

## Análise de chamada televisiva: multimodalidade e metafunções da linguagem<sup>1</sup>

Douglas Vidal Santiago<sup>2</sup>

### RESUMO

Este artigo tem como objetivo analisar as categorias das três metafunções da linguagem presentes na obra "A Gramática do Design Visual", de Kress e van Leeuwen (1996, 2006), isto é, as metafunções representacional, interacional e composicional, pela perspectiva da multimodalidade (KRESS, 2010; VAN LEEUWEN, 2011; LIMA-LOPES, 2012; RAMOS; CARMELINO, 2014; SILVA, 2016; KNOLL; FUZER, 2019), em uma chamada televisiva jornalística. Desse modo, o presente estudo tem como finalidade relacionar os modos semióticos com o valor social que trazem, observar em que medida os diferentes signos se articulam, explicitar os possíveis efeitos de sentido da integração de elementos de ordem verbo-visual, atentando-se para esses elementos, para o contexto sócio-histórico e para o propósito comunicativo do gênero em questão. Para tanto, propõe-se analisar a chamada para as edições especiais do Jornal da Cultura, da retrospectiva 2020, a partir das anotações da correlação de diferentes sistemas semióticos das produções multimodais em movimento utilizando o *software* ELAN (SLOETJES; WITTENBURG, 2008). Os resultados do estudo demonstraram que a chamada televisiva é um texto multimodal que agrega várias linguagens que juntas são processadas semioticamente pelo telespectador para a produção de sentido. Conclui-se, portanto, que o *corpus* analisado neste estudo é um texto audiovisual bastante produtivo que possibilitou dar visibilidade a maneira como os diferentes modos de linguagem são conjugados de forma a construir o significado em textos multimodais, permitindo observar minuciosamente a relação direta entre os estudos da multimodalidade e os de caráter linguísticos, corroborando à teoria sociosemiótica da verbo-visualidade.

**PALAVRAS-CHAVE:** Gramática do design visual. Metafunções da linguagem. Multimodalidade. Semiótica Social. Chamada televisiva.

---

<sup>1</sup> Este artigo foi a produção de avaliação final da disciplina *Multimodalidade e Construção de Sentidos no Meio Digital* (LP136) ministrada pelo Prof. Dr. Petrilson Pinheiro, no segundo semestre de 2020, do Programa de Pós-Graduação em Linguística Aplicada do Instituto de Estudo da Linguagem da Universidade Estadual de Campinas (IEL-UNICAMP).

<sup>2</sup> Doutorando pelo Programa de Pós-Graduação em Linguística do Instituto de Estudos da Linguagem da Universidade Estadual de Campinas (IEL-UNICAMP). Mestre em Letras (Estudos Linguísticos) pela Escola de Letras, Filosofia e Ciência Humanas da Universidade Federal de São Paulo (EFLCH-UNIFESP). Lattes: <http://lattes.cnpq.br/3856100803271323>. E-mail: [d233447@dac.unicamp.br](mailto:d233447@dac.unicamp.br). ORCID: <https://orcid.org/0000-0002-0998-9623>.

## 1. CONSIDERAÇÕES INICIAIS

Hodiernamente, nos meios de comunicação, o grande uso de imagens salienta a necessidade de compreendermos a significação da visualidade nos diferentes textos que circulam no contexto social. Os textos que mesclam a verbo-visualidade são conhecidos como multimodais, isto é, aqueles que reúnem dois ou mais modos semióticos<sup>3</sup> (Kress; Van Leeuwen, 2006), abrangendo assim diferentes modos de significação e uma multiplicidade de aspectos que compõem um universo sociocultural, como salientam Knoll e Fuzer (2019).

O objetivo do presente estudo é analisar as categorias das três metafunções da linguagem (Kress; Van Leeuwen, 1996, 2006), ou seja, as metafunções representacional, interacional e composicional, pela perspectiva da multimodalidade em uma chamada televisiva jornalística. Sendo assim, o estudo tem a finalidade de relacionar os modos semióticos com o valor social que trazem, além de observar em que medida os diferentes signos se articulam, explicitar os possíveis efeitos de sentido da integração de elementos de ordem verbo-visual, atentando-se para esses elementos, para o contexto sócio-histórico e para o propósito comunicativo do gênero em questão.

