta aos dois CEOs, mas sim a tipos executivos e *yuppies* oriundos de cada cultura, também com o objetivo de maior identificação entre o produto e o público específico daquele país.

Diferenças também são encontradas nos filmes feitos exclusivamente para a frente japonesa da campanha, como podemos observar na figura 41, na página anterior, que representa o filme *Steps*, cujo objetivo é discutir a facilidade com que programas Apple — no caso específico iWeb, parte do pacote multimídia iLife, que concentra aplicativos para produção amadora e

doméstica — constroem páginas de internet. O vídeo se inicia com a apresentação de Mac e PC; o primeiro, então, comenta que acabou de criar sua página na internet; o que faz com que PC pergunte qual o *workflow* para construção utilizado por ele (frames A até E). Segue-se então um constante questionamento de PC, que saca um bloco de notas e começa a anotar os passos ditados por Mac de forma sistemática e atenciosa. Mac, por sua vez, está descontraído narrando o processo de produção da página com naturalidade. O empasse se inicia quando Mac se recusa a citar o passo número 4 (frames M e N), encerrando a exposição:

(17) “The end. It is simple with a Mac”.

Isso desencadeia uma fala nervosa de PC, como se a facilidade dos computadores concorrentes não fosse por ele compreensível; Mac eventualmente necessita criar um quarto passo para o *workflow* (frames P e S), mesmo que ele não traga nada de construtivo para a criação da página em discussão.

Steps discute não apenas questões de usabilidade — ao tentar demonstrar que softwares produzidos pela Apple são mais fáceis de serem utilizados na realização de tarefas multimídia —, ele também discute uma metodologia de trabalho que simbolicamente acompanha cada uma das plataformas — PC é sistemático e ligado a um método de produção mais formal, sem necessariamente refletir sobre ele, ao contrário de Mac, que é livre e criativo —, uma vez que PCs estão relacionados ao mundo corporativo e Mac à área de criação. Além de trabalhar com esse estereótipo do executivo japonês como sistemático e supercentrado, frames F, H, K, L e M, ele traz algumas informações visuais importantes. A primeira é relativa ao olhar, PC e Mac interagem sem que nenhum deles encare diretamente o espectador, isso ocorre apenas na apresentação, outra diz respeito ao alinhamento de PC, que se mostra mais alinhado do que seus análogos nas campanhas dos EUA e do Reino Unido, no último, em especial, temos a personagem PC composto de forma pastiche e cheio de caras e bocas (figura 41, página 203). Efetivamente, esse personagem é o único PC das três campanhas que não mostra sinais de sobrepeso. Se pensarmos na forma como o enredo dos vídeos britânico e japonês são construídos, perceberemos que o aspecto tecnológico, tão presente nos vídeos estadunidenses,

é bem menos importante nessas peças recém estudadas. Na verdade, apesar da metáfora inicial ser mantida, o que há é a instituição de um humor que parece estar relacionado a elementos que pertencem à cultura que pretendem atingir, o que parece explicar o olhar de PC e Mac (JP), o caderno de notas e a concentração de PC (Japão) ou a postura bonachona de PC (Reino Unido).

Surge então uma questão importante a ser resolvida. Será que os vídeos adaptados mantêm todos os padrões da campanha dos EUA?

A figura 42, na página 208, traz uma comparação visual entre o vídeo *Fat HD* (V-11), estadunidense, que foi regravado tanto na campanha britânica, na qual é chamado *Stuffed*, como na campanha nipônica, na qual é chamado de *Bloated*. A composição visual e o conteúdo dos vídeos — todos versando sobre o fato dos HDs estarem lotados — parecem deixar claro que as obras nipônica e britânica são refazimentos a partir do americano *Fat HD* (V-11). Todavia, há uma diferença na natureza da causa desse problema: tanto no vídeo estadunidense como no britânico o HD está superlotado graças à presença de trial softwares que o deixam, por conseguinte, lento e pouco operante; enquanto a lentidão e superlotação do HD no vídeo nipônico é resultado do *upgrade* do sistema operativo, que ocupa maior espaço. Em um primeiro momento, pode-se inferir que essa adaptação talvez esteja relacionada ao valor que a comunidade-alvo dê para essa questão, ou mesmo à oportunidade dos produtores de veicularem uma crítica diferente da que já haviam realizado.

Como também podemos observar na figura 42, na página 208, não parece haver uma diferença muito significativa dos roteiros das versões britânica e estadunidense, apenas adaptações idiomáticas. Já com relação à campanha japonesa, percebe-se, além da adaptação temática já mencionada, que ela é a única das três que não menciona produtos Apple como forma de solucionar o problema. Nesse sentido, a construção discursiva dessa peça talvez seja uma das únicas peças publicitárias da campanha que não parece se firmar na superioridade de Mac, ao contrário, ela se firma apenas em mostrar uma deficiência do sistema operacional Windows. Na verdade, o que se pode também notar é que em *Fat HD* (V-11) as diferenças discursivas não incluem domínios totalmente distintos: tanto nas campanhas dos EUA e do

Reino Unido como na do Japão, a questão envolvida é a superlotação do HD por motivos que são alheios à vontade do usuário e que, numa certa medida, estão conectados a deficiências da plataforma. Essas diferenças, no caso dos vídeos estudados aqui, são fruto da interação entre as culturas dos três países. Logo é interessante observar que essas diferenças na organização do discurso parecem não ter influência profunda na organização visual do discurso dos comerciais; elas parecem apenas adequá-lo às especificidades de recepção de cada público.

Outras diferenças importantes são guardadas em outros modos, como a postura corporal de Mac, o tom utilizado por PC em sua fala. Com relação ao primeiro, percebemos que Mac tem uma relação completamente distante de PC no vídeo destinado ao público dos EUA: ela passa o tempo todo com a mão no bolso sem expressar nenhum tipo de solidariedade.

Como citado logo acima, há também algumas diferenças entre a forma como a personagem PC é construída em cada uma das campanhas. Apesar de se mostrar cansado, no comercial estadunidense, sua postura é distante de Mac, tanto seu tom de voz como suas ações físicas não parecem demonstrar que ele queira algum tipo de ajuda, ou que esteja sofrendo algo além do cansaço. Já o nipônico assume um papel intermediário, apesar de seu pedido de espera inicial ter um tom mais pessoal, quase desesperado, ele realiza alguns movimentos corporais que expressam seu cansaço, mas nada além de um “dar de ombros” e um suspiro no meio do vídeo. Já PC britânico, por sua vez, é o que mais parece sofrer os traumas dessa exposição: além de suspirar e dar com os ombros, ele fica cabisbaixo, utiliza um lenço para enxugar o suor de seu cansaço e faz movimentos muito mais entristecidos ao perceber que deverá voltar para onde estava antes de iniciar o comercial. Essas diferenças parecem resultar de diferentes concepções técnicas — referentes ao motivo que causa a superlotação do HD — e atitudinais — em uma certa medida, PC representa a reação de seus usuários ao problema —, que também representam estereótipos: no primeiro temos uma perspectiva técnica; na segunda, um misto entre técnica e emocional, ao passo que na última a reação é quase totalmente emocional.



Figura 42: Comparação entre 3 vídeos — Japão/Reino Unido/EUA

Algumas das diferenças apontadas nos vídeos produzidos exclusivamente para o Japão e para o Reino Unido desaparecem nos vídeos comuns as três campanhas, entre elas estão o humor diferenciado, a relação que Mac e PC estabelecem com a câmera, assim como a importância da questão tecnológica, que é utilizada nesses *remakes* com a mesma intensidade do comercial análogo dos EUA. Efetivamente, isso mostra que, mesmo que o discurso se mantenha com algumas modificações visuais e temáticas — as quais possuem o objetivo de adaptar essas peças tanto visualmente, como tematicamente para a nova praça de exibição —, mantém-se uma identidade com o modelo de produção da matriz.

5.4 Resultados parciais

Este capítulo fez uma análise de como esse modelo pode ser aplicado no processo de construção discursiva dos vídeos pertencentes a campanha "Get a Mac", estudando-se algumas de suas estratégias de construção do significado. Buscou-se observar como os vídeos são organizados em termos dos seus conteúdos discursivos, de sua organização macrotextual além de refletir sobre como a presença da campanha em diferentes culturas poderia modificá-la.

Na seção 5.1, discuti como elementos construtores de significado poderiam ajudar a constituir o discurso da campanha publicitária "Get a Mac". Esse discurso se fia na tradução de conceitos computacionais por meio de metáforas que estão estruturadas tanto no modo de linguagem verbal como no visual. Essa metáfora tem por objetivo facilitar a comparação entre os produtos Apple e seus concorrentes: a ideia parece ser traduzir o discurso técnico da área de computação para um universo mais próximo dos possíveis espectadores da campanha; ao construir esse discurso baseado em metáforas, os autores estão humanizando as questões que poderiam ser de difícil compreensão. Esse processo se constitui pela tradução do domínio técnico — representado pelas especificidades de cada um dos vídeos presentes no corpus — para o domínio cotidiano dos usuários de computador, alvo da campanha, de forma a possibilitar a sedução necessária para a compra do produto. Fazer com que o espectador entenda

questões computacionais com simplicidade é uma estratégia de venda da marca: imagina-se que os computadores Apple sejam tão simples de operar e tão imunes a problemas como mostrariam as relações entre as personagens no vídeo. Logicamente, a plataforma PC se caracterizaria por ter as qualidades contrárias.

Além disso, a criação de estereótipos leva a uma tentativa de identificação entre o espectador e o produto humanizadamente retratado. A representação veiculada nos comerciais tem o objetivo de retratar os produtos Apple como mais contemporâneos e menos tediosos do que os da plataforma PC, ali ilustrados como computadores corporativos e cheios de defeitos. O viés maniqueísta e manipulador dos comerciais se torna claro, as possíveis questões da plataforma Mac são ignorados a favor de uma construção discursiva que procura enaltecer suas qualidades técnicas e comerciais. Todavia, tais qualidades não são expostas a partir de sua competência, mas sim da incapacidade de PC de lidar com determinadas situações; em termos gerais Mac parece ter suas características expostas por meio de uma estratégia matemático-filosófica chamada de prova por absurdo: tudo que parecemos saber é aquilo que ele não é. Em outras palavras, Mac parece ser caracterizado por ser um "não PC", estabelecendo-se, assim, uma relação dialética; por ser um "não Mac" Pc está sujeito a toda forma de problemas. Na perspectiva adotada nos comerciais, o mundo computacional parece se resumir apenas na relação quase que totalmente antagônica entre essas duas plataformas.

A análise da construção visual em 5.2, na página 179, mostrou que essas escolhas também não são aleatórias, sendo motivadas ora pela necessidade de criar um espaço de atuação que permitisse o tipo de interação necessário para a instanciamento do discurso, ora pelo estabelecimento de intertextualidade em pelo menos duas situações. Na primeira, vemos claramente que a empresa busca criar um laço entre suas peças e obras cinematográficas consideradas importantes no universo da ficção científica, em especial na estética *Cyberpunk*.⁶⁹ Mas por que um comercial de computadores, que prima pelo incentivo do uso da tecnologia, citaria filmes que se caracterizam exatamente por uma análise niilista? Em primeiro lugar, é importante

⁶⁹Ver nota de rodapé 65 na página 181.

observar que esses filmes são considerados clássicos para os admiradores do gênero que, normalmente, circulam pela área de TI; logo ao associar seu produto a tais obras, a Apple está tentando relacionar-se de forma estética com seus compradores em potencial. Incentiva-se a compra dos produtos da marca não apenas por sua qualidade ou diferenciação dos demais, mas também por seu parentesco com obras cinematográficas consideradas especiais para esses mesmos compradores: nesse sentido, os produtos Macintosh parecem herdar as qualidades que esses filmes têm.

Aqui também discuti como esses vídeos são estruturados enquanto textos. Como já colocado na seção 1.2.2.2, na página 60, optou-se conscientemente por procurar as regularidades constituintes desses vídeos sem, contudo, discuti-los a partir de uma teoria dos gêneros do discurso. Isso se deu, principalmente, pelo fato de textos do domínio publicitário possuírem uma variação significativa de sua estrutura textual e discursiva: a pressão por inovação e pela diferenciação entre as peças cria obras audiovisuais bastante diferentes entre si. Com relação a estes resultados, observei que os vídeos parecem se dividir em três fases, todas com regularidades verbais e visuais. Entretanto, surgem algumas quebras, todas motivadas pela constituição discursiva dos vídeos em que ocorrem. É importante ressaltar que essas quebras não acontecem em conjunto, quebras visuais não estão atreladas a quebras verbais e vice-versa.

Finalmente, na seção 5.3, comparou-se a campanha dos EUA, centro de análise desta tese, com duas campanhas derivadas: uma no Reino Unido e outra no Japão, com o objetivo de verificar se haveria diferenças temáticas ou estruturais na adaptação dessa campanha para outras culturas, e entre elas estariam:

Troca de atores: ao invés de usar os mesmos atores, optou-se por utilizar dois duos de comédia já conhecidos da população local em ambos os países.

Vídeos específicos: tanto a campanha do Japão como a do Reino Unido possuem alguns vídeos exclusivos que não estavam na campanha dos EUA.

Adaptação de enredo: mesmo nos vídeos que são comuns as três campanhas, pode haver algumas adaptações no enredo do filme, como é o caso de *Fat HD* (V-11), que sofreu modificações no texto ao ser produzido no Japão.

Transformações nos modos de linguagem: mesmo que os comerciais sejam idênticos nas três campanhas, a interpretação dos atores e sua constituição física mudam.

Todas essas modificações, como mostrou a análise, parecem motivadas por questões culturais específicas de cada um dos contextos das novas campanhas, que pareceram buscar uma maior identificação entre o público local e os produtos Apple. Naturalmente, elas transformam a tônica geral das campanhas do Japão e do Reino Unido, sem, contudo, fazer com que elas perdessem sua identidade com a produção original. Essa identidade se deu, principalmente, pela manutenção de alguns elementos-chave da campanha, entre eles:

- Identidade visual: a fundo infinito minimalista não foi modificado nas campanhas japonesa e britânica.
- Identidade discursiva: a construção dicotômica entre Mac e PC se mantém nas campanhas, Mac sempre termina os vídeos com superioridade em relação a PC.
- Estereótipos: apesar da dicotomia Bill Gates vs. Steve Jobs ter sido perdida, a questão corporativo vs. criativo ainda é mantida, mas agora adaptada às realidades locais.
- Estrutura do vídeo: as fases e a estrutura de constituição dos turnos, Mac e PC se alternando na discussão do seu problema, se mantém nos vídeos das 3 campanhas.

Esses elementos de construção da identidade permitem que espectadores reconheçam os vídeos e percebam que, apesar das diferenças apontadas, todas fazem parte da mesma campanha. Isso é importante porque a deixa reconhecível para o seu público: não só os produtos Apple são os mesmos em todos os lugares do mundo, mas as campanhas também, mostrando que talvez haja uma similaridade entre o discurso de todas as subsidiárias.

Coesão e fluxo de informações no vídeo

In this post-modern age when
boundaries tend to become increasingly
blurred, spurred by intermixing of
cultures and the increasingly unfettered
flow of information [...]

R. Martinec & A. Salway

Introdução

Aqui discutirei algumas questões relacionadas à construção do processo de coesão e fluxo de informações em textos vídeo-fílmicos. Como já dito anteriormente, textos audiovisuais possuem uma característica similar aos verbais, eles ocorrem no tempo/espaço: ao vermos o trecho de um filme ou um vídeo, a sequência imediatamente anterior já não é mais recuperável, a não ser que estejamos em um equipamento reprodutor, como um DVD ou Blu-Ray. Programas de televisão e filmes de cinema, por exemplo, só podem ser assistidos novamente se adentrarmos a uma nova sessão ou assistirmos a uma reprise.

Devido a esta característica, observo que cabe, de certa forma, ao produtor dos vídeos construir um sistema de referências internas que permita conectar aquilo que já foi exibido

com aquilo que está aparecendo naquele exato momento na tela. Na verdade, uma análise mais próxima desses textos mostra que existem estratégias que ajudam esse leitor a construir esses processos: a estrutura do texto audiovisual também "amarra" suas proposições de forma a oferecer elementos que permitam ao espectador realizar essas conexões. Observou-se que existem, pelo menos, dois tipos de redundâncias presentes nesses vídeos, uma natural e relacionada à repetição dos frames em cada sequência de imagens e outra, que traz à cena elementos presentes em frames e sequências anteriores. Aqui também discutiram-se os processos de preenchimento de vazios em frames e sequências, o que parece ser possível graças aos processos de repetição presentes no texto.

Parto, então, para a seção 6.2, Fluxo de informação e construção do significado, na página 224, que se divide em duas. Na primeira (seção 6.2.1, Componentes estruturais, na página 224) discuto como os componentes estruturais podem estar relacionados com o processo de sequencialização da informação. De fato, observou-se a existência de alguns padrões que podem ser considerados significativos, em especial no que tange ao uso de enquadramentos. A seguir (seção 6.2.2, Componentes atitudinais, página 234) realizo essa mesma análise, mas agora com os componentes atitudinais. Resultados também mostram padrões de uso em, pelo menos, dois tipos de ações: apontar e olhar.

Seguem-se, então, os Resultados parciais do capítulo, na página 242.

