

Análise dos recursos semióticos em filme publicitário

Josenia Antunes Vieira (UnB)
Maria Lílian de Medeiros Yared (UnB)

Introdução

Este texto tem por objetivo analisar um trecho de comercial televisivo da Caixa Econômica Federal — uma instituição híbrida, misto de banco público com banco privado. O trecho do comercial será analisado e interpretado à luz da Teoria Semiótica Social da Multimodalidade, de Gunther Kress (2010) e da Linguística Sistêmico-Funcional, de Halliday (2014). O vídeo é uma peça publicitária de 60 segundos, veiculada em 2014 na televisão durante o chamado horário nobre — entre 21 horas e 22 horas.

Esta pesquisa pretende elucidar como os recursos semióticos – posição da câmara, texto no modo fala, distância do participante para o *viewer*, olhar dos participantes, gestos, cores etc – foram utilizados pelo autor do comercial e integrados no texto multimodal para a construção do todo significativo. Consideramos que o autor do comercial é o Governo e, dessa forma, ele também é o fazedor de signos, aquele que tem os recursos semióticos e econômicos à sua disposição para utilizá-los e, com isso, transmitir a mensagem com o efeito que deseja, o que equivale a realizar uma ação semiótico-social.

Recursos teóricos

Transcrever e analisar imagens em movimento (vídeo) não é tarefa simples, porque requer instrumentos teóricos de campos diferentes para a sua consecução. Como o vídeo é uma entidade semiótica multimodal, ele precisa ser transcrito de tal modo que sejam colocados em evidência os vários modos de expressão nele contidos e isso requer uma inevitável tentativa de convergência de teorias.

Para o presente texto, recorreremos a duas teorias: à Teoria Semiótica Social da Multimodalidade (TSSM) e à Linguística Sistêmico-Funcional (LSF). Pensamos ser relevante apresentar esses dois construtos teóricos, uma vez que a TSSM recorre a alguns termos referentes à LSF para a análise multimodal. Além disso, no dado analisado, existe o texto no modo fala, que necessita ser interpretado à luz da Linguística Sistêmico-Funcional, especificamente quanto aos processos relacionais no âmbito da posse, às metáforas gramaticais e às avaliações expressadas pelas imagens.

Em 2010, Gunther Kress, em seu livro *Multimodality*, esboça uma Teoria Semiótica Social da Multimodalidade, na qual procura estabelecer uma abordagem que descreva de forma precisa conceitos e definições relativos à comunicação humana em geral, com os seus vários sistemas semióticos. Preocupa-se em construir um arcabouço teórico capaz de analisar os diversos textos multimodais que estão se proliferando em praticamente todos os atos comunicativos das sociedades modernas. Agora, não existe mais a predominância do texto verbal na comunicação em geral e, sim, uma integração do sentido realizada pelo texto verbal, imagens, som, cores, expressão corporal etc.

A Teoria Semiótica Social da Multimodalidade defende que o homem é um fazedor de signos. Ela insere a noção de recurso semiótico, mais abrangente que o conceito de signo, fazendo referência ao fato de que nem todos possuem os mesmos recursos em uma sociedade. As sociedades humanas se valem dos recursos semióticos para realizar a comunicação. Assim, para Gunther Kress, autor da teoria, os recursos semióticos são a verdadeira ferramenta da comunicação humana. E entre esses vários modos de se expressar, existem a linguagem verbal, as imagens, os gestos, o tom, o ritmo etc. Neste trabalho, podemos observar como os vários recursos semióticos se integram para a formação de um todo significativo.

Podemos dizer ainda que a TSSM foca a sua atenção no significado construído durante o processo da comunicação – seu papel e função na ampla rede social de transmissão, reiteração e perpetuação de valores e visões de mundo. Ela parte do princípio de que todo autor/escritor de textos é um fazedor de signos e, dessa forma, ele escolhe os recursos semióticos de acordo com os seus interesses e com o que está ofertado pelo ambiente social.

A teoria da multimodalidade vê a comunicação humana como um sistema complexo e rico de produções semióticas. Ela surgiu graças à necessidade de se desenvolver uma teoria que conseguisse abarcar os textos multimodais da modernidade, principalmente aqueles veiculados na televisão e na Internet, ou aqueles feitos para os novos suportes, como os *tablets*. Mesmo os chamados livros didáticos estão se transformando para atender à demanda cada vez mais crescente da sociedade por luzes, cores, movimento, interação, enfim, por textos definitivamente multimodais.