A justificativa por estudar os textos multimodais se dá pela atualidade do tema, haja vista que os multiletramentos<sup>4</sup> têm como fito compreender o texto enquanto artefato constitutivo das relações sociais. Segundo Knoll e Fuzer (2019), ainda que, de um modo geral, a leitura de imagens seja apreendida de forma intuitiva pelas pessoas, persistem dificuldades na composição e na interpretação das mensagens em gêneros textuais que integram as práticas sociais cotidianas.

Com o interesse de minimizar tal problemática, *A Gramática do Design Visual*<sup>5</sup> surge assim como uma teoria com relevância que busca elaborar categorias analíticas para compreender o uso de imagens, tanto estáticas, quanto móveis, em textos de caráter multimodal.

De acordo com Ramos e Carmelino (2014), o texto multimodal tende a se apresentar de duas formas: uma fixa e outra em movimento. Segundo os autores, as

---

<sup>4</sup>O trabalho produzido pelo The New London Group (1996) introduziu a noção dos multiletramentos, argumentando que todas as formas de fazer sentido são multimodais. Destarte, a linguagem é fundamentalmente multimodal.

<sup>5</sup> Título original da obra em inglês: *Reading Images: The Grammar of Visual Design*, de Gunther Kress e Theo van Leeuwen (1996, 2006).

fotografias, as publicidades publicadas nas revistas, as histórias em quadrinhos e tantos outros exemplos possíveis serviriam, assim, para ilustrar o primeiro caso. Quanto à segunda forma, o rol de possibilidades também não é pequeno. Seriam os casos das animações, dos seriados de TV, dos filmes e dos variados vídeos que circulam pelas mídias virtuais. Desse modo, a chamada televisiva para as edições especiais do Jornal da Cultura<sup>6</sup> para a retrospectiva 2020, o *corpus* escolhido para análise neste artigo, se enquadra, evidentemente, como um tipo de produção multimodal em movimento. Esta chamada jornalística se constitui de uma mescla de modos, sendo eles: i. a oralidade (falas e canção), ii. os sons (ruídos, instrumental, melodia), iii. a visualidade (imagens fixas e em movimento de fatos contemporâneos) e iv. a escrita (legendas, títulos), sendo, essencialmente, um texto multimodal móvel. Noutras palavras, "trata-se de uma produção que reúne em si diferentes modalidades" (Ramos; Carmelino, 2014, p. 64).

A chamada televisiva é um gênero que faz parte do contexto social nas esferas midiáticas e tem sido disseminada, atualmente, também nos meios digitais de comunicação, tais como redes sociais e plataformas de compartilhamento de vídeo.

O leitor, ou, mais especificamente, o telespectador, precisa compreender as marcas do gênero em questão (a chamada televisiva jornalística) e os seus propósitos comunicativos, ler as imagens, bem como os textos verbais orais e escritos apresentados de modo a produzir sentido. Desse processo, "que alia tanto elementos verbais escritos e orais, quanto de ordem visual, seria construída a significação do texto multimodal" (Ramos; Carmelino, 2014, p. 68).

A chamada é um texto multimodal relativamente curto que antecipa os assuntos de destaque do telejornal com o propósito de despertar a atenção do telespectador. Frequentemente as chamadas são transmitidas dentro da programação regular de um determinado canal televisivo, no entanto, com o advento das mídias digitais, este gênero passou também a ser compartilhado pelos portais oficiais das emissoras na *internet*, bem como pelos usuários.

Conforme podemos observar, a chamada televisiva para a retrospectiva é um gênero multimodal que dialoga, evidentemente, com assuntos jornalísticos

---

<sup>6</sup>A Fundação Padre Anchieta é uma rede midiática "custeada por dotações orçamentárias legalmente estabelecidas e recursos próprios obtidos junto à iniciativa privada, a Fundação Padre Anchieta mantém uma emissora de televisão de sinal aberto, a TV Cultura; uma emissora de TV a cabo por assinatura, a TV Rá-Tim-Bum; e duas emissoras de rádio: a Cultura AM e a Cultura FM. Situada na Rua Cenzo Sbrighi, 378 - Caixa Postal 66.028 CEP 05036-900. São Paulo/SP - Tel: (11)2182-3000". Disponível em: <<https://cultura.uol.com.br/>>. Acesso em: 05 de jan. de 2023.