6.1 Repetição e preenchimento de vazios

Nesta seção discutirei o processo de redundância e preenchimento de vazios nos textos audiovisuais que compõem o corpus de pesquisa. Como já foi discutido na seção 3.2, Noções básicas de linguagem vídeo-filmica, na página 118, a imagem videográfica se constitui essencialmente da projeção de fotografias em sequência, sendo que é o acúmulo dessas fotografias no espaço-tempo que nos faz ter a ilusão de movimento. Cada segundo dos vídeos analisados nesta tese possui 29.97 frames (ou fotografias) por segundo (FPS) em sua composição, em

uma conta rápida, somariam-se 899.1 frames para cada vídeo com 30 segundos de duração. Isso faz da constituição de qualquer peça videográfica, algo relativamente redundante: a cada frame há apenas um pequeno movimento, quase imperceptível em sua unidade, mas que se torna significativo apenas no acúmulo de imagens que formam um segundo.

Ao observar a figura 43, página 215, que traz a instanciação de um gesto por PC em *Surgery* (V-08), percebe-se que as redundâncias são muitas: enquanto PC movimenta suas mãos, todos os demais elementos do quadro não se modificam. Há um acúmulo de elementos visuais repetidos, enquanto uma parte pequena do quadro se movimenta (ver figura 43). Isso permite não apenas que o efeito da ilusão de movimento se estabeleça, mas também a existência de um elemento determinante no processo de compreensão desses textos. É aqui que se estabelece uma primeira instância de estabelecimento da tecitura, da mesma forma que um texto verbal progride a partir de uma base estabelecida pela redundância de conceitos constantemente retomados por escolhas lexicais, a imagem em movimento passa pela redundância constantemente transformada: são pequenas modificações que ajudam a manter um lastro de percepção naquilo que já passou, ao mesmo tempo que permitem a progressão desses textos. Da mesma forma que uma palavra possui colocados que formam o seu cotexto imediato, cada frame possui imagens que o cercam, construindo uma espécie de coimagem.⁷⁰ Cada segundo seria, então, o resultado de uma combinatória, sendo impossível interpretar um frame em isolamento. Esse tipo de estratégia é, na verdade, um resultado técnico das próprias características desse tipo de texto multimodal.

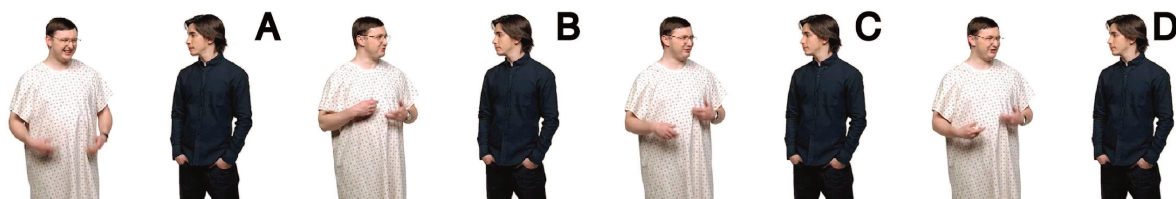


Figura 43: Repetição em *key-frames* em *Surgery* (V-08)

⁷⁰Gostaria de agradecer a José Luis Meurer (*in memoriam*) pelas proficuas discussões sobre esse aspecto.

Esse tipo de repetição é essencial para a construção da base articulatória dos textos audiovisuais, sem essa repetição não haveria a formação de imagens contínuas. Essas imagens formam um alicerce imagético que nos permite construir as bases para interpretação de quaisquer outros modos de linguagem que estejam presentes em tais textos. Apesar da independência de alguns desses modos de linguagem — como é o caso da linguagem verbal, presente no mundo externo e alheio ao vídeo —, ao se construir o amálgama de significação, como discutido nas seções 4.2, página 138, e 4.3, página 149, a interpretação do significado é dependente de uma percepção simultânea desses modos. O componente visual acaba ganhando um destaque importante na produção audiovisual uma vez que ela parece ser o suporte para que uma série de outros modos de linguagem possam ocorrer: vestimenta, cores, gestos, olhares, entre muitos outros parecem só poder existir nesses textos graças à existência desse componente. Nesse sentido, a estabilização da imagem em forma de redundância pelo acúmulo dos frames seria o elemento abalizador que permitiria a construção da mensagem.

Sugiro aqui que esse terreno estabelecido para o espectador é importante para que se possa construir as modificações responsáveis pela sequencialização da imagem, que se daria no fluxo de informações (ver seção 6.2, página 224). Existem algumas raras experiências que tentam quebrar essa lógica estruturada na repetição, a técnica é chamada de *single frame* (quadro único) e consiste em não permitir que o número de frames necessários para a constituição de um segundo ocorra. Os trabalhos que utilizam dessa técnica tendem a ser de caráter experimental e puramente artístico, um bom exemplo seria a obra de Paul Sharits e Eric Andersen que participaram ativamente do *Grupo Fluxus*, um importante conjunto de artistas de vídeo que revolucionou o cenário estadunidense durante a contracultura e trabalhou com críticas à mercantilização da mulher, ao consumo exagerado, ao conservadorismo social e político. No contexto do uso da técnica *single frame*, alguns trabalhos de Sharits e Andersen merecem destaque, do primeiro gostaria de citar *Word Movie* (de 1966) e do segundo *Opus 74 Version 2* (de 1966).

Em *Opus 74 Version 2*, Andersen faz uma crítica ao consumo exacerbado e descartabilidade de bens. Contudo, ao invés de nos permitir ver uma imagem por segundo, o que atribuiria 24 frames a cada uma delas,⁷¹ ele substitui cada um dos frames que comporiam um segundo por uma imagem diferente. O efeito resultante é de uma extrema confusão visual, mal formamos uma imagem em nossas retinas já temos uma outra projetada. Por conta disso, há uma imensa dificuldade em registrar em nossa retina, assim como em nossa memória, quais são os objetivos ali apresentados. Já o trabalho de Sharits se caracteriza pela mesma técnica, mas com palavras compondo cada um dos frames, sendo que, de um frame para outro, há a manutenção de apenas uma letra. Ao mesmo tempo que essa vertiginosa sucessão de palavras é projetada, ouve-se um locutor lendo-as a partir de um roteiro. O resultado é que não conseguimos associar o que é dito ao que aparece na tela: a velocidade é tão grande que não retemos informação visual alguma, apesar do vídeo se realizar em *looping*.⁷² A dificuldade em entender essas obras advém do fato de elas não possuírem referências cumulativas no tempo vs. espaço que nos permitam estabelecer um grau mínimo de redundância, necessária para sua leitura nos moldes tradicionais. Ao justapor as imagens dessa forma, não há a construção de uma tecitura baseada na redundância: as coimagens não são mais o suporte para a compreensão e avanço da discussão central do vídeo. A variedade de imagens acaba sempre sendo tão grande que é muito difícil representá-la em um frame estático: ou teríamos uma fotografia isolada ou uma mancha de transição, razão pela qual esses vídeos não foram representados em figuras nesta tese.

Esses exemplos são relevantes por nos mostrar que esse constituinte de repetição é algo que não pode ser ignorado em obras vídeo-fílmicas que possuam um propósito de construção baseado na mimetização do mundo real, mesmo que ficcional ou metafórica, como é o caso da campanha "Get a Mac". Isso porque é ela que permite a construção de um processo de

⁷¹Gostaria de lembrar que o vídeo analógico é composto por 24 quadros por segundo e não 29.97, como no digital.

⁷²Um *looping* pode ser definido como um trecho de vídeo ou áudio que se repete infinitamente, sem um começo ou um fim marcados.

coesão e coerência capaz de criar a narrativa baseada na repetição, retomada e progressão, aos moldes da linguagem verbal.

Esse processo de repetição é também responsável pela manutenção da coerência temática dentro dos vídeos, ela garante que ele seja um texto coerente e com uma narrativa estruturada em uma base de significados sólida e que está presente durante toda a peça comercial. Como é possível observar na figura 43, página 215, que representa os *key-frames*⁷³ de um segundo em *Surgery* (V-08), PC inicia seu movimento no frame A, levantando as mãos e levando-as à altura do peito. No frame B, esse movimento já está definido, e ele passa a girá-las uma de frente para a outra, um gesto comum para significação de movimento e troca; de fato a iconicidade deste gesto parece estar relacionada ao próprio tópico do vídeo, que é a atualização lógica (programas) e física (hardware). Essa mesma significação de atualização parece ser o elemento motivador de uma série de metáforas presentes no vídeo, entre elas a roupa usada por PC, típica de hospitais, sugere uma possível cirurgia, que é confirmada pela própria fala de PC, como podemos ler na transcrição a seguir.

- (18) (a) **PC**: Well, I' upgrading to Vista today, what is great, but a get along nervous when they mess with me inside. . .
- (b) **Mac**: What you mean? Is not just straight forward?
- (c) **PC**: No, Not really. Like a lot of PCs, I have to update my graphics card, my memory. If I want a premium package, I need a faster processor it's major surgery.

Em 18 a, observamos que PC justifica suas roupas graças a uma necessidade de atualização do sistema operacional, o que levaria a uma série de modificações internas, fato que o preocupa. Mac (em 18 b) se espanta e observa que o processo deveria ser simples; já PC retruca (em 18 c), colocando que essa atualização é algo "realmente sério". A instanciación dessa tradução de conceitos computacionais em uma metáfora verbo-visual⁷⁴ pode ser interpretada como o resultado da construção da peça audiovisual em torno de um único assunto,

⁷³Em vídeo, cinema ou animação *key-frame* é um frame que define pontos de início e fim de um movimento, sendo que, por vezes, eles também representam momentos intermediários. Sua função é entender o ponto de saída e chegada de um movimento, eliminando as repetições de uma transcrição frame a frame.

⁷⁴Ver seção 5.1, Constituição do discurso: tradução e metáfora, página 164, para uma discussão mais profunda desses conceitos.

ou ponto a ser explicado. Se esse ponto a ser explicado ou discutido influencia as escolhas em todos os modos de linguagem presentes na obra — e o trecho analisado de *Surgery* (V-08), exemplificado visualmente na figura 43, página 215, mostra exatamente isso: a roupa, a fala, os gestos são todos influenciados por ele —, não se poderia pensá-lo como um HiperTema?

Em minha opinião, esse HiperTema é um conceito que parece pairar sobre o vídeo, condicionando as escolhas significativas e as tornando, até certo ponto, redundantes. Essa redundância se manifesta em elementos diferentes, faz com que a mesma ideia esteja presente em diversas formas de representação, criando um processo de tecitura multimodal que se manifesta em todos os modos de linguagem e, conseqüentemente, em todas as escolhas. Assim, ao selecionar, mesmo que inconscientemente, quais elementos instanciam um determinado significado, tem-se esse HiperTema como elemento abalizador da estruturação dessa significação.

O apoio a essa repetição multimodal estaria na repetição de frames discutida logo acima. Essa construção de base propiciada pelo acúmulo repetitivo das coimagens permite a construção de outra base sólida para a compreensão. A manifestação do HiperTema nos gestos, na roupa e nos olhares de *Surgery* (V-08) são elementos que se repetem por pegar, de alguma forma, uma "carona" nessa repetição natural dos frames: é graças a ela que o pijama está presente em quase todos os frames do vídeo, é graças a ela que as falas duram uma quantidade de frames que nos permita associá-la à imagem e criar o sentido necessário para a compreensão do vídeo.

Na figura 44, página 220, há uma ilustração dos *key-frames* representando os 4 primeiros segundos de *Vista Choose* (V-07). Apesar de possuir uma construção de significados mais simples que *Surgery* (V-08), o mesmo princípio de relação entre o HiperTema e as escolhas em todos os modos de linguagem continuam presentes. *Vista Choose* (V-07) simula um *game show*⁷⁵ com o objetivo de ajudar PC a escolher qual a nova versão do sistema operacional vista ele deve utilizar.

⁷⁵Um *Game Show* seria um tipo de programa televisivo no qual pessoas ou equipes realizam provas de caráter físico ou intelectual com objetivo de ganhar prêmios. Exemplos contemporâneos seriam *Show do Milhão* e *Roletrando*.

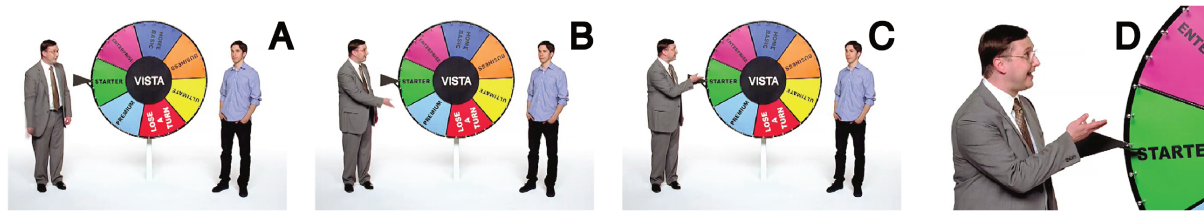


Figura 44: Repetição em *key-frames* em *Vista Choose* (V-07)

A posição central da roleta no frame A (figura 44, página 220) coloca-a em posição de destaque, determinando qual será o foco de todos os olhares durante esse vídeo. Isso se dá porque ela é um elemento imprescindível para realização do jogo, sua posição no quadro a manterá constantemente visível para o espectador, mesmo em planos mais fechados, como também pode-se ver na figura 36, página 194. Pode-se observar que no frame A (figura 44), PC inicia um gesto de apontar que tem por objetivo apresentar a tal roleta, sua mão vai levantando lentamente como podemos observar em B e em C. O gesto culmina com um corte em plano mais fechado, no qual a roleta ainda está presente; ao mesmo tempo que esse gesto e corte acontecem, PC está falando:

(19) (a) **Mac:** Hello, I'm a Mac...

(b) **PC:** And I'm a PC. And it's time to play... Cho:::se a Vista!!!!

Evidentemente, essa apresentação ajuda o espectador a entender a função desse objeto para a coerência local do vídeo; sua presença faz parte de uma ampla metáfora que trata com ironia a existência de tantas versões de um mesmo sistema operacional. Essa temática agregada ao *gameshow* está presente em todo o vídeo, construindo elementos visuais, verbais e gestuais; momento em que todas as escolhas parecem condicionadas por esse HiperTema.

Além desse processo cumulativo de imagens no tempo vs. espaço, existe um outro elemento que não pode ser esquecido. Como se pode observar no frame C da figura 46, na página 223, há um pedaço da manga de PC, identificável graças a sua redundância com os frames A e B, anteriores a ele. Esse tipo de estratégia é comum a diversos vídeos da campanha "Get a Mac". Por exemplo temos um pedaço da roleta presente no frame D da figura 44, página

220, resgatando seu plano geral, um pedaço da manga de PC que sobra no frame B da figura 48, página 228, as pernas cruzadas da analista no frame D e o rosto de PC no frame E da figura 50, página 230, o braço de PC que ainda fica em quadro depois do corte para o frame E, na figura 51, página 231, o braço de Mac, que permanece presente no frame A da figura 25, página 176, e o braço e a mão de PC e o braço de Mac que permanecem em quadro nos frames C, D e F da figura 26, página 177, respectivamente.

A importância desse tipo de redundância está em relacionar os frames, amarrando-os de forma a forçar a relação entre eles para o espectador: uma espécie de lastro cuja função é garantir a relação entre os frames, sem confiar necessariamente na memória do espectador para realizar todas as conexões entre eles. Em minha opinião, esse tipo de lastro parece se relacionar a uma espécie de coesão referencial. Trata-se de um processo muito similar a uma catáfora visual, uma vez que essa repetição mantém constante um elemento do HiperTema que estava presente frames anteriores. Cria-se então uma nova instância da construção da repetição nesses vídeos: uma conexão local entre frames. Essa conexão é importante não apenas por garantir que cada frame ecoe individualmente elementos do HiperTema do vídeo, como também contribui para a construção de uma tecitura local, que se estabelece entre os frames. Claro que a repetição de elementos do HiperTema e a constante justaposição de frames são importantes e também garantem essa coesão, mas a repetição de pequenos elementos — como a roleta, a manga e os braços — garantem uma conexão mais direta entre as coimagens, algo que complementa e amplia a relação já estabelecida pela justaposição natural dos frames.