O conceito mais fundamental da Teoria Semiótica Social da Multimodalidade é o conceito de *recurso semiótico*, que amplia a ideia de signo e torna claro que o uso, o contexto e o interesse do autor são os fatores que definem o rumo funcional e, assim, o uso do recurso semiótico em determinado ato comunicativo. O recurso semiótico é qualquer aparato (cor, fonte, palavra, gesto, fisionomia) usado para comunicar. O recurso semiótico é o conceito que redefine o signo como sendo motivado – e não arbitrário –, de acordo com o interesse do autor. Em relação a esse conceito basilar da semiótica social, afirma van Leeuwen:

Então, na semiótica social, recursos são significantes, ações observáveis e objetos que foram desenhados no domínio da comunicação social e que têm um potencial semiótico teórico, constituído por todos os seus usos passados e por seus usos potenciais, e um potencial semiótico real, constituído por aqueles usos passados que são conhecidos e *considerados relevantes pelos usuários do recurso* e por usos potenciais que podem ser descobertos pelos usuários, com base em suas necessidades e interesses específicos¹. (van LEEUWEN, 2006, p. 4, tradução nossa, grifo nosso).

Essa afirmação de acordo com a qual o usuário do recurso semiótico escolhe o recurso com base em seus interesses coaduna-se com a Linguística Sistêmico-Funcional, pois esta sustenta que a linguagem evolui de acordo com as necessidades sociais.

Ao estabelecer o esboço de sua Teoria Semiótica Social da Multimodalidade (TSSM), Kress reiterou o caráter multimodal da maioria dos textos circulantes nas sociedades modernas. Como afirmam Vieira e Silvestre:

À parte as mudanças globais, que exerceram marcante influência sobre a linguagem, pertencemos a uma sociedade da imagem; somos cidadãos multimodais a ponto de descansarmos quando vemos imagens em frente à TV. Somos fruto de uma sociedade digital, uma sociedade multimodal. Foi nesse favorável contexto que o discurso monomodal encontrou terreno fértil para se ressemiotizar e compor os atuais discursos multimodais (2015, p. 38).

É possível constatar que nossa sociedade não se encontra meramente num mar informativo verbal, mas num oceano cultural multimodal. Os textos são multimodais porque são realizados por vários modos: o modo escrita, o modo fala, o modo imagem, o modo cor. E modo, segundo a TSSM, é o recurso semiótico trabalhado, evoluído e moldado histórica e culturalmente em uma dada comunidade. Segundo Kress,

[...] o modo carrega e é moldado por profundas orientações ontológicas e histórico-sociais de uma sociedade e carrega sua cultura com ele em cada signo. O modo nomeia os recursos materiais, moldados em histórias frequentemente longas de esforço social e disponíveis como recursos de significado (KRESS, 2010, p. 114)².

Dessa forma, o modo é o recurso semiótico histórico-social, com a carga semântica de vários usos passados e reconhecido pela comunidade. E o modo pode servir de expressão para o discurso. Assim, por exemplo, o par verde/amarelo no Brasil incorporou uma dimensão muito grande de significação, dentro da qual se encontra necessariamente a ideia de Nação, de identidade brasileira; nesse sentido, também podemos observar que o modo potencial típico do sorriso do brasileiro, presente em vários comerciais, por exemplo, pode expressar a cordialidade, nossa identidade cultural básica.

1. So in social semiotics resources are signifiers, observable actions and objects that have been drawn into the domain of social communication and that have a theoretical semiotic potential constituted by all their past uses and all their potential uses and an actual semiotic potential constituted by those past uses that are known to and considered relevant by the users of the resource, and by such potential uses as might be uncovered by the users on the basis of their specific needs and interests.

2. Mode is meaningful: it is shaped by and carries the 'deep' ontological and historical/social orientations of a society and its cultures with it into every sign. Mode names the material resources shaped in often long histories of social endeavour and available as meaning resources. (Kress, 2010, p. 114).