contemporâneos e exige o acionamento de conhecimentos prévios por parte do telespectador à sua realidade sociocognitiva. Desse modo, o mecanismo inferencial é elemento imbricado com o ato de leitura (Ramos; Carmelino, 2014). Segundo Marcuschi (2008), compreender um texto é “realizar inferências a partir de informações dadas no texto e situadas em contextos mais amplos” (op.cit., p.239). Sendo assim, a compreensão pressupõe o processo cognitivo e o acionamento de uma série de informações armazenadas na memória dos sujeitos das mais variadas formas de interação. Noutros termos, o ato sociocomunicacional prevê conhecimentos de diferentes ordens: linguístico, enciclopédico (ou de mundo), imagético, compartilhado e multimodal.

Neste sentido, a chamada televisiva jornalística, por exemplo, exige um grau elevado de inferências por parte do telespectador, sendo necessário estabelecer todas as relações possíveis e preencher as lacunas por meio de inferências. Esse processo cognitivo de produção de sentido deve ser feito levando-se em consideração os conhecimentos das metafunções da linguagem (representacional, interacional e composicional), a situacionalidade e os dados explícitos na superfície multimodal. Estes termos são constitutivos da perspectiva da Semiótica Social (*Social Semiotics*), presentes na obra *Gramática do Design Visual*, de Kress e van Leeuwen (1996, 2006), cuja teoria visa a elaborar categorias de análise visual para relacionar os estudos da multimodalidade com os estudos linguísticos.

A vertente social da semiótica (ou sociosemiótica) defende que as produções multimodais, assim como as verbais, devem ser lidas sempre dentro de uma situação de ação comunicativa para serem compreendidas, observando as trocas de mensagens de maneira vinculada com o contexto social de uso. Ademais, nesta abordagem, entende-se que a escolha dos signos e a construção dos textos são movidas por propósitos comunicativos específicos. Os autores apreendem assim que o signo é motivado, isto é, "qualquer modo, qualquer forma de comunicação parte de uma base, de uma necessidade social e, portanto, está relacionado aos interesses de instituições e de seus produtores, às ideologias e a questões de poder" (KRESS; VAN LEEUWEN, 1996, 2006 apud SILVA, 2016, p. 62).

Com base nesse arcabouço teórico, propomos analisar a chamada televisiva, *corpus* do presente estudo, a partir das anotações da correlação de diferentes sistemas semióticos das produções multimodais em movimento utilizando o *software Eudico*

*Linguistic Annotator*, ferramenta profissional que possibilita a criação de anotações complexas em arquivos digitais de áudio e vídeo, com o fito de compreender as estratégias de como os variados recursos são empregados para a construção do significado.

Este artigo está composto pelas seções de fundamentação teórica: 2.a multimodalidade em uma perspectiva da Semiótica Social, 2.1.as três metafunções da linguagem; metodologia: 3.o tratamento do dado audiovisual e o uso da ferramenta *Eudico Linguistic Annotator*(ELAN); análise: 4.a multimodalidade na chamada televisiva jornalística; e, por fim, 5.considerações finais.

## **2. A MULTIMODALIDADE EM UMA PERSPECTIVA DA SEMIÓTICASOCIAL**

No campo dos estudos linguísticos, as pesquisas em multimodalidade vêm ganhando um grande espaço em distintas perspectivas teórico-analíticas: textual, interacional, cognitiva, psicolinguística, para citar algumas. Neste mesmo viés, as investigações em Linguística Aplicada (LA), área fundamentalmente inter e multidisciplinar com influências teóricas vindas da Pedagogia, Psicologia, Filosofia e Linguística Teórica (Lima-Lopes, 2012), tem se destacado pela preocupação em investigar como os diferentes modos de linguagem se combinam no processo de formação de significado, ou seja, como acontece a interação entre esses variados modos semióticos.