Por fim, outro fator importante para a construção da tecitura relaciona-se com o processo de preenchimento de vazios deixados pela modificação de enquadramentos. Para que possa discutir esse tipo de construção coesiva, é necessário retomar alguns dos conceitos já discutidos na seção 3.2, Noções básicas de linguagem vídeo-fílmica, na página 118, em especial a discussão em torno da figura 15, página 121. Enquadramentos podem ser encarados como cortes da totalidade da cena, cada vez que a objetiva é fechada, temos uma aproximação da

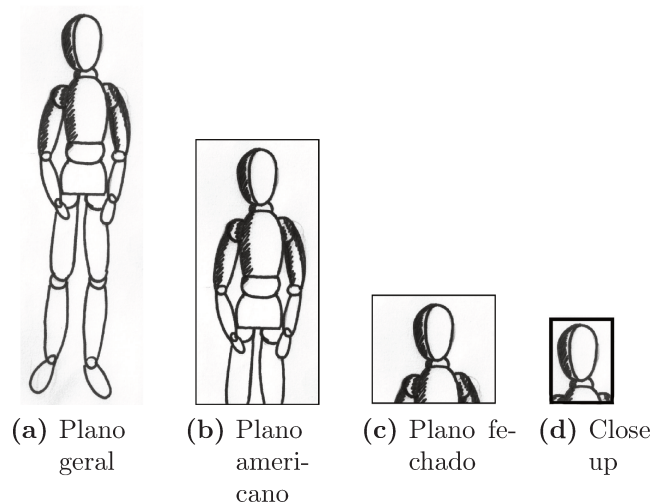


Figura 45: Planos e preenchimento

personagem fotografada. Essa aproximação, necessariamente, deixa de fora outros elementos que poderiam estar em cena, um exemplo disso é trazido pela figura 45, página 222. Ali percebemos que à medida que a objetiva é fechada, os planos mostram menos partes da modelo e de seu contexto imediato. A figura 45 ajuda a perceber que existem diferentes graus nesse tipo de corte. O plano representado por 45a, que instancia um plano geral, realiza um tipo de recorte de realidade diferente de 45b, que se mostra mais fechado e, portanto, mais restritivo. O mesmo pode ser observado em relação a 45b, uma vez que o plano americano necessariamente elimina uma quantidade de elementos menor que o plano fechado, retratado em 45c, ou ainda mais que o *close up*, que pode ser visto em 45d.

Uma questão importante que a figura 45, página 222, traz à baila é como o espectador, ao ver o enquadramento presente em 45d, é capaz de saber que ele se trata da mesma modelo presente na figura 45a? Em uma fotografia, por exemplo, a única possibilidade seria reconhecer a face da pessoa ali representada, claro que isso em uma imagem real e não em um desenho conceitual como o que utilizei. Todavia, os demais elementos do contexto estariam perdidos em definitivo: muitas vezes, mas não necessariamente, não se sabe a totalidade do cenário em uma fotografia, muitas informações não são recuperáveis. No vídeo, muitas

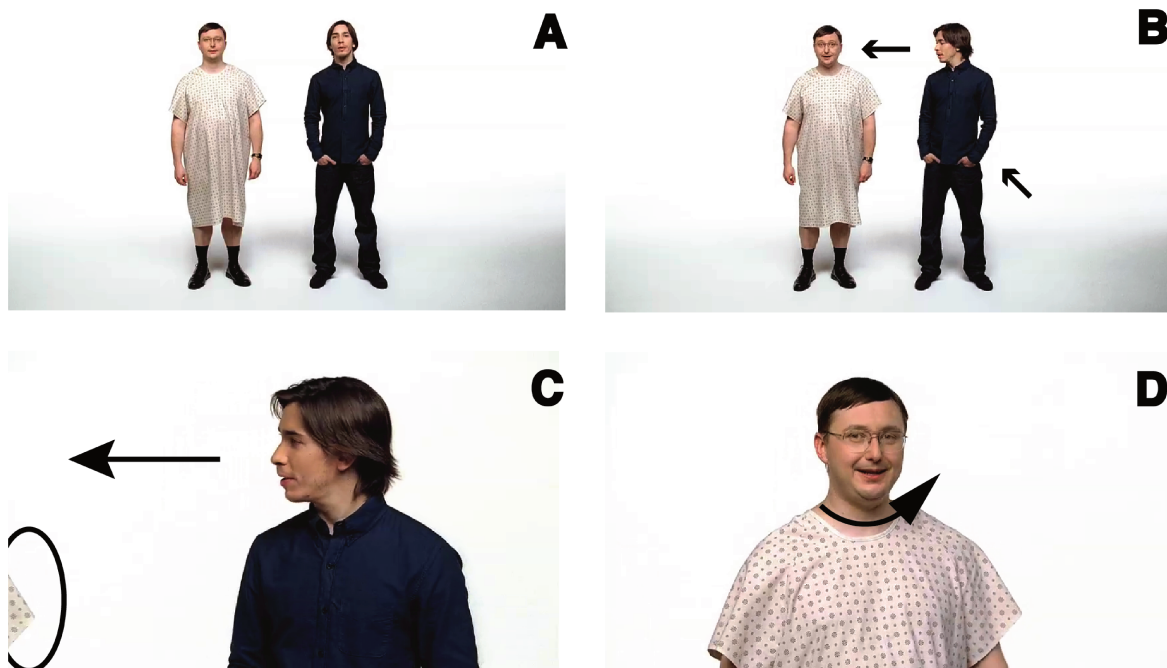


Figura 46: Planos para preenchimento em *Surgery* (V-08)

vezes, a realidade pode ser diferente, isso porque um enquadramento mais fechado pode ser antecedido por outro mais aberto. Nesse sentido, a troca de enquadramentos garante que o espectador ainda se lembre, com maior ou menor detalhe, dependendo, claro, da atenção dispensada ao vídeo, da situação retratada no enquadramento anterior.

Uma das consequências desse tipo de realidade está no fato de que a interação ocorrida na imagem em movimento pode pressupor essa lembrança. Isso se dá em especial quando as personagens interagem com elementos ou outras personagens que estão fora de quadro. Do ponto de vista puramente técnico, ao gravarmos uma cena como o frame C (figura 46, na página 223), que é antecedido por um plano geral, que vemos nos frames A e B (figura 46), nem sequer precisamos dos dois atores em cena. Ou seja, pela lógica de produção todos os *takes* de plano mais geral devem ser gravados antes daqueles do plano mais fechado.

Em outras palavras, os planos mais fechados podem ser produzidos sem que os elementos de um plano aberto estejam efetivamente presentes. Entretanto, sua justaposição a um plano

mais aberto gera uma transferência direta para nossa memória. Imaginamos que essas ações são efetivamente simultâneas e completamos os vazios visuais que nos são oferecidos. Ao se retomar a figura 46, na página 223, pode-se perceber que o frame D nos é compreensível porque, ao lho vermos, imaginamos que ele esteja sendo instanciado simultaneamente a C, sendo que, o contexto geral e completo seria algo similar ao frame B. Isso é possível graças ao acúmulo de coimagens anteriormente projetadas na tela: relacionamos o frame D com nossa memória do que vimos nos frames C, B e A.

Isso gera uma nova função para a repetição nesse processo de construção significativa. Ela garantiria não apenas a construção de uma tecitura de base para o desenvolvimento dos significados, como também garantiria a possibilidade de interpretação de omissões ou substituições por zero. De certa forma, ao se fechar a objetiva, estamos construindo um processo de eclipse visual ao mesmo tempo em que se direciona a atenção do espectador para um ponto específico do quadro.

6.2 Fluxo de informação e construção do significado

6.2.1 Componentes estruturais

Nesta seção discutirei alguns componentes estruturais⁷⁶ que se mostram relevantes para a análise do fluxo de informações dos vídeos da campanha "Get a Mac", corpus desta tese. Pode-se observar que existem algumas estratégias comuns a todos os textos audiovisuais aqui utilizados, elas estão relacionadas a um constituinte estrutural em especial: o enquadramento. É possível identificar um padrão de escolhas que parecem guiar o espectador para onde a informação está se originando. Isso ajuda a deixar claro para ele qual personagem das peças publicitárias em vídeo estudadas é o ponto de partida da informação naquele momento específico.

⁷⁶Ver capítulo 4, na página 127, para a definição de componentes estruturais.

A partir disso, poderíamos dividir os vídeos em duas categorias, partindo de seu uso dos recursos estruturais:

1. vídeos com uso regular dos recursos estruturais;
2. vídeos com uso diferenciado dos elementos estruturais.

A primeira categoria seria representada por uma parte significativa dos vídeos (8 entre os 12 que compõem o corpus), que se assemelham muito na forma de utilização desses recursos, criando a referida regularidade. Já os do segundo grupo, se caracterizariam ou por um uso de enquadramentos e ângulos acima do regular, mostrando uma característica mais arrojada na produção, ou por uma sequência de enquadramentos mais simples e estável. O quadro 8, na página 225, mostra quais vídeos são pertencentes a cada uma das categorias. O fato da grande maioria dos vídeos recair na primeira, poderia ser resultado da necessidade de regularidade; é ela que garante a percepção de pertencimento dos vídeos à campanha. Em outras palavras, ao manter o uso dos recursos técnicos com uma certa regularidade, está-se, na verdade, utilizando uma entre as tantas possibilidades combinatórias para o espectador que, nesta campanha, talvez perceba tal padrão.

Quadro 8 Vídeos e a estratégia estrutural no fluxo de informações

Tipo de Uso	Nome
Regular	<i>Viruses</i> (V-01)
	<i>Out of the box</i> (V-02)
	<i>Restarting</i> (V-03)
	<i>Better Results</i> (V-05)
	<i>Security</i> (V-06)
	<i>Vista Choose</i> (V-07)
	<i>Surgery</i> (V-08)
	<i>Meant for Work</i> (V-10)
Diferenciado	<i>Networking</i> (V-04)
	<i>Boxer</i> (V-09)
	<i>Fat HD</i> (V-11)
	<i>Breakthrough</i> (V-12)

Mas como esses elementos se manifestam? Como podemos observar na figura 47, página 226, há uma sequência de enquadramentos que tende a se repetir em diversos dos vídeos. No



Figura 47: Componentes estruturais em *Meant for Work* (V-10)

frame A da figura 47, podemos observar que *Meant for Work* (V-10) se inicia com um plano geral, que coincide invariavelmente com a abertura do vídeo. Segue-se então um plano médio-fechado em Mac (frame B), que então toma a palavra. Esse plano será mantido por alguns segundos até que tenhamos um corte para o frame C, que traz um plano também médio fechado, mas agora para PC, de forma a demonstrar que é ele que toma a palavra. O frame D restabelece o contexto geral do vídeo em um plano aberto, sendo quase que imediatamente substituído por E, em ambos os frames PC ainda está com a palavra e a manutenção em quadro claramente se dá graças a isso.

Mac toma a palavra novamente em *Meant for Work* (V-10), figura 47, página 226, como é representado pelo frame F em plano geral. É essa retomada que leva o vídeo ao frame G, no qual Mac está em plano médio fechado, dirigindo sua palavra para PC — que justifica o frame H, em um corte rápido, cuja função é mostrar que PC está quieto ouvindo o que é dito por ele. Volta-se rapidamente para Mac, frame I, que diz algumas palavras antes do corte para o frame J, no qual PC inicia a fala final. Essa fala se estende até o frame L, que, em plano geral, marca o encerramento do vídeo.

Esses resultados mostram que há uma tendência de que os cortes direcionem a visão do espectador para a personagem que está falando. Ele é considerado o Tema visual, enquanto o Rema parece não ser mostrado, omitido. Esse Rema é omitido e recuperado por um processo de coesão por ausência: não vemos, na maioria das vezes, a outra personagem, mas acreditamos que ela esteja ali devido a nossas lembranças dos *takes* anteriores. Alguns frames ajudam a reforçar essa ideia de simultaneidade, como é o caso de H: ao exibir a personagem pretensamente prestando atenção naquele que fala, está-se dando para o espectador a ideia de que o Rema visual, aquele a quem a fala é dirigida, está realmente ali. Frames como o H não acontecem a toda hora nos vídeos, apenas de tempos em tempos de forma a reforçar a memória do espectador.

A sequência de planos presente na figura 47, página 226, parece guiar o espectador perante aquilo que os produtores julgam ser relevante, ela é didática na medida que estabelece um claro "pingue-pongue" entre os falantes com eventuais retomadas do contexto geral, gerando um vídeo no qual não há desencontros entre o que se fala e aquilo que se vê. Observa-se, então, um segundo nível de simplificação nesses vídeos: o primeiro, já discutido na seção 5.1, página 164. Esta segunda, por sua vez, tem por objetivo de se certificar que o espectador está prestando atenção naquilo que os produtores esperam, gerando uma exposição direta da mensagem sem que o espectador necessite preencher muitos vazios para a sua compreensão. Isso deixa a mensagem mais simples de ser entendida, o espectador não precisa construir seu próprio processo interpretativo do fluxo de informações.

A forma como esse tipo de organização estrutural ocorre é algo constante em todos os vídeos. Um exemplo poderia ser *Viruses* (V-01), representado na figura 48. Esse vídeo, que discute a não infecção de Mac por vírus, também possui uma sequência de planos muito similar à discutida na figura 47. O vídeo se inicia com um plano geral, retratado pelo frame A; neste frame, após a apresentação das personagens, temos um grande acesso de tosse por PC, esse acesso ainda está em andamento quando temos o corte para B, no qual Mac inicia a sua fala e, então, prepara um lenço que será passado para PC. Essa simultaneidade parece

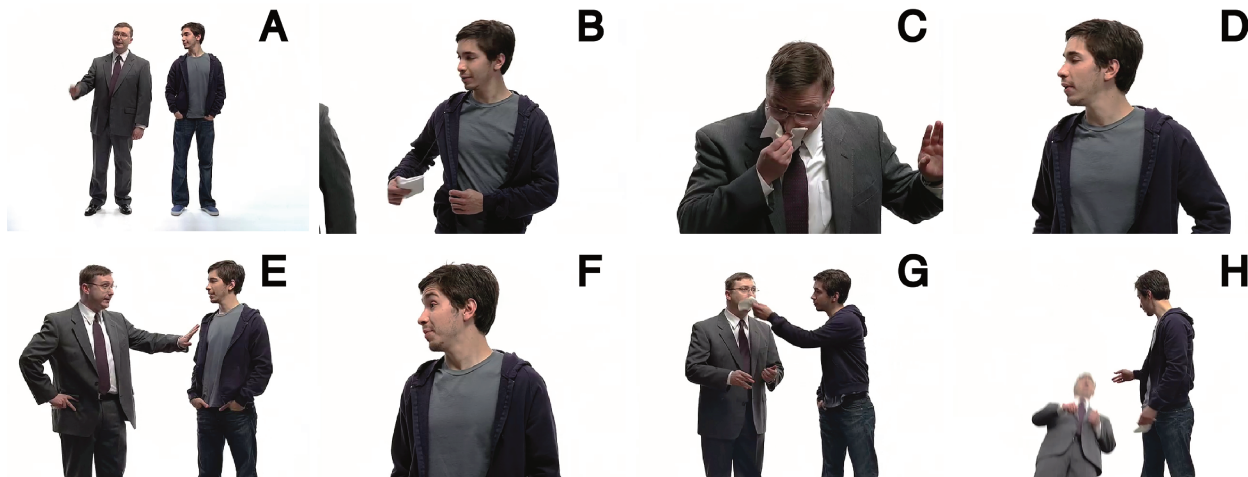


Figura 48: Componentes estruturais em *Viruses* (V-01)

mostrar que a ação de solidariedade de Mac ocorre graças à existência da tosse por conta de PC, a manutenção dessa tosse está ali para nos servir como um elemento coesivo e temático,⁷⁷ a tosse e a presença de PC se tornam a informação dada, sobre a qual Mac desenvolve o Rema, aqui relacionado com a informação nova, seu lenço. Esse mesmo lenço está presente no frame C, no qual PC limpa o seu nariz e pede para Mac que se afaste, pois poderia ser infectado, essa ação se dá entre os frames C e E, sendo que D tem a função apenas de nos mostrar que Mac ainda está ali prestando atenção no que é dito por PC. O frame D parece ter uma função quase apositiva, intercalado na realização de uma ação, ajuda a construir o processo coesivo, que parece se dar não apenas pela repetição de imagens e sons, mas também pela sequencialização de planos e informações.

O frame F marca a retomada da palavra por Mac, que explica sua imunidade em caráter definitivo para PC; o corte para o plano geral tem, então, o objetivo de encaminhar o vídeo para seu final em G e H, com a saída habitual de PC de quadro, agora desmaiando. Novamente, há o mesmo tom de "vaivém" entre o plano geral e o plano fechado, isso sempre acompanhando a personagem que detém a palavra. Algo que merece atenção é a inexistência de movimentos de câmera nesses dois vídeos, de fato a linguagem habitual desses comerci-

⁷⁷Para mais detalhes sobre a perspectiva de coesão audiovisual, ver seção 6.1 na página 214.

ais é instanciada com uma câmera estática apenas com variação de planos. Isso acaba por conferir um papel ainda mais importante para a linguagem verbal, uma vez que nela recai um peso informacional significativo; aparentemente ela é um elemento central não apenas na constituição da metáfora verbo-visual — como já discutimos anteriormente —, mas também na constituição dos elementos estruturais.