Os recursos semióticos servem ao interesse do autor/escritor (*fazedor de signos*) para que este possa construir a sua unidade comunicativa, que pode ser um comercial. Em um vídeo publicitário, a linguagem, as imagens, as cores, os movimentos, a música, trabalham integralmente na construção do significado do texto. É o chamado “princípio da integração dos recursos semióticos”, expressão cunhada por Baldry e Thibault. Segundo esses autores, o significado de um texto multimodal não é a soma dos diferentes recursos semióticos, mas a sua integração (BALDRY e THIBAUT, 2006, p. 18).

Lembremo-nos de que normalmente os comerciais para televisão podem apresentar várias pequenas narrativas, como é o caso do comercial da CEF. Nesse trecho analisado, há duas narrativas (o gerente abrindo a conta do cliente e a participante convidando o *viewer* para abrir uma conta na Caixa). Em relação a representações de narrativas, podemos dizer que, assim como em um enquadramento de uma paisagem, a percepção humana deduz o que está em volta, em um enquadramento de uma imagem “em movimento”, em uma “narrativa visual”, nossa percepção deduz o que aconteceu antes e o que acontecerá depois. Nesse processo, o *viewer* pode ser induzido a deduzir determinados fatos de acordo com os recursos semióticos utilizados e com a representação social que está sendo transmitida.

Um dos principais recursos semióticos utilizados na produção desse tipo de comercial ainda é linguagem. A Linguística Sistêmico-Funcional define a linguagem como uma ferramenta/teoria social desenvolvida pelo homem para conhecer a realidade (metafunção ideacional), interagir com os outros (metafunção interpessoal) e produzir textos (metafunção textual). Trata-se de ferramenta extremamente complexa e sofisticada, formada por vários sistemas, cada um com a sua funcionalidade dentro do universo da sociedade humana.

Embora Halliday tenha enfatizado o caráter simultâneo das metafunções da linguagem em uma comunicação verbal, é relevante observarmos que, neste tipo de gênero publicitário – propaganda de um banco público e, portanto, propaganda do Governo –, existe uma tendência à ênfase na metafunção interpessoal, porque a finalidade maior de uma peça publicitária dessa espécie é o convencimento e a inculcação de uma visão de mundo para a venda de um produto ou de uma ideia, por meio do afloramento das emoções do *viewer*. Por isso, é importante atentarmos para os processos utilizados e as metáforas gramaticais (ordem ou convite expresso por uma afirmação), ambos no texto no modo fala. O sistema da metafunção interpessoal da linguagem diz respeito aos atos de fala (demanda, oferta), que, no caso do dado em análise, pode ser realizado visualmente.

Das três metafunções de Halliday, a função que está ligada ao emissor/autor da mensagem é a função interpessoal, ou o significado interpessoal. Ela diz respeito a três características: indexalidade (finito, o aqui e o agora

estabelecido pelo autor da mensagem); modalidade (a intromissão do autor da mensagem na questão dos graus de certeza da sua representação ou de obrigação), também relacionada à dêixis interpessoal, e funções de fala (a ordem ou a oferta; a afirmação ou a pergunta, conforme estabelecido pelo autor). Todos esses campos categoriais estão relacionados à metafunção interpessoal, muito relevante para este estudo. Para a análise desse comercial, é procedente considerarmos como prioritárias essas quatro grandes áreas de categorias interpessoais de análise: o finito (localização temporal e espacial), a modalidade, o tipo de função de fala e a avaliatividade.

A metafunção interpessoal está conectada à variável de contexto denominada de “relações” (tenor), semanticamente ligada à troca, relacionada ao *quantum* de interação. Essa função é expressa também pelos recursos lexicogramaticais do sistema do MODO, no qual se encontram as escolhas determinantes de uma ordem (imperativo), uma pergunta ou uma afirmação (indicativo).