A noção de multimodalidade surgiu em 1920 como um termo técnico no campo da psicologia da percepção, denotando os efeitos que diferentes percepções sensoriais têm uns sobre outros (VAN LEEUWEN, 2011). Na LA, conforme Lima-Lopes (2012), a perspectiva multimodal originou-se a partir das pesquisas realizadas pelo linguista funcionalista britânico Halliday (1985, 1994), sobretudo dentro do estudo de Letramento e da Formação de Professores de Línguas, e teve seu desdobramento em duas vertentes complementares:

- i. A primeira delas tem Kress e van Leeuwen (1996, 2006) e Lemke (1998) como principais autores, cujos estudos advogam que "o desenvolvimento tecnológico tem criado novas formas de significação textuais que, de certa forma, causam uma perda de *status* por parte da escrita" (LIMA-LOPES,

2012, p. 24). Isto é, uma vez que vivemos em uma sociedade dinâmica e multimidiática, a escrita "tem perdido seu poder [...] como forma preferencial de divulgação do conhecimento chamado letrado" (op. cit., p. 24). Assim, esses pesquisadores da linguagem passaram a desenvolver trabalhos teóricos que possibilitaram a expansão da sociosemiótica para outros modos de linguagem para além da verbal. Já em *AGramática do Design Visual* (GDV), Kress e van Leeuwen (1996, 2006) advogam “a necessidade de se desenvolver métodos de análise textual que contemplem as imagens como modo semiótico para que se possam descrever todos os significados veiculados textualmente” (ARAÚJO, 2011, p. 14). Nesta perspectiva, “a língua constitui um sistema semiótico que está em interação com outros sistemas que, de forma conjunta, constituem o significado” (LIMA-LOPES, 2012, p. 18). Descrever a língua é, portanto, descrever um conjunto específico de recursos semióticos atribuindo sentido a todos os modos de comunicação.

- ii. A segunda vertente refere-se então à *Teoria Multimodal* de Kress (2010), que buscava independência das categorias analíticas da *Gramática Sistêmico-Funcional* (GSF) visando a ampliar a forma de compreensão do processo de comunicação humana, mas com base epistêmico-teórica hallidiana, cujos estudos orientam-se ainda hoje pelo significado ampliado por Halliday em 1994.

A GSF, modelo de análise linguística com forte vocação semântico-funcional, tem como preocupação principal a forma como a linguagem é estruturada para o uso (Eggins, 1994, p. 23-24 apud Lima-Lopes, 2012, p. 26), buscando compreender como os vários recursos são empregados para produção do significado.

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 apud LIMA-LOPES, 2012, p. 26).

Halliday (1994) elege assim o texto como a unidade mínima de significado, na qual é possível observar a manifestação de estruturas e da gramática de uma determinada língua (espécime), pensando-se assim no sistema per se, ou como os vários recursos foram empregados para produção do significado (artefato). Sendo assim, essas duas perspectivas são indissociáveis e contribuem para a compreensão da real função que um texto deve ter. A GSF se preocupa, portanto, com a linguagem em sua totalidade, argumentando que qualquer escolha no nível estrutural reflete em uma escolha no nível semântico e vice-versa, defendendo, assim, a não-arbitrariedade do signo. Noutros termos, toda forma linguística é usada de uma maneira mediadora e não arbitrária na expansão do significado. A este conjunto de possibilidades significativas de uma seleção em um dado contexto, Halliday (1994) nomeia de potencial semiótico (ou potencial significado).

Conforme observamos, a escolha tem um papel importante nessa teoria, visto que se modificarmos qualquer estrutura linguística, modificamos, conseqüentemente, seu significado. Destarte "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" (LIMA-LOPES, 2012, p. 30).

A abordagem multimodal expande essa teoria estratificadora dos diferentes níveis de escolha a outros sistemas semióticos (visual, gestual, videográfico ou qualquer outro), demonstrando que significado e estrutura "estão diretamente interligados, são interdependentes e motivados por objetivo social da comunicação" (HALLIDAY; MATHIESSEN, 2004, p. 3-4 apud LIMA-LOPES, 2012, p.27). As noções da Linguística Sistêmico-Funcional têm sido assim aplicadas a outros modos de representação com ênfase na dimensão social, sob a ideia da linguagem como recurso para se construir significados que exerçam funções sociais.