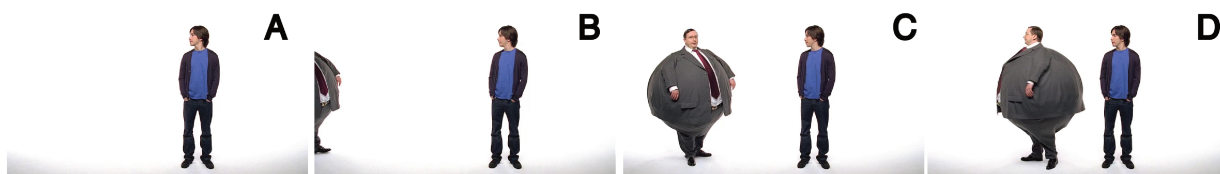


Figura 49: Baixo uso de componentes estruturais em *Fat HD* (V-11)

Nesses últimos, a importância do componente verbal parece ser ainda maior: ele parece facilitar a recepção do fluxo de informações visuais, o que parece subordinar as escolhas de planos a ela. Contudo, é importante observar que, talvez, essa não seja característica de todo e qualquer texto audiovisual. Nos vídeos aqui analisados ela parece ser importante pela natureza minimalista e simplificadora que é imposta a eles, comerciais que devem ser facilmente entendidos por consumidores leigos. Devemos sempre lembrar que a linguagem verbal e a imagem movimento nem sempre caminharam em conjunto, exemplos seriam os filmes mudos nos quais prescindia-se do som e, por conseguinte, da fala. Vale lembrar que os primeiros filmes sonoros só começaram a ser produzidos em meados dos anos 20. Até essa época, atores tinham que receber treinamento especial para serem capazes de demonstrar suas emoções com exagero na tela, de forma a serem compreendidos. Os intertextos — notas explicativas que eram projetadas entre as cenas — eram raros e só traziam o contexto geral. Nesses filmes, por exemplo, a quantidade de vazios interpretativos era muito maior, a ausência quase que absoluta de linguagem verbal fazia com que a interpretação do visual fosse completamente dependente do espectador. Filmes publicitários, como é o caso dos que compõem o corpus dessa pesquisa, não podem caminhar por essa seara, eles precisam ser os mais claros dentro das limitações temporais — comerciais têm em média trinta segundos

— e técnico-orçamentárias — muitas campanhas têm gastos previamente fechados —, sem perder a criatividade e o senso de inovação.

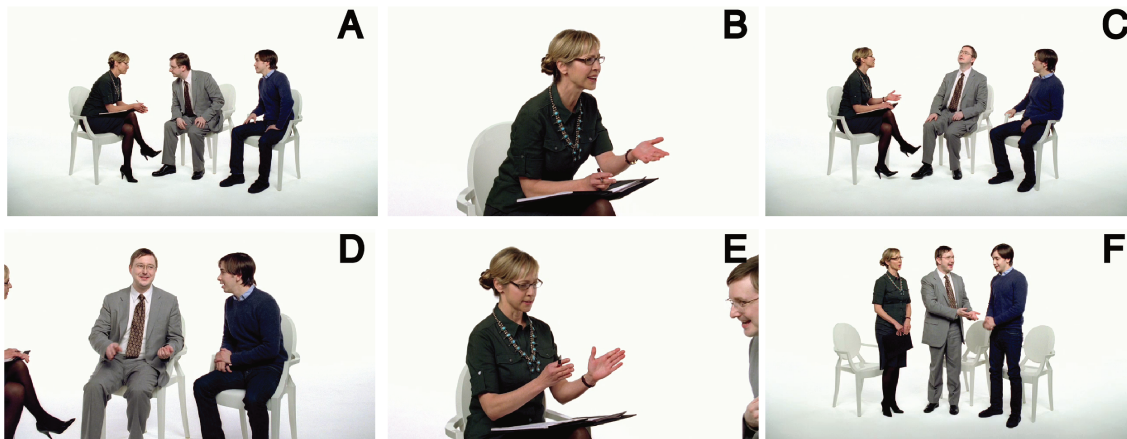


Figura 50: Baixo uso de componentes estruturais em *Breakthrough* (V-12)

Já os vídeos que instanciam os componentes estruturais de uma forma diferenciada podem ser divididos em duas categorias. Na primeira estão vídeos que fazem um baixo uso desses componentes, como seria o caso de *Boxer* (V-09), *Fat HD* (V-11) e *Breakthrough* (V-12); ao passo que a segunda é composta por apenas um vídeo *Networking* (V-04) que traz um exemplo de um super uso desses mesmos componentes.

A figura 49, página 229, que retrata quatro frames de *Fat HD* (V-11) — vídeo que tenta comparar a utilização de *trial softwares*⁷⁸ nos sistemas Mac e PC — é um exemplo de baixo uso desses constituintes. Cada um dos frames ali presentes retrata uma fase do vídeo, o frame A, por exemplo, é referente à abertura que se estende até C, momento em que PC começa a articular a razão de estar tão gordo. Logo em seguida, no frame D, PC relata ter esquecido algo e inicia sua saída de quadro, de forma bastante lenta. Como se pode observar, todos os frames de 49, página 229, possuem o mesmo enquadramento, plano geral.

Esse tipo de escolha, no caso de *Fat HD* (V-11), parece ser motivada pela instanciação de outros modos de linguagem: ao escolher retratar PC como alguém gordo e com dificuldades de locomoção — o que seria uma metáfora verbo-visual relacionando o HD cheio e lento com um indivíduo pesado —, planos mais fechados tornariam mais difícil a percepção do estado físico

⁷⁸Para explicação sobre *Trial Softwares*, ver nota de rodapé 67 na página 187.

de PC. Some-se a isso a necessidade de realizar uma dilatação do tempo, para representar a lentidão de PC, escolheu-se deixar espaços não preenchidos na tela. Esse lugar seria o de PC, sendo possível inferir isso de duas formas; primeiramente, espectadores habituados com os vídeos da campanha já saberiam que o lado direito do quadro é sempre ocupado por ele; outra possibilidade seria a direção do olhar de Mac que conjugado ao lento aparecimento de PC à esquerda do quadro nos ajudam a perceber de onde vem essa informação. Observa-se, então que o quadro já dado pode ser interpretado como Tema, ao passo que o Rema seria representado por PC e a construção da metáfora que se dá em torno dele.

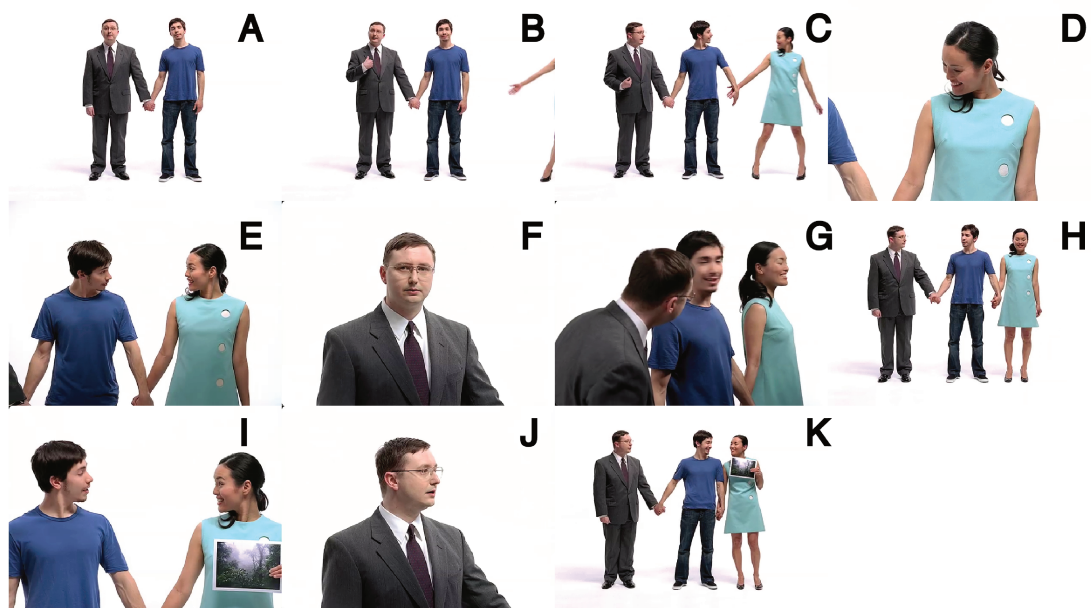


Figura 51: Uso de constituintes estruturais exclusivos em *Networking* (V-04)

Outro exemplo de utilização diferenciada no corpus é *Breakthrough* (V-12), aqui representado pela figura 50, na página 230. Ali temos Mac, PC e uma analista discutindo as inoperâncias de PC, tentando livrá-lo da culpa. Como é observável na figura 50, na página 230, o frame A representa um plano aberto em que estão os 3 em quadro. Há o primeiro corte para o frame B, que representa a analista que está com a palavra, ela é a única personagem que será retratada nesse plano, o que ocorrerá apenas mais uma vez, no frame E. Todos os planos em que PC e Mac aparecem são abertos, variam entre o geral, como é o caso de A, C e F, e o americano, frame D. Além dessa variação, cabe ressaltar que *Breakthrough* (V-12)

não segue o padrão natural de abertura, como discutido em 5.2, página 179, uma vez que é a analista que inicia a conversa, sem que haja a tradicional abertura "Hello I'm a Mac... and I'm a PC". Essa configuração visual também é motivada pelo tipo de significado que se escolheu instanciar: para que se possa mostrar o relacionamento entre as personagens na totalidade em uma seção de psiquiatria, é necessário que planos mais abertos ocorram, de forma a colocar todos em quadro. Os cortes mais fechados para a analista parecem se justificar pelo fato de ela ser o elemento catalisador da seção, é ela que coloca toda a informação nova, sempre repetida pelos outros dois, como podemos observar na transcrição a seguir:

- (20) (a) **Analista:** So PC, this issues are not your fault...
(b) **PC:** Not my fault...
(c) **Mac:** Of course not, Pal...

Já *Networking* (V-04) merece destaque especial por ser o vídeo do corpus, e talvez de toda a campanha, que possua a maior variedade de planos e movimentos de câmera. Como se pode observar na figura 51, página 231, o vídeo se abre, no frame A, com um plano geral no qual Mac e PC se apresentam como o usual. Segue-se então a sequência representada pelos frames B e C na qual uma nova personagem, a câmera digital, entra em quadro. Percebe-se que há um rápido movimento de *truck*⁷⁹ para a esquerda, de forma a aumentar o tamanho do quadro e permitir a entrada da nova personagem. Segue-se, então, um plano médio-fechado em D, que mostra a câmera digital em detalhe, corte que coincide com sua fala no vídeo. Já o frame E, mostra ela e Mac conversando em japonês, algo muito próximo do que seria um plano americano, e excluindo PC, que parece perplexidade em F. O frame F é o momento que PC inicia uma nova fala que é terminada no frame G. Há novamente um corte para plano geral em H, momento em que a câmera digital faz uma solicitação para Mac, o que propicia seu corte para o frame I, novamente com apenas Mac e a Câmera em um plano mais fechado. O frame J mostra novamente PC com tom de perplexo, prestando atenção na interação. Nesse frame, ele tenta balbuciar algumas palavras em uma língua que não aquele

⁷⁹Ver discussão a respeito da figura 16 na página 122 para uma explicação sobre os movimentos de câmera.

utilizada por Mac e pela câmera digital. Paralelamente a isso, temos o corte para K, que encerra o vídeo com Mac e a Câmera rindo de PC.

Diferentemente dos outros vídeos, *Networking* (V-04) tem um uso bastante específico dos componentes estruturais. É a primeira vez que a câmera deixa seu eixo estático para realizar um movimento de *truck*, como o que vemos entre os frames B e C da figura 51, na página 231, sendo que ele tem uma função clara: possibilitar a entrada da personagem sem que uma das marcas identificadoras da campanha seja quebrada. Outro ponto importante de se observar é o frame G: ele traz um ângulo de câmera chamado de POV (*Point of View* ou Ponto de Vista), que nos mostra a visão que uma personagem teria do *setting* de gravação e dos demais atores que com eles contracenam; no caso de *Networking* (V-04), há a visão que PC teria de Mac e da Câmera Digital.

O vídeo *Networking* (V-04) tem por objetivo representar a capacidade de Mac de se conectar de forma mais rápida e eficiente com equipamentos periféricos, como é o caso da câmera fotográfica. Talvez seja por isso que esse vídeo possua uma estruturação mais experimental. A quantidade de cortes não é necessariamente maior que os demais, mas observa-se que os cortes não parecem seguir um sistema de "pingue-pongue" como citado em quase todos os vídeos analisados nesta seção: Mac e a Câmera interagem à revelia de PC, os cortes não parecem refletir os falantes, necessariamente, eles se organizam por blocos de solidariedade e conexão, como já analisado na seção 4.3, página 149, em especial na discussão sobre a figura 19, página 150, que isolam PC e formam um bloco com as demais personagens. Isso faz com que PC seja invariavelmente retratado em um plano mais fechado que as demais personagens, parecendo interagir a distância. De certa forma, a representação do POV de PC nos leva a ter a sua percepção visual que passa por sua pretensa subjetividade e, por conseguinte, ajudando a perceber ainda mais esse isolamento.

6.2.2 Componentes atitudinais

Nesta seção discutirei como alguns componentes atitudinais podem contribuir para o fluxo de informações em textos vídeo-fílmicos. Minha hipótese é que gestos e olhares contribuem para a sinalização de mudança de quadro, ajudando-nos a prever cortes de câmera e câmbios nos turnos de fala. A utilização de componentes atitudinais, diferentemente da análise dos componentes estruturais (ver seção 6.2.1, página 224), não é classificável em termos de frequência e quantidade de uso. Sua variação é muito grande e parece estar relacionada com o espaço físico disponibilizado pela abertura da objetiva, cada tipo de enquadramento parece possibilitar que a instanciação ocorra de forma particular. Além disso, a natureza do Hiper-Tema em alguns vídeos parece ser também um elemento determinante dessas escolhas, além de interações mais e menos solidárias entre as personagens.

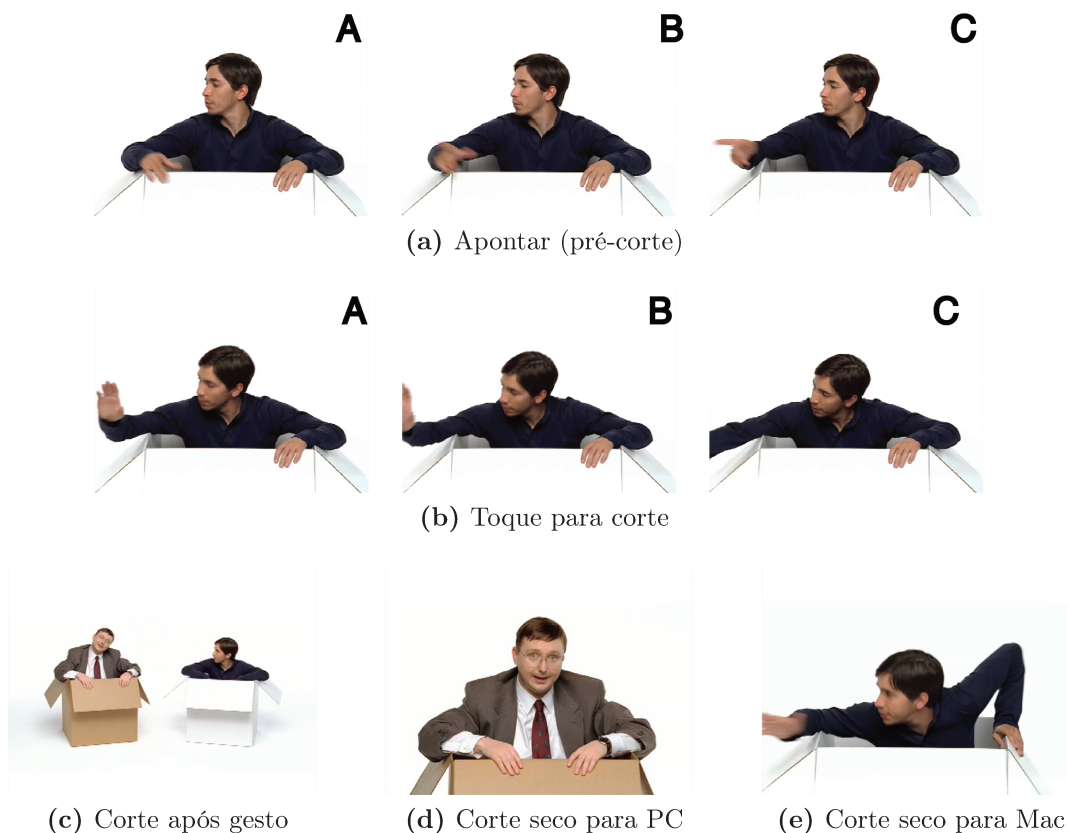


Figura 52: Gestos e o fluxo de informação em *Out of the box* (V-02)

Um exemplo de como esses constituintes são aplicados é oferecido pela figura 52 na página oposta, que traz alguns frames de *Out of the box* (V-02). Tanto em 52a como em 52b, ambas na página 234, há gestos de Mac que sinalizam a passagem da fala para PC. No caso de 52a, há um gesto de apontar que está em processo de preparação no frame A, e tem sua direção definida em pré-realização no B, sendo finalmente instanciado no frame C. Esse gesto icônico ainda não marca o corte para PC, mas dá uma prévia para o espectador avisando-o que isso acontecerá; efetivamente o corte é simultâneo à instanciação do toque presente em 52b: a preparação (frame A) e a pré-realização (frame B) já mostram a direção do pequeno tapa que dará na caixa de PC. No momento exato desse toque (frame C), a mão de Mac já saiu de quadro, ouvimos o toque de sua mão na caixa e, simultaneamente a isso, temos o corte para 52c, página 234. Esse corte leva a um plano geral, no qual PC toma a palavra e começa a explicar sua situação: estar em "várias caixas", alguns frames depois temos um corte direto, sem a indicação de gestos para 52d, no qual ele continua sua explicação. Mac sobrepõe-se e toma a palavra, novamente num corte seco, agora para 52e.