Como fazer para projetar o sistema de transitividade, pertencente à metafunção ideacional, na metafunção interpessoal? Nesse caso, é procedente identificar o autor da mensagem como dizente de uma *verbiage* (o comercial). Se considerarmos o vídeo uma *verbiage* do Governo Federal, teremos que o vídeo é uma representação e, dentro dele, há outro dizente não participante (apresentador cuja imagem não está presente), que constrói representações das práticas sociais. Assim, a voz do Governo fica camuflada na voz do apresentador não participante e, desse modo, as práticas sociais são recontextualizadas como uma representação da representação. Dessa forma, considero que o Governo Federal é o dizente de uma *verbiage* específica (o vídeo com a propaganda). O caminho percorrido entre a ideologia governamental, sustentáculo do comercial, e o produto final pode ser descrito à luz do conceito de ressemiotização de Iedema, um conceito muito útil para entendermos as modificações ocorridas em um conjunto semiótico ao longo da produção de um texto multimodal. Esse caminho percorrido, em escala institucional, no caso dos dados do nosso estudo, seria: ideologias governamentais, departamento de marketing da Caixa Econômica Federal, agência de publicidade, transmissão pelos canais de comunicação e recepção pelos *viewers*. Tal como a brincadeira infantil do telefone sem fio, o produto final frequentemente não diz muito de sua origem.

Recursos metodológicos

Os *corpora* desta pesquisa compõem-se de um trecho de 4 segundos de um vídeo de 30 segundos referente a um comercial da Caixa Econômica Federal de 2014. Esse vídeo foi seccionado em trinta imagens, mas selecionamos para o presente artigo apenas quatro quadros, inseridos na matriz de transcrição multimodal, o que permitiu a percepção de vários recursos semióticos utilizados nos comerciais.

Foi utilizada como ferramenta interpretativa a matriz de recursos semióticos elaborada por Baldry e Thibault (2010), ligeiramente modificada para a consecução da pesquisa em tela. Como essa matriz separa o filme quadro a quadro e considera cada quadro uma imagem a ser analisada, ela permite vermos a integração dos vários recursos semióticos para a formação do significado (gestos, postura, distância do *viewer*, olhar direcionado ou não ao *viewer* etc.)

A matriz original de Baldry e Thibault define em suas colunas os seguintes elementos: o tempo em segundos, o quadro visual (um quadro por segundo), a imagem visual (com glosas notacionais sobre as perspectivas horizontais e verticais, sobre as distâncias entre os participantes e o *viewer*, olhar dos participantes), cinestesia, trilha sonora e interpretação metafuncional. Para este texto, nossa matriz contém quatro colunas, a saber: *tempo*, *quadro visual*, *imagem visual e texto no modo fala*. Não inserimos a coluna interpretação metafuncional, porque nosso objetivo é apontar pontualmente os recursos referentes à metafunção interpessoal. Também optamos por mudar o nome “trilha sonora” por “texto no modo fala”, uma vez que o filme não contém música de fundo, mas a voz do apresentador.

O construto matricial mostra a evolução das imagens e as unidades que formam o todo, ratificando o fato de que o significado de um texto multimodal não é produzido pela mera soma dos recursos semióticos, mas pela *integração* desses recursos. Em cada quadro visual, essa matriz se utiliza das seguintes categorias: PH (perspectiva horizontal); PV (perspectiva vertical); D (distância entre o participante e o *viewer*); PV (participante do vídeo); CV (conteúdos do vídeo), cor e cinestesia. A matriz terá quatro linhas, correspondentes a quatro quadros.





Análise dos dados

Será feita a análise de excerto de comercial em forma de vídeo da Caixa Econômica Federal, veiculado na TV a cabo, Canal SportTV, no dia 09/02/2014, às 21h30min, dia e horário em que parcela significativa da população brasileira assiste à televisão, porque o futebol é a ‘paixão nacional’, como nos é reconhecidamente repetido e inculcado. Esse comercial tem 30 segundos. Ele mostra que a Caixa Econômica Federal é diferente de todos os outros bancos, porque *é de todos*

os *brasileiros*, o que descreve uma situação irreal, mas provocadora de emoções. Nenhum banco público pertence ao cliente, mas oferece um serviço a esse em troca de dinheiro; ademais, há muitíssimos brasileiros que não possuem conta corrente por absoluta falta de condições financeiras.

A sequência de imagens da matriz de recursos semióticos devolve a lembrança da percepção de movimento do filme. Embora haja perda de significado na transcrição do comercial, uma vez que as imagens visuais matriciais são estáticas, ganha-se no aumento da consciência dos recursos semióticos utilizados simultaneamente. Quando assistimos a um comercial, quase não percebemos as nuances dos recursos semióticos utilizados pelo fazedor de signos, pois as imagens (cerca de 14 por segundo) passam rapidamente pelas nossas retinas. Essa matriz permite uma análise no sentido pleno da palavra, uma “quebra” das partes para visualizar melhor o todo. Veremos a seguir, a matriz de integração dos recursos semióticos (Quadro 1).