Mais especificamente, em *AGramática do Design Visual*, Kress e van Leeuwen (1996, 2006) apresentam sistematicamente orientações para a leitura de imagem. Nesta obra, a perspectiva de multimodalidade está filiada, como dito, à Semiótica Social, que está calcada na produção e recepção de significados visuais, sendo de indiscutível relevância para a compreensão da comunicação visual orientada por um viés social. Assim, o foco dessa teoria semiótica está no processo de produção de signos em que a forma e o sentido são relativamente independentes antes de serem trazidos pelo produtor do signo para se tornar um novo signo produzido, isto é, "o que o produtor de *Revista Gatilho*, Juiz de Fora, v. 25, p. 290-317, 2023 – ISSN: 1808-9461

signos quer dizer e como dizê-lo, em qualquer meio, não se limita ao sistema, e sim, abrange suas funções e contextos” (SILVA, 2016, p. 52).

Portanto, ao pensarmos em semiótica social, precisamos ter em mente que: a) indivíduos, com suas histórias sociais, socialmente formados, situados em ambientes sociais, com recursos culturalmente disponíveis são vistos como agentes na produção de significados e na comunicação; b) os signos são sempre produzidos em uma interação social; c) os signos são motivados e nunca arbitrárias relações de significado e de forma; d) a relação motivada de forma e de significado está sempre baseada nos interesses dos produtores de signos; e) as formas/significantes que são usadas na produção de significados são feitas numa interação social e se tornam parte dos recursos semióticos de uma dada cultura; f) todos os signos são metafóricos; g) em uma perspectiva multimodal da semiótica social, todos os signos em todos os modos são significativos; h) poder diz respeito a relativas posições sociais (KRESS; VAN LEEUWEN, 1996, 2006; KRESS, 2010 apud SILVA, 2016, p. 53).

Como podemos notar, portanto, para a Semiótica Social não existe texto sem discurso, tampouco discurso sem texto, há assim entre ambos uma relação constitutiva (Knoll; Fuzer, 2019).

Neste artigo, para análise de uma chamada de televisão de cunho jornalístico, utilizamos as três metafunções da linguagem (representacional, interacional e composicional) da *Gramática do Design Visual*, conforme veremos na seção seguinte.

### 1.1 AS TRÊS METAFUNÇÕES DA LINGUAGEM

Sem dúvida, a *Gramática do Design Visual* (Kress; Van Leeuwen, 1996, 2006) é hoje um dos estudos mais importantes na descrição da estrutura que organiza a informação visual dos textos. Assim, como verificamos anteriormente, do mesmo modo que a gramática da língua, a gramática da comunicação visual pode ser descrita como um sistema de escolhas semântico-funcionais.

As metafunções da linguagem da GDV (Kress; Van Leeuwen, 1996, 2006) derivam da GSF propostas por Halliday (1985, 1994): as metafunções ideacional, interpessoal e textual, instrumentos de análise que têm como finalidade observar como a linguagem se organiza funcionalmente.

A metafunção ideacional refere-se ao modo como se organizam as experiências de mundo na linguagem; a metafunção interpessoal

*Revista Gatilho*, Juiz de Fora, v. 25, p. 290-317, 2023 – ISSN: 1808-9461



abrange o modo como os participantes da interação se relacionam socialmente por meio da linguagem (as formas de aproximação ou afastamento em relação ao leitor, por exemplo); já a metafunção textual corresponde aos modos de organização do texto (HALLIDAY, 1985). Kress e van Leeuwen (2006) trataram essas metafunções como representacional, interacional, e composicional, respectivamente (KNOLL; FUZER, 2019, p. 591).

Conforme vimos, Kress e van Leeuwen (1996, 2006) assumiram, em sua gramática do *design* visual, que, da mesma maneira que a língua, a comunicação visual pode realizar as metafunções.