Primeiramente, essa análise parece nos mostrar que o corte de câmera do plano fechando de Mac para o plano geral está sincronizado com o ápice do gesto descrito em 52b — tal gesto serve como indicação desse acontecimento para o espectador, já o preparando para a mudança que está por vir — e com o som da mão de Mac tocando na caixa. Descorporificado, este último marcando sonoramente a transição como um tiro de largada que nos dá o instante de sincronia perfeito. Em segundo lugar, é importante reparar que apenas Mac utiliza gestos e sons para marcar sua transição. A exemplo dos frames representados em 52d e 52e, ambas na página 234, todos os cortes que partem da imagem de Mac são diretos sem a sinalização gestual ou sonora. Isso também nos mostra um critério de diferenciação das personagens nesse vídeo, Mac parece orquestrar o fluxo de informações, ao passo que PC é sujeito a ele passivamente.

Uma observação minuciosa do corpus mostra que uma parte significativa dos vídeos relaciona seus componentes atitudinais ao fluxo de informações, e isso pode acontecer de duas

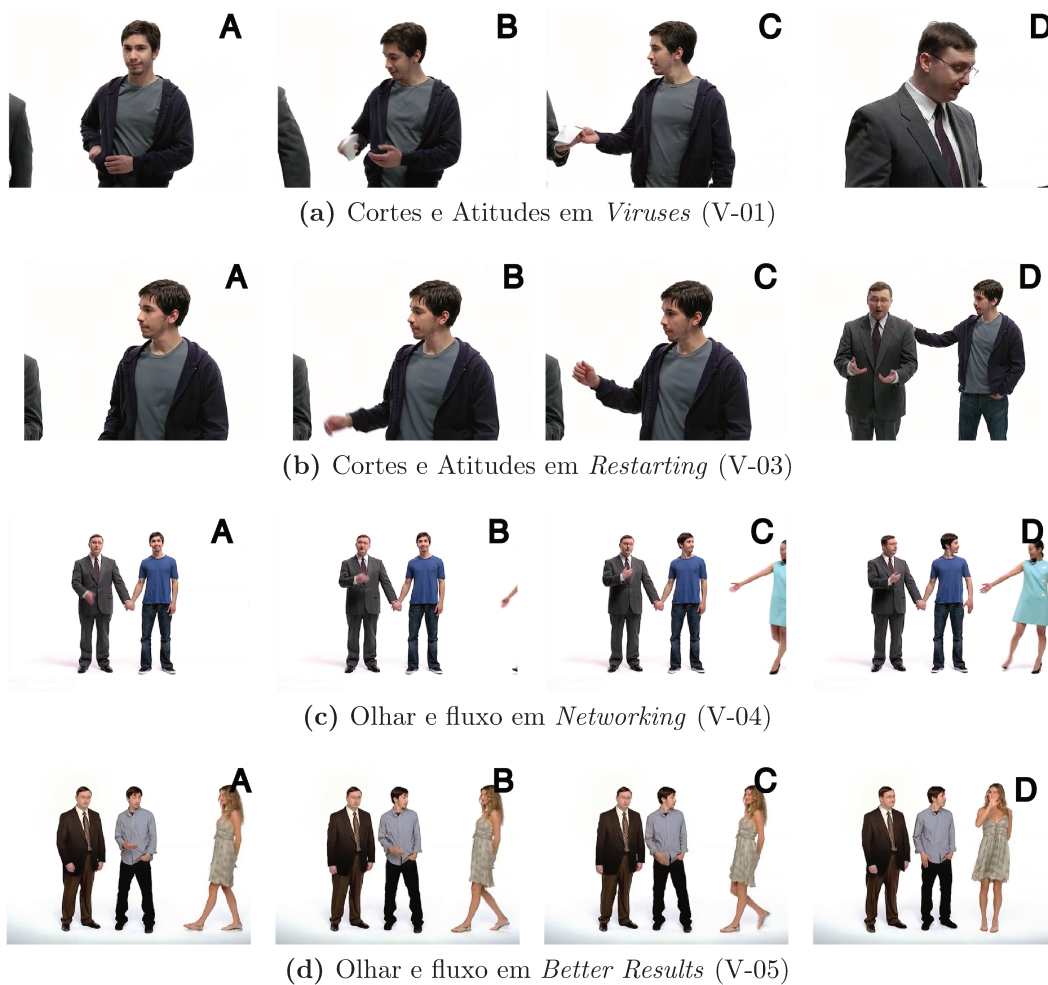


Figura 53: Olhares, gestos e sua relação com o fluxo de informação

formas. A primeira, já descrita, está relacionada com uma identificação entre componentes atitudinais e estruturais (ver análise da figura 52b, na página 234), sendo que a segunda relacionada a forma como estes mesmos componentes identificam qual a origem da informação. A figura 53 traz um resumo de como os componentes atitudinais podem se relacionar em diversos vídeos desse corpus. Na subfigura 53a, página 236, vê-se um exemplo de um constituinte atitudinal sendo utilizado como elemento indicativo de corte entre diferentes frames de *Viruses* (V-01). No frame A, Mac está iniciando um gesto de retirada de um lenço, que se completa no frame B; logo na sequência ele direciona sua mão com o lenço em direção a PC, que tem apenas sua mão em quadro enquanto recebe o lenço no frame C. No momento

exato que PC pega o lenço, temos o corte para o frame D. Como já discutido anteriormente, ver análise da figura 25, na página 176, esse componente atitudinal também é utilizado na expressão de significados de solidariedade entre as personagens. De fato, esse tipo de multifuncionalidade dos componentes pode ser um elemento comum a textos multimodais; não se deve pensar que um grupo de escolhas feito em um ou mais modos de linguagem tenha apenas uma função. Essa possibilidade de multissignificação faz com que as escolhas em um texto multimodal sejam de interpretação bastante complexa e, ao mesmo tempo, demonstrando que seus significados sociais e estruturais ocorrem de forma simultânea.

Na subfigura 53b, página 236, também temos a construção de um gesto multissignificativo: ele pode tanto indicar uma ação de solidariedade, como também um constituinte que coopera com a percepção do corte que está por vir. No frame A, vê-se Mac com sua mão direita em posição de descanso, seu braço parte em direção a PC no frame B, sem que ainda haja uma definição totalmente clara de qual será esse movimento. É apenas em C que Mac demonstra que irá dar um tapa nas costas de PC, no caso de *Restarting* (V-03), vídeo retratado em 53b, o objetivo é reanimá-lo e ajudá-lo a reiniciar o sistema. Esse tapa efetivamente ocorre no frame D, no qual temos o corte para PC; ambos ocorrendo de forma simultânea. No caso do frame D, o plano geral é necessário para mostrar que foi a ação de Mac que reanimou PC naquele contexto.

Já a subfigura 53c, página 236, mostra o trecho de *Networking* (V-04), vídeo que discute a conectividade de Mac com periféricos, tais como máquinas fotográficas, representada pela garota que aparece em quadro. Além do já analisado significado de solidariedade vs. não solidariedade — ver discussão em torno da figura 19 —, a subfigura 53c, página 236, traz o uso de um componente atitudinal, o olhar. Ao se observar o frame A, vemos que Mac e PC estão de mãos dadas olhando para o espectador, ao dirigir seu olhar para direita, Mac observa a entrada que se inicia, ambos no frame B. A câmera desloca-se para esquerda em um movimento sutil de *tilt* (ou *travelling*), permitindo a entrada da atriz, que é finalizado em D. Poder-se-ia afirmar que o olhar que é dirigido por Mac para a mulher, que está prestes a se

conectar a ele, é uma espécie de anúncio da origem da informação nova naquele contexto. Mac e PC representariam, assim, o Tema e a informação dada, ao passo que a nova personagem representaria o Rema e a informação nova. Apesar de não haver uma identificação direta entre Tema e informação dada, nos vídeos aqui analisados ela parece se instanciar com uma certa constância; talvez resultado de uma questão importante na imagem em movimento: não é possível ver um frame sem que outro o tenha antecedido, como já discutido na seção 6.1, na página 214. Isso tem implicações importantes para o processo de construção dessa linguagem, pois sempre trabalhar-se-á com um conjunto de elementos estabelecidos previamente pelo contexto. Talvez o único frame 100% inédito e novo para o espectador seja o primeiro frame de um texto videográfico ou filmico.

Algo bastante semelhante também é apresentado pela subfigura 53d, página 236, que traz alguns frames de *Better Results* (V-05), cujo objetivo é mostrar que computadores Apple realizam tarefas multimídia mais competentemente que os da plataforma PC. No frame A pode-se observar que Mac sincroniza seu olhar para sua direita, enquanto uma modelo entra pelo mesmo lado da tela. Ao mesmo tempo, ainda no frame A, percebemos que Mac aponta com sua mão para a modelo, e esse gesto vai sendo retraído à medida que ela entra em quadro, como pode-se observar no frame B; já no frame C, a mão está quase que totalmente em posição de descanso; há uma sincronia entre o final do gesto de Mac e a apresentação da modelo, no exato momento que sua mão atinge o descanso ela fica de frente ao público e acena, como se pode observar no frame D.

A subfigura 53d traz à baila uma discussão importante, a sequencialização de gestos, ações e olhares como elemento de contribuição do fluxo de informações. Se forem observados novamente os frames ali representados, será fácil perceber que cada uma das ações acontecem ou simultaneamente, como é o caso do olhar e do apontar para a modelo, assim como sua entrada, ou sequencialmente, como é o caso do fim do ato de apontar com o início do ato de cumprimentar. Isso demonstra que ações — sejam gestos, olhares ou quaisquer outros — parecem contribuir tanto para a indicação da fonte geral da informação, como para sua

sequencialização. No caso das ações, entendo que elas podem estar conectadas ou por justaposição, como é o caso das analisadas em *Better Results* (V-05), como por uma contingência de ação e reação. No caso desta última, gostaria de chamar a atenção do leitor novamente para a subfigura 53b, que representa o vídeo *Restarting* (V-03); ali o pequeno tapa que Mac dá nas costas de PC, representado pelo frame D. Essa ação desencadeia o reinício de PC, esta última nitidamente ocorrendo como uma reação à ação instanciada por Mac. Esse tipo de ação não apenas mostra qual será a próxima fonte de informação, uma vez que ela coincide com o corte e o início da fala de Mac, como também desencadeia essa ação propriamente dita. Uma consequência inicial desse tipo de conjugação de modos de linguagem é a construção de proposições que navegam por diferentes formas de expressão: a ação física de Mac desencadeia uma ação verbal de PC.

Em outras palavras, poderíamos afirmar que um Tema atitudinal desencadeou um Rema verbal? Na minha opinião analítica, sim, acredito que ao constituirmos textos multimodais, como os que são analisados nesta tese, essa possibilidade é real. Isso porque, ao possuir uma gama grande de recursos de linguagem para constituir seus significados, os autores de um determinado texto tenderiam a utilizar os que se demonstram mais afeitos à instanciamento. Se observarem-se, novamente, alguns dos frames instanciados na figura 53, na página 236, notar-se-á que a relação temática de ações se configura de forma bastante rica. Em 53a, representado *Viruses* (V-01), a ação de entregar o lenço para PC gera não apenas o corte de câmera, mas também duas ações que ocorrerão mais adiante do vídeo: PC assoa o nariz no referido lenço — o que parece ser uma relação Tema vs. Rema apoiada em causa e consequência — e, posteriormente, explica sua situação para Mac — ver análise ao redor da figura 26, na página 177 para discussão de mais frames —, ou seja, a relação Tema vs. Rema instanciada por um processo de justaposição assoar vs. falar. Outro exemplo seria 53c, na página 236, representado *Networking* (V-04), ali temos a simultaneidade entre o olhar e a entrada da atriz, que acumula, com a postura de sua mão, a intenção de dar as mãos para Mac; ações justapostas que estabelecem uma relação de Tema vs. Rema.

Essa relação também pode ser ampliada para contextos em que o componente verbal é adicionado como um elemento que se conjuga no processo de construção desse fluxo, um exemplo disso seria o vídeo *Security* (V-06), que trata do uso de sistemas firewall por computadores PC, representado pela figura 54, página 240. No frame A, estão Mac, PC e seu Firewall que tem por objetivo intermediar as interações de PC; ao se apresentar, Mac percebe que todas as suas interações serão intermediadas por ele (transcrição 21 a e 21 b, na página 241). Ainda no frame A, Firewall vira sua cabeça para o lado de PC, mostrando que é dali que deve partir a informação de autorização, uma informação nova verbal que parte de um tema visual. A voz de autorização de PC sobrepõe a de Firewall (transcrição 21 b, na página 241, e frame B, figura 54) que vira seu rosto para Mac, sentido que a mensagem enviada por PC toma. Contudo, esse é um dos poucos momentos em toda a campanha que a linguagem verbal, o componente estrutural e o atitudinal não estão correndo na mesma direção; nesse momento, apesar do movimento de cabeça indicar Mac, PC começa a falar e o corte não segue o olhar de Firewall. Percebemos então que Mac é o objetivo da comunicação, mas não seu Tema, que é representado por PC.

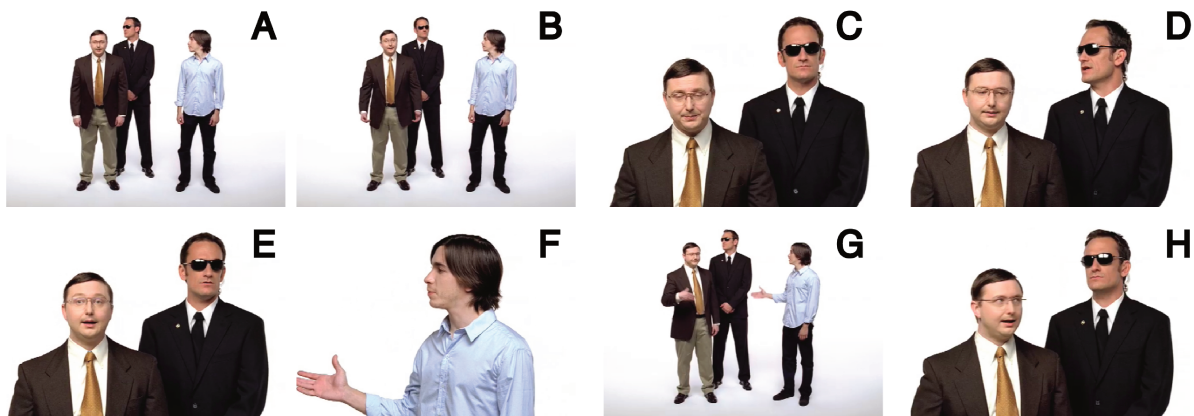


Figura 54: Fluxo de informações em *Security* (V-06)

No frame C da figura 54, página 240, PC retorna a fala saudando Mac, essa saudação é interrompida por Firewall (ver transcrição 21 c na próxima página) que novamente pede sua autorização para divulgar a informação, o que ocorre nos frames D e E (ver transcrição 21 d e 21 e). Temos, então, um corte seco para o frame F, no qual Mac, em plano fechado, aponta

para Firewall enquanto olha para PC, ambos fora de quadro e percebidos graças ao processo de lembrança do espectador, que recorda formação do quadro geral. No frame G temos um corte para o plano geral, no qual Mac termina de fazer seu questionamento (transcrição 21 f) e Firewall vira seu rosto para PC, de forma a pedir autorização para entrada da mensagem, que ocorre em H (transcrições 21 g e 21 h).

- (21) (a) **Mac:** Hello, I'm a Mac. . .
(b) **Firewall:** Mac requests for salutation, cancel or allow?
(c) **PC:** Allow, and I'm a PC.
(d) **Firewall:** You're returning Mac's salutation: cancel or allow?
(e) **PC:** Allow.
(f) **Mac:** OK, why [this]. . . ?
(g) **Firewall:** Mac is asking question: Cancel or Allow?
(h) **PC:** Allow! He's part of Vista, my new operating system. . .

Um ponto importante a ser observado é que em *Security* (V-06) os componentes atitudinais, estruturais e a linguagem verbal estão interligados, todos funcionando conjuntamente no sistema de fluxo de informações. Dada a complexidade da interação nesse vídeo, uma série de modos de linguagem parece acionada de forma a completar diferentes níveis de significado: temos a fala, que externaliza os processos de autorização e comunicação entre as personagens, os gestos e expressões faciais que, como vê-se nos frames C, D e F da figura 54, página 240, parecem refletir certa indignação de PC e Mac a respeito da existência de Firewall, ou que, como se vê nos frames C, D, E e H, mostram certa indiferença de Firewall. O movimento de cabeça de Firewall é um indicativo da personagem para quem a mensagem é direcionada. São estabelecidos assim, dois fluxos paralelos: um é o da mensagem dentro do vídeo, tendo as personagens e sua interação como referência e outro para o espectador, relacionado aos cortes de câmera e à autorização verbal de Firewall. Isso gera um filme publicitário que, apesar de simples em sua superfície, possui diversos matizes de significado, codificados pelos diferentes modos de linguagem.