Quadro 1 – Trecho do comercial da Caixa Econômica Federal

T	FRAME VISUAL	IMAGEM VISUAL	TEXTO MODO FALA
24		PH – oblíquo PV – mediano D – tomada média (social) CV – mesa de gerente de banco PV – gerente e cliente Cor – azul Cinestesia: sorriso e aperto de mão.	<i>Você precisa</i>
24		PH – oblíquo PV – mediano D – close médio CV – persiana da janela atrás PV – cliente Cor – azul Cinestesia: sorriso	<i>De um banco</i>
25		PH – oblíquo PV – baixo D – social (cintura para cima) CV – agência da CEF, placa, céu azul PV – moça vestida de azul Cor – azul Cinestesia: sorriso e gesto com a mão.	<i>Que é de todos</i>
26		PH – oblíquo PV – baixo D – social (cintura para cima) CV – agência da CEF, placa, céu azul PV – moça vestida de azul Cor – azul Cinestesia: sorriso e gesto com a mão.	<i>Os brasileiros</i>

Fonte: Elaborado pelas autoras.

A primeira coluna da matriz, intitulada T, de tempo, indica o momento do filme em que o quadro apareceu e a duração de quadros visuais semelhantes. Ela permite localizar a imagem na sequência e ver a duração de determinada sequência. No caso do comercial, podemos ver que a duração dos quadros visuais finais (T25 e T26) é de

dois segundos, o que é muito para um filme. Isso significa que a mesma composição visual (uma moça vestindo azul, sorrindo, perto de um prédio com fachada da Caixa, sob um céu azul) foi enfocada por meio da utilização de 28 imagens, na verdade. Levando em conta que, para cada segundo, são 14 quadros que passam rapidamente por nossos olhos, podemos constatar que essa sequência dos quadros em T25 e T26 apresentou 28 quadros com a mesma estrutura. Assim, é possível afirmar ser bastante relevante para a formação do significado desse texto multimodal, utilizar recursos como o sorriso da participante, a pequena distância dela para o *viewer*, o céu azul, e a cinestesia corporal da participante convidando o *viewer* a abrir uma conta na Caixa.

O quadro visual nos mostra a imagem, que pode ser relacionada com o texto no modo fala (há um apresentador), e com os componentes da imagem que revelam as categorias de interação entre os participantes e de interação entre os participantes e o *viewer*. Nos dois primeiros quadros não há a interação dos participantes com o *viewer*, mas pode-se perceber que o texto no modo fala é dirigido ao *viewer*, informando-o sobre o que ele precisa. Essa forma verbal “precisa” significa “tem que ter”, realizando um processo relacional no âmbito semântico de posse.

No que diz respeito à utilização da cor como recurso semiótico, prevalece a cor azul; trata-se de um matiz que na cultura brasileira indica felicidade (“tudo azul”; “vesti azul, minha sorte então mudou”), tranquilidade e paz. O azul remete à calmaria do mar e à bondade do céu, temas muito presentes em nossa cultura. Por si só, essa cor tem uma carga simbólico-emotiva muito grande, e sua utilização no comercial enfatiza o acolhimento quase maternal da Caixa aos brasileiros.

Quanto à cinestesia, ou movimentos corporais, há uma relação de proximidade crescente entre os dois participantes (o cliente e a gerente) e entre a participante e o *viewer*. Essa distância encontra-se no limiar entre a distância social recomendada e uma distância menor, característica de relações menos formais, quase de amigos.

Considerando a perspectiva vertical dos quadros visuais, percebemos uma prevalência da perspectiva vertical alta, com o participante ficando em um plano superior em relação ao *viewer*, especialmente nos quadros 25 e 26. Em princípio, isso poderia indicar uma relação assimétrica, denotando que a Caixa Econômica Federal é uma instituição bancária que tem mais poder que os telespectadores, o que é verdade. Mas essa aparente assimetria pode indicar também uma suposição de que o *viewer* não é ainda cliente da CEF e, como tal, de certa forma é “inferior” ao participante que o encara. Entretanto, ela o convida a mudar de posição e a ficar no mesmo plano que ela, o que resolveria a assimetria inicial. É possível afirmar que se está diante de uma proposta de se manter no mesmo nível que os outros clientes da Caixa.