Sumariamente, podemos definir as três metafunções da seguinte forma:

*i. Metafunção Representacional:*

Metafunção ideacional é definida como “a habilidade dos sistemas semióticos de representar objetos e suas relações em um mundo fora do sistema representacional ou nos sistemas semióticos da cultura” (KRESS; VAN LEEUWEN, 1996, 2006, p. 47)<sup>7</sup>. No campo das imagens, os vetores (*vectores*) se encarregam do processo de ação e de interação entre os participantes, que são diferenciados no plano representacional em dois tipos:

- **os interativos** (*interactive participants*), aqueles presentes na ação de comunicação, isto é, que falam e escutam ou que escrevem e leem; os que produzem imagens ou as visualizam;
- **os representados** (*represented participants*), que constituem o assunto/problema da comunicação; ou seja, pessoas, lugares e coisas (incluindo as abstratas) representadas em e pela fala ou escrita ou nas imagens produzidas (Kress; Van Leeuwen, 1996, 2006, p. 47).

Além dessas duas subdivisões, a metafunção representacional é desmembrada também em representações narrativas e conceituais.

Por um lado,

nas **representações narrativas**, os participantes estão sempre envolvidos em eventos e ações. Os vetores indicam a ação contida nos fatos apresentados em uma representação imagética, podendo ser

---

<sup>7</sup> "The ability of semiotic systems to represent objects and their relations in a world outside the representational system or in the semiotic systems of culture" (KRESS; VAN LEEUWEN, 1996, 2006, p. 47).

representados através de setas ou até mesmo pelo posicionamento dos participantes representados ou objetos, e que leva o olhar do observador para determinado ponto da imagem (BRITO; PIMENTA, 2009 apud SILVA, 2016, p. 64).

Kress e van Leeuwen (1996, 2006) ressaltam que tais características de direcionalidade devem estar presentes, caso a estrutura realize uma representação narrativa. Neste sentido, significa algum evento como “está conectado a”, “está conjugado a”, “está relacionado a” (KRESS; VAN LEEUWEN, 1996, 2006, p. 59 apud SILVA, 2016, p. 64)

Por outro lado,

Nas **representações conceituais**, a imagem não se apresenta como uma narrativa, mas representa uma relação de taxonomia entre os seus participantes. Pelo menos um conjunto de participantes fará o papel de subordinado (*Subordinate*) em relação a pelo menos um outro participante, o superordinado (*Superordinate*). Desse modo, essas representações ocorrem de forma classificacional, analítica ou simbólica. Segundo Kress e van Leeuwen (1996, 2006), representações conceituais representam participantes em termos de sua essência mais ou menos generalizada. Classificação conceitual, conforme os autores supracitados é, assim, representada pela mesma estrutura como na hierarquia social (KRESS; VAN LEEUWEN, 1996, 2006, p. 59 apud SILVA, 2016, p. 67, grifos nossos).

## *ii. Metafunção Interacional:*

A comunicação visual apresenta base para construir e manter outros tipos de interação entre o produtor e o observador da imagem. Deste modo, como ressalta Silva (2016), participantes interativos são pessoas reais que produzem sentido com imagens em contextos de instituições sociais que, em diferentes graus e de diferentes maneiras, regulam o que pode ser dito com imagens, como deve ser dito e como deve ser interpretado (Kress; Van Leeuwen, 1996, 2006).

A metafunção interacional estabelece estratégias de aproximação ou de afastamento do produtor do texto em relação ao seu leitor (um participante que é exterior à imagem), buscando estabelecer um elo imaginário entre ambos (Silva, 2016, p. 69).

Segundo Silva (2016, p. 70), Kress e van Leeuwen (1996, 2006) classificam as imagens em três dimensões: *i.* o olhar, *ii.* o enquadramento e *iii.* a perspectiva.

i. O olhar (*The gaze*) é dividido em: uma imagem de demanda (*demand*), aquela em que o participante representado (aquele presente na imagem) se coloca olhando diretamente para o leitor; e uma imagem de oferta (*offer*), a qual se dirige ao leitor de forma indireta, que cumprirá o papel de um observador invisível.

ii. O enquadramento (*Size of frame*): a escolha de distância pode sugerir diferentes relações entre os participantes representados e os observadores. Kress e van Leeuwen (2006) dividem essa segunda dimensão de significados interativos como plano fechado (*close-up/close shot*), plano médio (*medium shot*) e plano aberto (*long shot*).

iii. A perspectiva (*Perspective*): refere-se ao trabalho com a imagem através de um ângulo específico, de um determinado ponto de vista, a partir do qual os participantes são representados e que indica uma atitude mais ou menos subjetiva por parte do produtor da imagem em relação aos participantes (Silva, 2016, p. 72).