6.3 Resultados parciais

Este capítulo teve por objetivo discutir os processos de construção da tecitura textual e do fluxo de informações nos textos audiovisuais que compõem o corpus de estudo desta tese. Na seção 6.1, na página 214, discuti como os processos de repetição estrutural, repetição de elementos e personagens no quadro e a coesão por omissão no fechamento da objetiva se constituem nas obras analisadas. Ali, refleti sobre a existência de um alicerce de repetição que se forma pela justaposição de diversos frames, algo necessário para a constituição da impressão de imagem em movimento. Essa repetição permite que estabeleçamos uma base interpretativa para os pequenos movimentos presentes em cada frame, sendo que é o acúmulo dessas pequenas modificações que nos permitem entender o movimento que se dá em quadro. Outro tipo de redundância estaria relacionada à manutenção de objetos e personagens em um frame, sua função é permitir um processo de ligação entre frames, auxiliando na coesão em situações em que o frame de saída de um corte tem um enquadramento diferente do frame de chegada. Esse processo parece reforçar a repetição de caráter mais estrutural, adaptando-a às necessidades locais de cada texto audiovisual por ser motivada pelo HiperTema de cada um dos vídeos. Esta seção, então, se encerra com uma discussão sobre como o espectador de um vídeo completa as informações que lhe são omitidas cada vez que a objetiva fotografa um plano mais fechado. Diferentemente da fotografia, o vídeo e o cinema podem contar com as repetições anteriormente citadas para que haja, pelo menos, recuperação parcial dessa informação na memória do espectador. Esse processo, em minha opinião é algo similar a uma elipse de caráter visual, ao se fechar a objetiva estamos propositadamente omitindo uma série de informações que passam a ser pressupostas por quem está assistindo.

Já na seção 6.2.1, na página 224, estudei como os componentes estruturais do vídeo podem ser utilizados no fluxo de informações; resultados mostram que parece haver um padrão entre a natureza do corte realizado e o tipo de foco que o realizador quer. No caso específico dos vídeos que compõem esse corpus pôde-se observar que a estrutura prima por um relativo didatismo: os cortes e planos tendem a destacar a personagem que tem a posse do turno. Isso

demonstra que os realizadores do vídeo estão preocupados em construir um texto que, pelo menos no que diz respeito ao fluxo de informações, seja mais facilmente compreendido. As escolhas de cortes e planos possuem uma relativa regularidade, fazendo com que uma grande parte dos vídeos da campanha possuam seqüências muito similares no uso de elementos estruturais, o que parece também contribuir para a constituição da identidade visual da campanha. Essa identidade não se constitui apenas nas roupas das personagens, no tipo de texto falado — que compara as características de cada plataforma —, mas também em escolhas de linguagem, uma vez que a constituição de padrões também é um elemento que ajuda no reconhecimento e construção de significados dessa campanha.

Por fim, a seção 6.2.2, na página 234, discute como elementos atitudinais são utilizados no processo de construção do fluxo de informações. Resultados mostram que eles servem tanto para sinalizar cortes de câmera como também a origem da informação em determinado frame. Por essa perspectiva, eles reforçariam os significados já construídos pelos cortes, mas com sinalizações ainda nos frames pré-corte, anunciando a transição que está por vir; além disso eles ajudariam em situações em que a origem da informação não é necessariamente óbvia, como casos em que a personagem que tem a posse do turno não está em quadro.

Entre os resultados da análise realizada neste capítulo, gostaria de destacar a pluralidade funcional; muitos dos gestos, cortes e enquadramentos podem realizar diversas funções dependendo do critério de observação. Por exemplo, em vídeos como *Networking* (V-04), representado nas figuras 51, na página 231, e 19, página 150, utilizam alguns gestos com dupla função: o olhar entre as personagens determina não apenas a origem da informação, como também um processo de não solidariedade e isolamento de uma das personagens, como já analisado na seção 4.3, na página 149. Esse tipo de multifuncionalidade pode ser pensada como uma questão importante em textos multimodais; se cada escolha em um modo de linguagem não possui uma única interpretação de sua significação, seu condicionamento à construção do significado no texto audiovisual pode propiciar uma relação multifuncional. Por essa razão, a percepção desses significados e funções vai depender da consciência do espectador em relação a essas possibilidades e à constituição dessa linguagem.

Considerações Finais

Na abertura ao esforço da significação
que vem do outro, trabalhando,
esburacando, amarrotando, recortando
o texto, incorporando-o em nós,
destruindo-o, contribuímos para erigir
a paisagem de sentido que nos habita.

Pierre Lévy

Meu objetivo nesta tese foi refletir sobre orientações teóricas e metodológicas adequadas para a análise multimodal em textos audiovisuais, discutindo um conjunto de categorias e procedimentos analíticos que foram testados na análise de uma campanha publicitária da Apple Computer Inc. Mais do que uma campanha publicitária, "Get a Mac" revela uma agressiva estratégia de marketing que visa a comparar os computadores Macintosh à principal plataforma concorrente, sendo considerada um marco na publicidade. O corpus desta tese foi composto por 12 vídeos publicitários da referida campanha, que ficou no ar por quatro anos e produziu quarenta e duas peças publicitárias em vídeo, sem contar os diversos *banners* animados que foram veiculados em sítios de revistas de TI.

O desenvolvimento deste estudo passou por duas etapas bastante claras. Na primeira, composta pelos 2 primeiros capítulos (Capítulo 1, na página 23, e Capítulo 2, página 85),

reflito criticamente sobre as bases teórico–epistemológicas nas quais este estudo se insere. Na segunda, composta pelos 3 capítulos finais (Capítulos 4, na página 127; 5, na página 163; 6, na página 213), realizo uma reflexão de caráter metodológico–analítico, que busca explorar a construção do significado em textos multimodais. Meu objetivo foi o desenvolvimento de critérios de levantamento e interpretação desses vídeos, de forma a propiciar um caminho analítico que permita uma compreensão mais apurada da forma como os diversos modos de linguagem se conjugam no processo de construção do significado.

A revisão da literatura indicou que os estudos iniciais que teorizaram sobre multimodalidade assumem pressupostos básicos da Gramática Sistêmico-Funcional (GSF) da tradição hallidiana. Dada essa tradição, a pesquisa explicita como a GSF foi inicialmente aplicada a textos verbais, tendo com ênfase tanto os conceitos básicos de linguagem desenvolvidos por essa perspectiva teórica, passando pela definição e evolução da abordagem a partir da GSF e, depois, centrando-se nos conceitos de coesão e tecitura textual, além da definição funcional Tema vs. Rema. Na sequência, ainda no plano teórico, busquei entender como os teóricos ampliaram o escopo da teoria original para aplicá-lo nas pesquisas sobre a construção de sentido em imagens estáticas e em movimento. Constatei nessa revisão da literatura que a maioria dos estudos, em especial Kress e van Leeuwen (1996) e van Leeuwen (2001), levam apenas a noção posicional em conta na análise de imagens estáticas, mas deixam de explorar alguns critérios básicos de análise de elementos técnicos da imagem, em especial da imagem em movimento. Foi a necessidade de entender tais questões que me motivou a aprofundar a reflexão e discussão sobre métodos adequados para a análise multimodal, aprofundadas no capítulo 4 (página 127) do presente estudo. Esse capítulo enfatiza a importância tanto dos elementos técnicos como das ações realizadas pelos participantes no processo de construção do significado dos vídeos analisados. Isso contribuiu para a análise da tecitura e do fluxo de informações, como foi discutido no capítulo 6 (página 213).

O estudo realizado leva a algumas conclusões importantes sobre a GSF, modelo teórico que norteou esse trabalho, além de salientar para a comunidade de linguistas aplicados, e

linguistas de um modo geral, a importância de discutir outros códigos de linguagem. Na área da Linguística Aplicada, as pesquisas mais recentes têm discutido muito sobre os múltiplos letramentos, mas esses estudos constatarem a realidade das mudanças, refletem sobre questões mais amplas de ordem social, mas não discutem a sistematização de significados estruturais de forma a perseguir uma interpretação semântico-discursiva. Nessa direção a GSF oferece parâmetros de análise e interpretação da materialidade textual que acredito trazer gratas contribuições para a análise videográfica. Defendo no presente estudo a necessidade de estabelecermos algumas ferramentas analíticas que não se baseiem apenas na interpretação estética do conteúdo dos vídeos.

No entanto, como salientei ao longo da discussão teórica desta tese, não podemos ignorar que as categorias desenvolvidas pela GSF, e herdadas pela abordagem multimodal, foram criadas tendo como referência a análise da linguagem verbal, e, por isso, necessitam de adaptações significativas para que possam funcionar em outros modos de linguagem. Um exemplo disso seria a natureza do estudo realizado nesta tese, que teve um forte apoio em estudos oriundos de áreas como a filosofia, semiótica e comunicação; esse tipo de interdisciplinaridade é vital para que o poder explicativo da teoria seja ampliado e também para nortear revisões nos preceitos básicos da GSF. No caso específico da abordagem multimodal, ressalto a necessidade de se constituir unidades que, sem perder sua capacidade categorizante, classificatória e estratificante, permitam explicar o processo de constituição do significado em modos de linguagem que não sejam o verbal. De fato, um dos elementos que dialoga a favor dessa teoria é sua adaptabilidade. No meu entender, ela não obriga o analista ou professor a "encaixar" todos os eventos do mundo em um conjunto fechado de unidades de análise.

A análise empírica do presente estudo, ao explicitar e justificar as escolhas realizadas na interpretação de cada modo de linguagem, pode contribuir para delinear um possível caminho de análise de vídeos publicitários e pode auxiliar análises empíricas críticas sobre a construção do significado em vídeos como os produzidos na campanha "Get a Mac". O modelo proposto e adotado nos procedimentos de análise permite a observação de elementos

estruturais específicos da linguagem videográfica em conjunto com modos de linguagem referentes às atitudes culturais retratadas nos comerciais analisados, além, claro, da linguagem verbal ali presente. Em termos gerais, esse processo metodológico se caracterizaria pelo mapeamento de cada um dos modos de linguagem; observação dos seus constituintes e do modo como esses estão interligados na constituição de um significado específico e estabelecimento de uma relação entre cada constituinte com a orientação discursiva do texto. Retomando em mais detalhes a metodologia de análise adotada, o mapeamento foi construído a partir da observação dos modos de linguagem presentes no texto, enquanto a observação dos constituintes está relacionada à determinação de quais elementos estão constituindo os modos de linguagem no momento de instanciação de um dado significado. O processo de relacionar o significado de cada constituinte com o texto, permitiu durante a análise descrever a forma que cada constituinte contribui individualmente para o sentido global almejado pela equipe produtora dos vídeos.

Revelou-se a necessidade do uso de ferramentas digitais e da realização de uma análise que contemple dois momentos distintos e complementares. O primeiro constitui-se da transcrição linguística, relacionando esse componente aos demais elementos multimodais que compõem os vídeos. Esse levantamento, no presente estudo, foi realizado através da construção de uma tabela, que permite associar a imagem de um *key-frame* à interpretação funcional, à transcrição linguística e possíveis regularidades estruturais que se apresentem. O segundo momento da análise está relacionado a um procedimento de transcrição que permitiu destacar quais elementos constituintes de cada modo de linguagem estão acontecendo simultaneamente. A importância deste segundo momento está em entender a sincronia da realização, traçando um paralelo entre os elementos de diferentes modos de linguagem e determinando como eles se combinam na instanciação do significado.

Embora essa perspectiva metodológica tenha se revelado como sendo bastante útil na análise dos vídeos selecionados, há algumas questões que precisam ser pensadas mais aprofundadamente. Em primeiro lugar, esse modelo não prevê o estabelecimento de unidades de análise muito

fechadas. Ele se restringe a ser um indicativo de como observar cada um dos modos de linguagem envolvidos na interação ali retratada. Em outras palavras, é necessário estarmos atentos ao fato de que cada realização videográfica pode contar com um diferente sistema de associações entre os modos de linguagem, exigindo do analista uma grande flexibilidade analítica em relação ao texto videográfico. Em princípio, essa questão levantada não é conflitante com a orientação geral da GSF que prevê a necessidade de mantermos uma flexibilidade analítica. Já nos primórdios da teoria, Halliday e Matthiessen (1999, p. 547-548) colocam que a indeterminação é uma característica comum a todas as línguas naturais, o que parece ser resultado do fato de não haver fronteiras exatas entre as categorias que configuram qualquer sistema de análise; eles são elementos artificiais, ficções teóricas criadas para nossa necessidade de compreensão do fenômeno da linguagem. Quando refletem especificamente sobre como esse princípio se aplica à análise do sistema de transitividade, Lima-Lopes e Ventura (2008, p. 17) colocam que:

[...] não existe uma fórmula exata para auxiliar na classificação dos processos, o que nos impede de traçar critérios exatos ou regras fixas. Por isso, cada vez que um analista classifica um processo ou elemento circunstancial, ele está operando com o seu conhecimento e com a sua experiência, o que pode mudar de indivíduo para indivíduo. Isso, muitas vezes, leva a várias classificações possíveis, as quais são estudadas de acordo com a função de um processo. Esse fato nos faz pensar que um trabalho de análise de transitividade pode estar ligado a vários fatores. Entre eles, temos a cultura e os valores sociais trazidos pelo linguista, os quais são, por vezes, diferentes daqueles trazidos pela comunidade estudada, o que obriga o analista a conhecer o contexto no qual um fraseado é produzido.

No caso específico dos resultados deste estudo, essa flexibilidade se dá na interpretação das próprias unidades. Apesar de haver um vocabulário próprio da construção videográfica — tais como os planos e os movimentos já estudados nos capítulos anteriores — e dos modos de linguagem corporificados (estes específicos de cada cultura), a percepção dessa associação é subjetiva e depende do conhecimento individual. O mesmo tipo de indeterminação se dá

em relação à construção das unidades narrativas, uma sequência, por exemplo, não tem um tamanho ou um ponto de corte exato, sua determinação também é de caráter subjetivo e se apoia, normalmente, na percepção de algum corte na temática discutida ou na forma de realização dos enquadramentos.

Em síntese, entendo que, mais que criar unidades de análise, o presente estudo ofereceu uma forma de constituir unidades próprias de interpretação que são aplicadas à cada situação discursiva. Boa parte disso advém de um respeito pela diversidade estrutural e combinatória que o vídeo e o cinema trazem. O modelo aparentemente aberto é, na verdade, algo favorável, pois permite sua adaptação a uma série de produções que não seguem os modelos de produção dos vídeos que analisei nesta tese, algo que amplia a aplicabilidade dos seus resultados. Se por um lado a generalidade do modelo é um ponto positivo, por outro ela exige muito do analista, em especial com relação ao conhecimento técnico; isso faz com que o analista necessite de algum conhecimento sobre linguagem videográfica para que seja possível uma análise dos constituintes de caráter mais técnico e formal do vídeo. Sua associação com os componentes de caráter mais atitudinal só será possível se o analista souber observar quais constituintes estruturais realizam quais funções. Isso em si não é um problema, apenas problematiza descrições e análises de vídeos que não adotam esse rigor.

Ou seja, apesar da realização de uma reflexão sobre questões relacionadas ao letramento tecnológico não ser o foco desta tese, a metodologia de análise aqui adotada tem importantes implicações para tal. Isso porque, nos últimos anos, parece haver uma grande vulgarização do vídeo como elemento de registro cotidiano não profissional. Proliferam-se as câmeras digitais capazes de gravar em alta definição, sendo que os programas de edição têm ficado cada vez mais acessíveis. Outrossim, refletir sobre a forma como esses vídeos constroem seus significados é um elemento importante para o uso desse recurso comunicacional; não basta apenas oferecer uma reflexão crítica sobre o papel que essa mídia possui na sociedade, é necessário também ter consciência de sua constituição significativa de forma a poder avaliar sua aplicabilidade para certos contextos. Uma consequência básica dessas colocações é a

necessidade de políticas de letramento específicas para que o vídeo e o cinema tenham suas gramáticas compreendidas e apropriadas por analistas e professores. Isso porque, no âmbito social, esse tipo de linguagem já está incorporada nas práticas de registro cotidiano, incluindo-se sua distribuição por canais online.

Um ponto que merece ser discutido é o fato de que uma abordagem como a multimodal quebra alguns preceitos básicos tanto da Linguística Teórica como da Aplicada, por trazer à baila sistemas semióticos e reflexões que extrapolam o componente linguístico da proposição. De fato, a linguística em geral tem ignorado por muito tempo tudo aquilo que não constitui a materialidade da linguagem verbal. Apesar de diversos avanços propiciados pela análise da conversação (MARCUSCHI, 1993), pela fonética e fonologia (ROACH, 1983; 2001) e, em especial, pelo crescente interesse de pesquisadores de LA em questões de hipermídia e letramento digital (GEE; HULL, 1996b; KRESS, 2003; BOLTER, 2006; HOBBS, 2006; LEMKE, 2006a; b; BAWDEN, 2008; BUCKINGHAM, 2008), pouco se tem estudado sobre os códigos que normalmente acompanham a linguagem verbal. Se toda interação é dependente desses múltiplos códigos, todas as áreas da Linguística Aplicada poderiam se beneficiar dessa nova perspectiva, acrescentando uma compreensão mais holística das práticas languageiras que estudamos. Isso, em princípio, pode gerar uma ampliação ainda mais significativa: não é apenas a relação entre a linguagem verbal e a tecnologia que pode interessar a Linguística Aplicada, mas todo o escopo de semioses que ocorrem conjuntamente à linguagem verbal. Para que esse tipo de barreira seja quebrada, é importante quebrar diversos preconceitos existentes dentro da própria área, em relação às temáticas válidas para pesquisa, permitindo ao analista, por conseguinte, expandir o escopo de suas análises e a aplicabilidade das ações pedagógicas. É importante trazermos para a escola a possibilidade de leitura crítica de vídeos na linha ilustrada pelo presente estudo.