Em relação aos conceitos da Linguística Sistêmico-funcional, pode-se observar, no trecho analisado, que há um comando (proposta) realizado pelo movimento corporal da participante nos quadros 25 e 26. O texto verbal (texto no modo fala) é composto de dois processos relacionais, ambos indicando posse: “Você *precisa (tem que ter)* de um banco que *é de (pertence a)* todos os brasileiros.”. Isso perfaz uma metáfora de modo: uma proposta realizada como afirmação com processos relacionais na área semântica da posse. Há ainda uma interação da participante com o *viewer*, quando a participante realiza uma metáfora interpessoal visual com o gesto das mãos, chamando-o para entrar, quando está “convidando-nos” a abrir uma conta na instituição. Trata-se de uma ordem-convite (papel de fala) realizada visualmente; é uma troca de serviços.

No que tange ao finito (*dêixis* temporal), os processos verbais estão no presente do indicativo, um tempo indefinido que indica a perpetuidade do que está sendo dito. As orações têm baixa modalização, porque não há incertezas no que está sendo dito — isso significa que, gramaticalmente falando, não foram usados os recursos que põem em dúvida o que está sendo comunicado, os modalizadores, especificamente advérbios de modo que indicam probabilidade ou possibilidade.

Nas imagens visuais, as avaliações positivas podem ser vistas no sorriso e na alegria dos participantes que são clientes da CEF. No texto falado, temos avaliações positivas em relação à instituição, basicamente introduzidas pela oração encaixada (um banco *que é de todos os brasileiros*). Essa oração encaixada funciona como suporte para a construção da identidade da instituição.

A integração entre as imagens do trecho do filme e o texto no modo fala (apresentador) é feita pelos processos relacionais que estabelecem a relação de posse. Para Halliday, os processos relacionais ligam estaticamente duas entidades, criando classificações, seja caracterizando, seja identificando. As relações possessivas incluem outros traços semânticos em determinados processos como precisar (necessitar de requerer, demandar). Na oração “Você precisa de um banco”, o verbo “precisar” não tem a conotação de uma necessidade, mas de uma exigência, um requerimento (precisar ter). Em relação a essas relações abstratas, afirma Halliday:

3. In addition to possession in the usual sense of ‘owning’, this category includes abstract relationships of containment, involvement and the like. Among the verbs commonly occurring in this function are include, involve, contain, comprise, consist of, provide. Some verbs combine the feature of possession with other semantic features; for example *exclude* [negative] + have, *owe* ‘have on behalf of another possessor’, *deserve* ‘ought to have’, *provide* ‘have as a resource’, *require* ‘need to have’, *lack* ‘fail to have’.

Em adição à posse, no sentido comum de ‘ter a propriedade de’, esta categoria inclui relações abstratas de contenção (containment) e semelhantes. Entre os verbos comumente ocorrendo nesta função estão incluir, envolver, conter, compreender (comprise), consistir de, aprovisionar. Alguns verbos combinam o traço de posse com outros traços semânticos: por exemplo ‘excluir’ [negative], ‘ter a propriedade de’ (ter em nome de outro possuidor), merecer (‘ter que ter’), prover, fornecer (ter como recurso), **requerer (precisar ter)**, faltar (‘falhar em ter’) (HALLIDAY, 2014, p. 296, tradução nossa, grifo nosso).³

Essa classificação de “precisar” como “ter que ter” está em consonância com a expressão “banco de todos os brasileiros”, também indicativa de posse. O comercial vende a ilusão de que os brasileiros são donos da CEF, o que não é verdade, mas essa retórica fantasiosa serve como um excelente recurso para despertar a emotividade restrita do *viewer*/telespectador.

Considerações

Textos em formato vídeo contêm vários sistemas semióticos interagindo ao mesmo tempo e no mesmo contexto para a formação do significado. O texto em formato vídeo é um grande signo, cujas formas de expressão podem ser sistemas semióticos complexos como a linguagem verbal falada ou a imagem. Há um entrelaçamento entre o construto linguístico e o construto visual, formando um todo textual em formato de redes de relações intersemióticas.

A análise do dado aponta para a utilização de recursos que facilitam a interação emotiva e a identificação do *viewer*/telespectador com a idealização do brasileiro realizada pelo comercial. São recorrentes no comercial as cores nacionais (especialmente o amarelo e o azul), o sorriso dos participantes, a amabilidade e o acolhimento realizados pelo movimento corporal dos participantes, pela curta distância entre o participante e o *viewer*, pelas perspectivas horizontal e vertical da câmera, e também especialmente pela recontextualização das práticas sociais, representadas sem as contradições que lhe são inerentes. Convém lembrar que, em uma interação como essa, entre o *viewer* e o vídeo, há uma assimetria evidente, em que o fazedor de signo teve todo o tempo para preparar o seu texto, podendo realçar a metafunção que lhe for mais conveniente. Além disso, o *viewer* não pode responder nada, visto que não há tomada de turno, nesse caso. Sendo assim, o fazedor de signo pode, de fato, realçar a metafunção interpessoal, apelando à emotividade do *viewer*. Mesmo se considerarmos o texto no modo fala, perceberemos que ele se limita a fazer declarações em forma de metáforas gramaticais (convite ou ordem em forma de afirmações: “você precisa de um banco que é de todos os brasileiros”) e não descreve a realidade do banco e de seu atendimento ao público. E como todo texto tem efeitos sobre o *viewer*, não podemos encarar esse tipo de afirmação (banco de todos os brasileiros) como um “modo de falar”, mas como um recurso semiótico muito bem usado para despertar a emoção no *viewer* e fazê-lo aderir inconscientemente a essa ilusão de posse.

Cada modo semiótico possui seu efeito no significado total do texto multimodal. No caso do modo verbal, as unidades componentes daquilo que é construído são as orações. A oração, unidade gramatical do sistema de transitividade, casa de força da linguagem verbal nos termos de Halliday, comporta apenas um processo, que, de modo geral, determina

os tipos de participantes. Caso haja mais de um processo, ou haja vários processos, teremos um complexo oracional. Por sua vez, as imagens podem realizar vários processos ao mesmo tempo, o que consideramos mais perto de situações gerais da comunicação humana, que utiliza simultaneamente os recursos de vários sistemas semióticos. Em uma imagem, podemos ter realizados processos mentais e materiais. Na análise dos comerciais, foi possível vermos vários processos realizados simultaneamente pelos quadros que compõem o filme. Isso indica uma capacidade maior de comunicação da imagem, o que ratifica a velha, batida, mas verdadeira frase feita: uma imagem vale mais que mil palavras.

As imagens visuais carregam de forma mais intensa a dimensão emocional. Dessa forma, os processos mentais são “mais bem realizados” pelas imagens do que pela linguagem verbal. As imagens dos quadros do filme são portadoras da emotividade dos processos mentais como apreciar, gostar, amar, para ficarmos em uma gradação. Nesse caso, a saturação das cores, a distância do mundo retratado para o *viewer*, o brilho dos olhos e o sorriso dos participantes do filme carregam de forma mais eficiente os traços de emotividade (afeto) dos processos mentais – o olhar de felicidade da gerente e do cliente em seu aperto de mão dizem muito mais do que a oração: “Você é muito bem-vindo nesta instituição” ou “Estou feliz por estar neste banco”.

A expressão corporal (incluída a expressão facial) é responsável por grande parte do sentido da mensagem do comercial e, sem ela, mais da metade do significado teria ido embora. A linguagem verbal, nos comerciais analisados, serve para se integrar e ratificar o significado das imagens visuais, nesse caso mais capazes de realizar os sentidos do afeto, no sistema de avaliatividade, emoções no *viewer*.