Vale a pena ressaltar algumas questões deste trabalho que mereceriam ampliação em pesquisas futuras, entre elas ampliar o corpus de estudo para produções audiovisuais que não sejam exclusivamente relacionadas à publicidade. Apesar de algumas produções cinematográ-

ficas terem sido citadas nos capítulos iniciais desta tese, elas foram analisadas de forma mais rápida, servindo apenas como forma de exemplificar e discutir os conceitos teóricos ali apresentados. Além disso, os vídeos analisados nesta tese são produto de uma estética que prima por um uso moderado de movimentos de câmera e efeitos de edição. De forma geral, sua montagem é simples, sendo que a câmera parece apenas alimentar o ponto de vista privilegiado e onisciente do espectador. Seria, logo, importante aplicar a metodologia aqui desenvolvida não apenas em vídeos e filmes que fazem parte da indústria do entretenimento, mas também em caráter mais experimental, em especial produções independentes e não *mainstream*, como seria o caso de trabalhos de videoarte. Isso seria teoricamente relevante por demonstrar se os processos analíticos aqui desenvolvidos seriam aplicáveis a discursos que privilegiam outras formas de conduzir o olhar do espectador. Esse tipo de ampliação ajudaria a determinar se o modelo aqui proposto é aplicável a outras realidades midiáticas.

Concluindo, entendo que o estudo realizado propicia alguns avanços na compreensão da construção do significado em textos multimodais e também instiga questões sobre processos de letramento. Em relação específica ao objeto de análise eleito, vídeos publicitários, entendo que a pesquisa indica possíveis reflexões e abre para questões importantes que devem servir como base de reflexão para novos estudos. Entre as possíveis contribuições que este trabalho traz para a área de estudos da linguagem e tecnologias, gostaria de ressaltar:⁸⁰

1. **Escolha multimodal.** As escolhas multimodais se dão em diversos modos de linguagem, todos eles contribuindo para a construção e interpretação do significado.
2. **Multissignificação.** Essas escolhas instanciadas em 1 tem como resultado primeiro o fato de signos dentro da produção audiovisual são, necessariamente, multissignificativos. Isso obriga os realizadores a pensar de antemão nas possíveis interpretações que determinadas escolhas possam ter.

⁸⁰A natureza da análise multimodal leva a necessidade de estudos que tentem abarcar uma série de modos de linguagem, o que leva, naturalmente, a necessidade de criação de uma metodologia de análise que também seja compreensiva. Talvez por isso que seja importante pensar essas contribuições não podem ser vistas com uma sequência estanque, mas como uma tentativa de reflexão sobre o fenômeno multimodal.

3. Necessidade de compreensão da relação entre os vários modos de linguagem.

O processo de construção de multissignificados nesses textos não pode ter como base de expressão um único modo de linguagem, mas sim a relação entre os diversos que podem ali estar presentes. Esses modos de linguagem, devem ser pensados como um conjunto resultante do processo de multissignificação.

4. Desenvolvimento de unidades de análise próprias para cada modo de linguagem.

A abordagem desenvolvida nesta tese mostrou que o desenvolvimento de unidades de análise devem seguir os princípios básicos da análise multimodal, mas levar em conta os elementos típicos de cada modo de linguagem que instancia um determinado texto.

5. Proposta de análise videográfica.

A proposta aqui desenvolvida busca mostrar a importância entre o componente verbal — representado pela fala —, o estrutural — representado pelos elementos técnicos de linguagem videográfica —, e os atitudinais — representados pelas ações corporificadas das personagens. Elementos que ajudaram na constituição das unidades de análise em cada contexto.

6. Importância do letramento em vários modos de linguagem.

Não apenas questões de linguagem verbal que devem ser levadas em conta, o processo de construção do vídeo é algo complexo, que necessita de um amplo planejamento composicional em vários sistemas de construção de significado.

7. Necessidade de compreensão da linguagem videográfica.

O vídeo possui uma estrutura e forma própria de organização enquanto linguagem, é importante um conhecimento mínimo dessas questões de forma a produzir algo que reflita as aspirações estéticas e conteudísticas de seus usuários.

8. Adaptação ao meio e à comunidade.

Como vimos nesta tese, mesmo produtos comunicacionais profissionais adaptam-se, na medida do possível, às relações discursivas e estéticas da comunidade a que se direcionam ou servem. É imprescindível, por

consequente, que qualquer produção, por mais simples que seja, também reflita essa questão.

9. **Planejamento e pré-produção.** Essa variável depende não apenas das disponibilidades técnicas do meio, mas também da intenção no ato da realização dos produtos audiovisuais, o que certamente também estará relacionado ao Fluxo de Informações e à redundância que constitui a tecitura de tais textos.
10. **Preocupação com o produto.** Apesar do processo ser, muitas vezes, o foco do ato de letramento, textos audiovisuais são julgados enquanto produtos acabados e independentes de seu ato de produção.

Referências Bibliográficas

2001 — UMA ODISSEIA NO ESPAÇO. Direção: S. Kubrick. USA: Warner, 1968. 1 DVD (148 min.), Widescreen, color, legendado.

ALMEIDA, B. L. D. **Icons of contemporary childhood: a visual and lexicogrammatical investigation of toy advertisements.** Doutorado em Letras UFSC, Florianópolis, 2006.

AUSTIN, J. L. **How to do things with words.** Oxford: OUP, 1962.

BAKHTIN, M. (V.N. VOLOCHINOV). Estudo das ideologias e filosofia da linguagem. In: BAKHTIN, M. **Marxismo e filosofia da linguagem.** São Paulo: Hucitec, 1929/1992.

BAKHTIN, M. Os gêneros do discurso. In: BAKHTIN, M. **Estética da criação verbal.** São Paulo: Martins Fontes, 1953/1992.

BALDRY, A. P. Phase and transition, type and instance: patterns in media texts as seen through a multimodal concordancer. In: O'HALLORAN, K. L. (Ed.). **Multimodal Discourse Analysis.** London/New York: Continuum, 2004. p.83–108.

BALDRY, A. P.; THIBAUT, P. J. **Multimodal Transcription and Text Analysis.** London: Equinox, 2006.

BARRET, C. **Digital video for beginners.** Cambridge: Ilex, 2004.

BAWDEN, D. Origins and Concepts of Digital Literacy. In: LANKSHEAR, C. e KNOBEL, M. (Ed.). **Digital Literacies: Concepts, Policies and Practices.** New York: Peter Lang, 2008. p.17–22.

BEAUGRANDE, R. D. **Linguistic Theory: The Discourse of Fundamental Works.** London: Longman, 1991.

BERBER SARDINHA, A. P. Linguística de corpus: histórico e problemática. **DELTA**, v. 16, n. 2, p. 326-367, 2000.

BERBER SARDINHA, A. P. **Metáfora.** São Paulo: Parábola, 2007.

- BETTON, G. **Estética do cinema**. São Paulo: Martins Fontes, 1987.
- BIBER, D. **Variation across speech and writing**. London: Cambridge University Press, 1988.
- BIBER, D. **Dimensions of register variation**. London: Cambridge University Press, 1995.
- BIBER, D.; CONRAD, S.; REPPEN, R. **Corpus Linguistics: investigating language structure and use**. Cambridge: Cambridge University Press, 1999.
- BÖHLKE, R. D. F. **Constructing ideal body appearance for women: a multimodal analysis of a TV advertisement**. Doutorado em Letras UFSC, Florianópolis, 2008.
- BOLTER, J. D. Hypertext and the question of visual literacy. In: MCKENNA, M. C., et al (Ed.). **International Handbook of Literacy and technology**. London: LEA, v.2, 2006. p.3–13.
- BRAGA, D. B. Letramento na internet: o que mudou e como tais mudanças podem afetar a linguagem, o ensino e o acesso social. In: KLEIMAN, A. B. & CAVALCANTI, M. C. (Org.). **Linguística aplicada: suas faces e interfaces**. Campinas: Mercado de Letras, 2007, p. 18–198.
- BRESSANE, T. B. R. **Processos e Produtos no ensino de Construção de Hipermídia**. Doutorado em Linguística Aplicada e Estudos da Linguagem, LAEL-PUCSP, São Paulo, 2006.
- BUCKINGHAM, D. Defining Digital Literacy: What Do Young People Need to Know About Digital Media? In: LANKSHEAR, C.; KNOBEL, M. (Ed.). **Digital Literacies: Concepts, Policies and Practices**. New York: Peter Lang, 2008. p.73–89.
- CAMPOS JR., C. D.; LEPIANI, G.; LIMA, M.; MOREIRA, D. **Nada Mais Que A Verdade: A Extraordinária História do Jornal Notícias Populares**. São Paulo: Carrenho Editorial, 2002.
- CANEVACCI, M. **Comunicação Visual**. São Paulo: Brasiliense, 2009.
- CUBITT, S. Codecs and Capability. In: LOVINK, G.; NIEDERER, S. (Ed.). **Video Vortex Reader — Responses to YouTube**. Amsterdam: Institute of Network Cultures, 2008. p.45–51.
- DISQUE M PARA MATAR. Direção: Alfred Hitchcock. EUA: Warner, 1954/1982/2004. 1 DVD (105 min.), color, legendado.
- DONDIS, D. A. **A primer of visual literacy**. Cambridge: MIT, 1973.
- EGGINS, S. **An introduction to systemic functional grammar**. London: Printer Publishers, 1994.

- EGGINS, S. **Analysing Casual Conversation**. 2nd. London: Equinox, 2006.
- EGGINS, S.; MARTIN, J. R. Genres and registers of discourse. In: VAN DIJK, T. A. (Ed.). **Discourse as structure and process**. Thousand Oaks: SAGE Publications, 1997.
- EISENSTEIN, S. **O sentido do filme**. Rio de Janeiro: Jorge Zahar, 1947/2002.
- ELWES, C. **Video Art: A Guided Tour**. London: I.B.Tauris & Co, 2005.
- FERREIRA S. N. **Semiotic change in modern and postmodern time advertisements: an investigation based on systemic functional semiotics and social theory**. Doutorado em Letras UFSC, Florianópolis, 2011.
- FLUSSER, V. **Filosofia da Caixa Preta**. Rio de Janeiro: Sinergia/Singular, 2002.
- FLUSSER, V. **Língua e Realidade**. São Paulo: Annablume, 2007a.
- FLUSSER, V. **O mundo codificado**. São Paulo: Cosac Naif, 2007b.
- FRIES, P. H. Lexico-grammatical patterns and the interpretation of texts. **Discourse Processes**, v. 15, 1992, p. 73–91.
- FRIES, P. H. On the status of theme in English: arguments from discourse. In: PETÖFI, J. S.; SÖZER, E. (Ed.). **Micro and macro connexity of texts**. Hamburg: Buske Verlag, 1983.
- FRIES, P. H. . Themes, methods of development, and texts. In: HASAN, R. e FRIES, P. H. (Ed.). **On subject and theme a discourse functional perspective**. Amsterdam: John Benjamin, 1995a.
- FRIES, P. H. A personal view of theme. In: GHADESSY, M. (Ed.). **Thematic development in English texts**. London: Printer, 1995b.
- GASNIER, A. Le “Jornal do Brasil” n’intéresse plus personne. **Le Monde**. Paris: Disponível em: <<http://www.lemonde.fr>> acesso em: 10 dez. 2010.
- GEE, J. P.; HULL, G. Fast Capitalism: Theory and Practice. In: GEE, J. P.; HULL, G. (Ed.). **The new work order: behind the language of the new capitalism**. Boulder: Westview Press, 1996a. p.24–48.
- GEE, J. P.; HULL, G. Sociocultural literacy, discourses and new work order. In: GEE, J. P.; HULL, G. (Ed.). **The new work order: behind the language of the new capitalism**. Boulder: Westview Press, 1996b. p.1–23.
- HALLIDAY, M. A. K. **Explorations in the Functions of Language**. London: Arnold, 1973.
- HALLIDAY, M. A. K. **Language as social semiotic**. London: Arnold, 1978.

- HALLIDAY, M. A. K. **An introduction to Functional Grammar**. 1st. London: Edward Arnold, 1985.
- HALLIDAY, M. A. K. Corpus Studies and probabilistic grammar. In: AIJMER, K. E ALTERNBERG, B. (Ed.). **English Corpus Linguistics**. London: Longman, 1991.
- HALLIDAY, M. A. K. Language as system and language as instance: the corpus as a theoretical construct. In: SVARTIVIK, J. (Ed.). **Directions on Corpus Linguistics**. Proceedings of the Nobel Symposium 82. Berlin: Mouton de Gruyter, 1992a.
- HALLIDAY, M. A. K. Some lexicogrammatical features of the Zero Population Growth Text. In: MANN, W. C.; THOMPSON, S. A. (Ed.). **Discourse Description**: Diverse linguistic analyses of a fund-raising text. Amsterdam: John Benjamin, 1992b.
- HALLIDAY, M. A. K. Quantitative studies and probabilities in grammar. In: HOEY, M. (Ed.). **Data, description, discourse**. New York: Harper Collins, 1993b.
- HALLIDAY, M. A. K. **An introduction to Functional Grammar**. 2nd. London: Edward Arnold, 1994.
- HALLIDAY, M. A. K. Modes of meaning and modes of expression: types of grammatical structure and their determination by different semantic functions. In: WEBSTER, J. J. (Ed.). **On Grammar**. London: Continuum, 2005/1979. p.196–218.
- HALLIDAY, M. A. K.; HASAN, R. **Cohesion in English**. London: Longman, 1976.
- HALLIDAY, M. A. K.; HASAN, R. **Language, context and text**: aspects of language in a social-semiotic perspective. Oxford: OUP, 1989.
- HALLIDAY, M. A. K.; MATHIESSEN, C. M. I. M. Construing **Experience through meaning**: A language approach to cognition. London/New York: Continuum, 1999.
- HALLIDAY, M. A. K.; MATHIESSEN, C. M. I. M. **An introduction to Functional Grammar**. 3rd. London: Edward Arnold, 2004.
- HASHIMOTO, M. R. **Experimental Re:coletivo**: Por uma outra divisão do trabalho no filme publicitário. Mestrado em Comunicação, Faculdade de Comunicação Social Cásper Líbero, São Paulo, 2010.
- HELLWIG, B.; UYTVANCK, D. V.; HULSBOSCH, M. **ELAN - Linguistic Annotator**. Nijmegen (Netherlands): Language Archiving Technology: Max Planck Institute for Psycholinguistics 2008.
- HERRINGTON, A.; MORAN, C. **Genre Across the Curriculum**. Logan: Utah State University Press, 2005.

- HOBBS, R. Multiple Visions of Multimedia Literacy: Emerging Areas of Synthesis. In: MCKENNA, M. C.; LABBO, L. D., et al (Ed.). **International Handbook of Literacy and technology**. London: LEA, v.2, 2006. p.15-28.
- HORKHEIMER, M.; ADORNO, T. A indústria cultural: o iluminismo como mistificação das massas. In: HORKHEIMER, M.; ADORNO, T. (Ed.). **Indústria Cultural e Sociedade**. São Paulo: Paz e Terra, 2002.
- IEDEMA, R. Analysing film and television. In: VAN LEEUWEN, T.; JEWITT, C. (Ed.). **A Handbook of Visual Analysis**. London: Sage, 2001, p.183–204.
- IEDEMA, R. Multimodality, resemiotization: extending the analysis of discourse as multi-semiotic practice. **Visual Communication**, v. 2, n. 1, 2003 , p. 29–57.
- IVINS JUNIOR, W. M. **Prints and Visual Communication**. Cambridge: MIT Press, 1969.
- JONES, R. H. "You show me yours, I'll show you mine": the negotiation of shifts from textual to visual modes in computer-mediated interaction among gay men. **Visual Communication**, v. 4, n. 1, 2005 , p. 69–92.
- KAHNEY, L. **The Cult of Mac**. San Francisco: Starch, 2004.
- KENNEDY, G. **An introduction to corpus linguistics**. London: Longman, 1998.
- KLINGER, B. The art film, affect and the female viewer: The Piano revisited. **Screen**, v. 47, n. 1, 2006, p. 19–41.
- KONG, K. A taxonomy of the discourse relations between words and visuals. **Information Design Journal**, v. 14, n. 3, 2006, p. 207–230.
- KRESS, G. Against Arbitrariness: The Social Production of the Sign as a Foundational Issue in Critical Discourse Analysis. **Discourse & Society**, v. 4, n. 2, 1993, p. 169–191.
- KRESS, G. **Literacy in the New Media Age**. London: Routledge, 2003.
- KRESS, G. Gains and losses: New forms of texts, knowledge, and learning. **Computers and Composition**, v. 22, n. 1, p. 5–22, 2005.
- KRESS, G. 'Screen': metaphors of display, partition, concealment and defense. **Visual Communication**, v. 5, n. 2, 2006 , p. 199–204.
- KRESS, G. **Multimodality: A Social Semiotic Approach to Contemporary Communication**. Oxon: Routledge, 2010.
- KRESS, G.; VAN LEEUWEN, T. **Reading Images: The Grammar of Visual Design**. Oxford: Routledge, 1996.