Finalmente, devemos discorrer um pouco, neste momento conclusivo do texto, sobre a transcrição multimodal de filmes. Esse tipo de transcrição esbarra em algumas dificuldades e a principal delas refere-se à impossibilidade de transcrever o movimento real nas imagens, a menos que a tecnologia se desenvolva a tal ponto que já nos permita inserir um vídeo completo em uma revista tradicional de papel, por exemplo. Mesmo se nos lembrarmos de que um vídeo pode ser inserido em um artigo publicado em uma revista eletrônica, devemos ter em mente também que esse tipo de inserção ainda requer grande quantidade de memória e uma excelente rede de transmissão de dados, o que pode inviabilizar esse intento. Mas essa “deficiência” é compensada pelo processo de transcrição e análise, visto que somente na matriz de recursos semióticos poderemos ficar cientes da integração dos vários modos de expressão. Ao assistirmos ao comercial na televisão, não percebemos a simultaneidade dos recursos semióticos utilizados. Assim, em uma transcrição multimodal, de certa forma fazemos o caminho inverso daquele feito pelo produtor do vídeo, explicitando as interações entre imagem, texto no modo fala, olhar dos participantes, movimento dos participantes, cores etc.

Outro impasse nas transcrições multimodais é o fato de que sempre nos expressaremos em linguagem verbal na transcrição, não havendo outra forma de fazê-lo obviamente, uma vez que não podemos explicar imagens utilizando imagens. Sendo assim, as transcrições de textos multimodais serão sempre traduções ou transduções, uma transformação modal, uma vez que transformamos um texto multimodal em um texto em que prevalece a linguagem verbal escrita. Kress diz o seguinte:

A transdução é vista como subordinada à – um tipo de – tradução. Ela nomeia o processo de mover o material-significado de um modo para outro – da fala para a imagem; da escrita para o filme. [...] Esse processo implica uma – geralmente total – rearticulação do significado das entidades de um modo para as entidades do novo modo (KRESS, 2010, p. 125).⁴

Por isso, não são bem definidos os limites entre análise e transcrição. Melhor dizendo, uma transcrição multimodal já carrega em si a análise, porque o pesquisador define certos rumos e não outros para explicar as imagens. Sobre isso, afirmam Flewitt, Hampel, Hauck e Lancaster:

O que é transcrito depende do contexto de pesquisa e do que o pesquisador está tentando encontrar, mas as transcrições devem ser consideradas como versões reduzidas da realidade observada, na qual alguns detalhes são priorizados e outros são deixados de lado (FLEWITT et al, 2009, p. 49).⁵

É relevante enfatizarmos o caráter parcial de toda transcrição. Se considerarmos que a transcrição é uma espécie de teoria parcial de uma realidade objetiva, necessariamente verificaremos que toda e qualquer transcrição carrega em si uma escolha de ênfase. Essa escolha pode significar que o papel do transcritor não é neutro, mas carregado de decisões que podem influir na interpretação de um texto multimodal.

4. Transduction is seen as subordinate to – as one kind of – translation. It names the process of moving meaning-material from one mode to another – from speech to image; from writing to film. [...] This process entails a (usually total) re-articulation of meaning from the entities of one mode into the entities of the new mode.

5. Transcription depends on research context and must be viewed as ‘reduced versions of observed reality where some details are prioritized and others are left out’.

Referências

BALDRY, A; THIBAUT, P. J. *Multimodal transcription and text analysis*. A multimedia toolkit and coursebook with associated on-line course. London: Equinox, 2006.

FLEWITT, R; HAMPEL, R; HAUCK, M; LANCASTER, L. What are multimodal data and transcription? In: JEWITT, C. *The Routledge handbook of multimodal analysis*. London: Routledge, 2009.

HALLIDAY, M; MATTHIESSEN, M. *An introduction to functional grammar*. London: Hodder Education, 2014.

IEDEMA, R. Multimodality, ressemiotization, extending the analysis of discourse as multi-semiotic practice. In: *Visual communication*. Stanford. vol. 2, n. 1, 2003, p. 29-57.

KRESS, G. *Multimodality: A social semiotic approach to contemporary communication*. London: Routledge, 2010.

KRESS, G; van LEEUWEN, T. *Reading Images: The Grammar of Visual Design* (second edition). London: Routledge, 2008.

van LEEUWEN, T. *Introducing Social Semiotics*. London: Routledge, 2006.

VIEIRA, J; SILVESTRE, C. *Introdução à Multimodalidade: Contribuições da Linguística Sistêmico-Funcional, Análise de Discurso Crítica e Semiótica Social*. Brasília: CEPADIC, 2015.

YARED, M.L.M. *A ação semiótico-social da publicidade governamental sob a perspectiva da Análise de Discurso Crítica e da Multimodalidade*. Tese de Doutorado defendida na Universidade de Brasília em 2015.