KRESS, G.; VAN LEEUWEN, T. **Multimodal Discourse: The Modes and Media of Contemporary Communication**. London: Hodder Arnold Publication, 2001.

KRESS, G.; VAN LEEUWEN, T. Colour as a semiotic mode: notes for a grammar of colour. **Visual Communication**, v. 1, n. 3, 2002, p. 343–368.

LAKOFF, G.; JOHNSON, M. **Metaphors we live by**. Chicago: University of Chicago, 1980.

LEMKE, J. L. Multiplying Meaning: Visual and Verbal Semiotics in Scientific Text. In: MARTIN, J. R. E VEEL, R. (Ed.). **Reading Science**. London: Routledge, 1998.

LEMKE, J. L. Travels in hypermodality. **Visual Communication**, v. 1, n. 3, 2002, p. 299–325.

LEMKE, J. L. **Critical Analysis across Media: Games, Franchises, and the New Cultural Order**. Paper presented at The First International Conference on Critical Discourse Analysis - Valencia, Spain. 2004.

LEMKE, J. L. Metamedia Literacy: Transforming Meanings and Media. In: MCKENNA, M. C.; LABBO, L. D., et al (Ed.). **International Handbook of Literacy and technology**. London: LEA, v.2, 2006a. p.283–301.

LEMKE, J. L. Toward Critical Multimedia literacy: Technology, Research and Politics. In: MCKENNA, M. C.; LABBO, L. D., et al (Ed.). **International Handbook of Literacy and technology**. London: LEA, v.2, 2006b. p.3–14.

LEMKE, J. L. **Transmedia traversals: marketing meaning and identity**. Manuscript of a paper to be published in an edited version in the Proceedings of the Third International Conference on Multimodality held in Pavia— Italy. Disponível em: <<http://www-personal.umich.edu/~jaylemke/>>. Acesso em: 12 ago 2007. 2006c.

LENHADOR. Direção: N. Kassel. EUA: Imagem, 2004. 1 DVD (1h 27min), color, legendado.

LÉVY, P. A. Revolução contemporânea em matéria de comunicação. **Revista FAMECOS**, n. 9, 1998, p. 37–49.

LÉVY, P. A. **O que é o virtual?** São Paulo: Editora 34, 1996.

LIMA-LOPES, R. E. A estrutura Genérica em Cartas de Venda. **Trabalhos em Linguística Aplicada**, v. 45, p. 293–309, 2006.

LIMA-LOPES, R.E. **Visualidade na escrita: explorações sobre a tipografia**. Mimeo, S.d.

LIMA-LOPES, R.E. **A internet e sua influência nos processos de criação e distribuição da videoarte**. Relatório Final - CIP (Centro interdisciplinar de Pesquisa). Faculdade

Cáspér Líbero, 2011. Disponível em: <<http://www.casperlibero.edu.br>> acesso em: 10 jan. 2012.

LIMA-LOPES, R. E.; VENTURA, C. S. M. A transitividade em Português. **DIRECT Papers** — CEPRIL, PUCSP, São Paulo, n. 55, p. 1–25, 2008.

LINZAMAYER, O. W. **Apple Confidential 2.0**: The definitive history of the world's most colourful company. San Francisco: No Starch Press, 2004.

LIFE ON MARS. Direção: Bharat Nalluri et. al. Reino Unido: Log On, 2006/2007. 8 Blu-Rays (972 min.), color, legendado.

MACHADO, A. R. **Os Anos de Chumbo**: Mídia, Poética e Ideologia no Período de Resistência ao Autoritarismo Militar. São Paulo: Sulina, 2006.

MACHADO, A. **Arte e mídia**. Rio de Janeiro: J. Zahar, 2007a.

MACHADO, A. R. As imagens técnicas: da fotografia à síntese numérica. In: MACHADO, A. R. (Ed.). **Pré-Cinemas e Pós-Cinemas**. 4 ed. São Paulo: Papyrus, 2007b. p.220–235.

MATHESIUŠ V. **On information-bearing structure of the sentence Slovo a slovesnost 5**. Cambridge, MA.: Harvard Univ. Linguistics Dep, 1939/1975.

MATRIX. Direção: Wachowski, A.; Wachowski, L. USA: Warner, 1999. 1 DVD (136 min.), Widescreen, color, legendado.

MCLUHAN, M. The Medium is the message. In: MCLUHAN, M. (Ed.). **Understanding Media**: the extensions of man. Berkeley: Gingko Press, 1994. p.17–36.

MACHIN, D. **Introduction Multimodal Analysis**. New York: Hodder Arnold, 2007.

MACHIN, D.; JAWORSKI, A. Archive video footage in news: creating a likeness and index of the phenomenal world. **Visual Communication**, v. 5, n. 3, 2006, p. 345–366.

MARCUSCHI, L. A. **Análise da Conversação**. São Paulo: Ática, 1993.

MANOVICH, L. **The Language of New Media**. Cambridge: Massachusetts, 2001.

MARTIN, J. R. Nominalisation in science and humanities: distilling knowledge and scaffolding text. In: VENTOLA, E. (Ed.). **Functional and Systemic linguistics**: Approaches and uses. Berlin: Mouton de Gruyter, 1991.

MARTIN, J. R. **English text** — Systems and structure. Amsterdam: John Benjamin, 1992.

MARTIN, J. R. Analysing genre: functional parameters. In: CRISTIE, F.; MARTIN, J. R. (Ed.). **Genre and institutions**: Social processes in workplace and school. London: Cassel Academic, 1997.

- MARTIN, J. R. **Grammar meets genre**: Reflections on the Sydney School: Inaugural Lecture at Sydney University Arts Association — 31 de agosto de 2000. Disponível em: <<http://linguistics.org>>. Acesso em: 10 de mar. de 2001, 2000.
- MARTIN, J. R.; MATHIESSEN, C. M. I. M.; PAINTER, C. **Working with functional grammar**. London: Arnold, 1997.
- MARTIN, J. R.; ROSE, D. **Working with Discourse**: Meaning beyond the clause. 2nd. London/New York: Continuum, 2007.
- MARTINEC, R.; SALWAY, A. A system for image-text relations in new and old media. **Visual Communication**, v. 4, n. 3, 2005, p. 337–371.
- MCDONAGH, D.; GOGGIN, N.; SQUIER, J. Signs, symbols, and subjectivity: An alternative view of the visual. **Computers and Composition**, v. 22, n. 1, 2005, p. 79–86.
- MCENERY, T.; WILSON, A. **Corpus Linguistics**. Edinburgh: Edinburgh University Press, 1996.
- MENG, A. P. K. Making history in From Colony to nation: a multimodal analysis of a museum exhibition in Singapore. In: O'HALLORAN, K. L. (Ed.). **Multimodal Discourse Analysis**. London/New York: Continuum, 2004. p.28–54.
- MILLERSON, G. **Video camera techniques**. Oxford: Focal Press, 1994.
- MORAES, D. **Planeta Mídia**. São Paulo: Letras, 1998.
- MOURA, C. B. DE **Vozes marginais: Uma análise de produção audiovisual de meninos de rua**. Mestrado em Linguística Aplicada. Instituto de Estudos da Linguagem - UNICAMP, Campinas, 2011.
- NASCIMENTO, R. G.; BEZERRA, F. A. S.; HEBERLE, V. M. Multiletramentos: iniciação à análise de imagens. **Linguagem & Ensino**, v. 14, n. 2, p. 529–552, 2011.
- NEGROPONTE, N. **A Vida Digital**. São Paulo: Companhia das Letras, 2002.
- NORRIS, S. The implication of visual research for discourse analysis: transcription beyond language. **Visual Communication**, v. 1, n. 1, p. 97–121, 2002.
- NORRIS, S. **Analyzing Multimodal Interaction**: a methodological framework. New York: Routledge, 2004.
- NORRIS, S. Multiparty interaction: a multimodal perspective on relevance. **Discourse Studies**, v. 8, n. 3, 2006, p. 401–421.
- NINETEEN EIGHTY-FOUR. Direção: Michael Radford. EUA: Lume Filmes. 1 DVD (113 min.), Widescreen, color, legendado.

- NUDD, T. Apple's "Get a Mac" the Complete Campaign. **Ad Week**. Disponível em: <<http://www.adweek.com/adfreak/apples-get-mac-complete-campaign-130552>>. Acesso em: 10 de maio 2011.
- O'HALLORAN, K. L. Visual Semiosis in film. In: O'HALLORAN, K. L. (Ed.). **Multimodal Discourse Analysis**. London/New York: Continuum, 2004.
- ÔNIBUS 174. Direção: José Padilha. Brasil: Vinny Filmes, 2003. 1 DVD (118 min.), Letterbox 4:3, color, dublado.
- O'TOOLE, M. **The Language of Displayed Art**. London: Leicester University Press, 1994.
- O'TOOLE, M. Opera Ludentes: The Sydney Opera House at work and play. In: O'HALLORAN, K. L. (Ed.). **Multimodal Discourse Analysis**. London/New York: Continuum, 2004. p. 11–27.
- PASCHOAL, M. S. Z.; CELANI, M. A. **Linguística Aplicada**. São Paulo: EDUC, 1992.
- PRIOR, P. Moving multimodality beyond the binaries: A response to Gunther Kress' "Gains and Losses" **Computers and Composition**, v. 22, n. 1, 2005, p. 23–30.
- PUDOVKIN, V. I. **Film technique and film acting**. Peterborough: Sim Press, 2007. Edição para Kindle.
- ROACH, P. **English phonetics and phonology: a practical course**. Cambridge: Cambridge University Press, 1983.
- ROACH, P. **Phonetics**. Oxford: Oxford University Press, 2001.
- SADUN, E. **Digital video essentials**. San Francisco: Sybex, 2003.
- SANTAELLA, L. **Corpo e Comunicação**. São Paulo: Paulus, 2004.
- SANTAELLA, L. **Por que as comunicações e as artes estão convergindo?** São Paulo: Paulus, 2005.
- SAUSSURE, F. de. **Course in General Linguistics**. Columbia University Press: Sussex, 1959.
- SDJIC, D. **A linguagem das coisas**. Rio de Janeiro: Intrínseca, 2008.
- SINCLAIR, J. **Corpus, Concordance, Collocation**. Oxford: Oxford University Press, 1991.
- SIQUEIRA, C. P. **Análise temática em estudos de tradução — O caso do relatórios anuais de empresas brasileiras**. Mestrado em Linguística Aplicado ao Ensino de Línguas, LAEL — PUCSP, 2000.

- SNYDE, I. Attuned to the truth. **Computers and Composition**, v. 22, n. 1, 2005, p. 39–47.
- STRAUBHAAR, J.; LAROSE, R. **Comunicação Mídia e Tecnologia**. São Paulo: Tompson, 2004.
- STUBBS, M. **Text and corpus analysis**. Blackwell, 1996.
- TAVARES, H. **Teoria literária**. Rio de Janeiro: Villa Rica, 1974.
- THOMPSON, G. **Introducing functional grammar**. London: Edward Arnold, 1994.
- THX 1138. Direção: George Lucas. EUA: Warner, 1971. 1 Blu-Ray (95 min.), color, legendado.
- TOLEDO FILHO, J. C. L. **WebCultura: do método à arquitetura no ambiente digital**. Dissertação de Mestrado em Educação, Arte e História da Cultura. Universidade Presbiteriana Mackenzie, São Paulo, 2008.
- VAN LEEUWEN, T. Rhythmic structure of the film text. In: VAN DIJK, T. A. (Ed.). **New approaches to the Analysis of Mass Media Discourse and Communication**. Berlin: de Gruyter, 1985.
- VAN LEEUWEN, T. Moving English: the visual language of film. In: GOODMAN, S.; GRADDOL, D. **Redesigning English: new texts, new identities**. London: Routledge, 2001. p. 81-105.
- VAN LEEUWEN, T. **Introducing Social Semiotics**. London/New York: Routledge, 2005.
- VAN LEEUWEN, T. Legitimation in discourse and communication. **Discourse and Communication**, v. 1, n. 1, 2007, p. 91–112.
- VAN LEEUWEN, T.; JEWITT, C., (Eds). **Handbook of Visual Analysis**. London: Sage, 2001.
- VIAN JR., O.; LIMA-LOPES, R. E. A perspectiva teleológica de Martin. In: MEURER, J. L.; BONINI, A., et al (Ed.). **Gêneros: teorias, métodos e debates**. Campinas: Parábola, 2005. p.19–44.
- WATTS, H. **On Camera: O curso de produção de filme e vídeo da BBC**. São Paulo: Summus, 1984.
- XAVIER, I. **A Experiência do Cinema**. Rio de Janeiro: Graal, 1983.
- XAVIER, I. **O discurso cinematográfico**. Rio de Janeiro: Paz & Terra, 1977.
- ZETTL, H. **Video Basics 4**. Belmont: Thompson/Wadsworth, 2004.

Anexo A: Exemplo de transcrição manual

Time (in/out)	Frame Visual	Sinestesia/Olhar/Ação	Trilha sonora	Interpretação metafuncional e metáforas	Fluxo de Informações
0:00:00 0:03:12		1) Mac e PC são apresentados lado a lado, dentro de caixas de papelão 2) A caixa de papelão de PC é cinza e sem cores, a de Mac é branca, o que é consistente com as cores dos produtos comercializados na época do comercial 3) PC abana suas mãos enquanto Mac fica mais calmo, com suas mãos sobre a aba da caixa 4) As cores de Mac são bem mais opacas, muito menos vivas que Mac	1) Música típica do comercial 2) Mac: Hello, I'm a Mac 3) PC: ...and I'm a PC.	1) Mac e PC estão dentro das caixas. PC tenta visivelmente chamar mais a atenção do espectador 2) Tanto Mac como PC são dizentes em processos verbais, a audiência é o receptor 3) Os processos ditos são os relacionais atributivos	1) PC é o tema posicional, Mac é o tema verbal 2) O fluxo de informações se dá pelo término do turno
0:03:12 0:13:04		1) Mac inicia a nova ação: convida PC para iniciar, bate na caixa e faz menção de saída: 2) PC para a ação e declara que não poderá sair 3) Mac volta para sua caixa e fica estático 4) PC está fazendo exercícios de aquecimento em sua caixa	1) Música típica do comercial 2) Mac: Ready to get started? 3) PC: O no, not quite. . . have a lot to do. 4) Mac: I might a::: make a home movie, maybe create a website, try my built in camera, I do it already out of the box... So, how about you	1) Nem sempre informação verbal e não verbal coincidem, nesse caso, a declaração de PC é realizada quando ele não está em quadro 2) O ato de fala de PC tem um resultado sobre Mac, que para de realizar sua ação 3) Nesse momento, estabelece-se a metáfora do "out of the box" com pessoas que, dentro da caixa, representam a capacidade do Mac de trabalhar automaticamente	1) Há um corte seco para Mac em plano médio. 2) Mac olha e dirige sua fala para PC, que a responde fora de quadro: a informação nova que é a não partida vem de fora do quadro 3) Mac passa a olhar para a direita de onde vem a info nova, ainda de dentro da caixa 4) Corte seco para PC em plano médio 5) O mesmo acontece com Mac, ele inicia seu novo turno ainda com PC em quadro: o fluxo se dá pela passagem do turno e pelo enquadramento 6) Corte seco para PC

0:13:04 0:25:23		1) Ambos os computadores continuam dentro da caixa 2) PC inicia a interação fazendo exercícios de alongamento 3) Mac gesticula, e salta da caixa energicamente 4) PC mantém-se com feição demonstrando inferioridade	1) Música típica do comercial 2) PC: Well, first I have to download some new drivers, I have to erase the trial software that comes on my hard drive, and I got a lot of manuals to read 3) Mac: [sobrepondo PC]: Sweet!! 4) You know it sounds that you have a lot of stuff to do before you do any stuff, I gonna get started because I am a kind of excited.. Let me know when you're ready!!	1) Toda a informação dada por PC e por Mac tem por objetivo ampliar e sedimentar a metáfora do "out-of-the-box". Isso se confirma pelas falas negativas de PC, que mostram sua capacidade de fazê-lo, e as de Mac, positivas demonstrando sua capacidade de fazê-lo.	1) Passagem de turno, toque na caixa e mudança de enquadramento (plano geral) marcam o avanço do fluxo 2) Em um ponto específico da informação, quando PC diz: "I got a lot of stuff to do" há uma mudança para plano médio, 3) Mudança em plano enfatizando a questão médio-fechado para Mac 4) Mudança para plano aberto para que Mac possa marcar sua saída para a direita, PC abandonando PC	que ainda faz exercício e escuta falar
0:25:23 0:28:29		1) PC fica estático com feição de inferioridade em sua caixa. Não consegue sair e fala	1) Música típica do comercial 2) PC: Actually, the rest of me is in some other boxes, so...	Idem a anterior	Informação estática, PC aponta para fora de quadro mostrando de onde deverão vir suas peças	
0:28:29 0:30:00		Frame congelado de produto Apple				