### iii. *Metafunção Composicional:*

A Metafunção Composicional, assim como auto explicitado em sua terminologia, se refere a composição do todo, da forma como os elementos representacionais e interativos são postos para se relacionarem uns com os outros e como se integram dentro de um todo de significados. Essa metafunção tem como papel central organizar/combinar os elementos visuais de uma imagem, ou seja, integrar os elementos representacionais e interativos em uma composição para que ela faça sentido. De acordo com a teoria, os três sistemas que se inter-relacionam, e por meio dos quais tais elementos realizam significados composicionais, são: o valor da informação (*Information value*), o enquadramento ou a estruturação (*Framing*) e a saliência (*Saliency*) (Kress; Van Leeuwen, 1996, 2006; Almeida, 2009, p. 23).

As metafunções oferecem base para uma análise estrutural das imagens, mas não se restringe apenas a ela, uma vez que, por conta do caráter social da abordagem proposta, os elementos imagéticos são vistos como portadores de significados potenciais que podem ser lidos, interpretados, acordados ou não pelos seus leitores. Ou seja, “as estruturas visuais não são meramente formais: elas têm uma dimensão semântica profundamente importante” (KRESS; VAN LEEUWEN, 1996, 2006, p. 47)<sup>8</sup>.

---

<sup>8</sup> “Visual structures are never merely formal: they have a deeply important semantic dimension” (KRESS; VAN LEEUWEN, 1996, 2006, p. 47).

Apresentamos a seguir os procedimentos metodológicos realizados para este estudo com o fito de compor o registro audiovisual para a análise das metafunções.

### **3. O TRATAMENTO DO DADO AUDIOVISUAL E O USO DA FERRAMENTA *EUDICO LINGUISTIC ANNOTATOR (ELAN)***

Em sua tese de doutorado, Lima-Lopes (2012) propõe uma metodologia de análise para comunicação multimodal em produções audiovisuais, salientando as principais contribuições para a análise de vídeos multimodais com base nos conceitos epistêmico-teóricos da teoria multimodal. De acordo com ele,

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 (LIMA-LOPES, 2012, p. 34).

O pesquisador sugere assim que os elementos técnicos, como as ações efetivas dentro da imagem em movimento, sejam levados em consideração no processo de análise e compreensão de como o significado é construído em textos audiovisuais. Além disso, ressalta que é necessário 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.

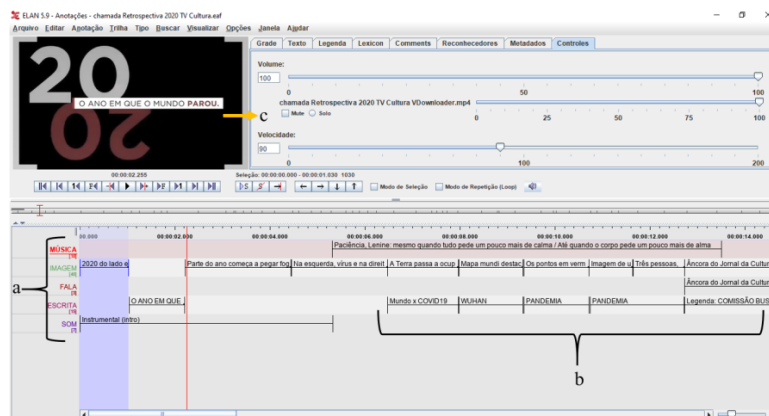
A necessidade de desenvolvimento deste tipo de modelo se justifica não apenas pela construção de modelos analíticos que nos permitam compreender a construção de tais textos, mas também pela premência no desenvolvimento de modelos de letramento que deem conta de uma realidade midiática que congrega diversas formas de expressão de significados, todas elas motivadas culturalmente (LIMA-LOPES, 127-128).

Como podemos notar, há uma série de estratégias que podem ser mobilizadas na compreensão e produção do sentido de textos multimodais. Particularmente neste estudo, para realizar as anotações do material audiovisual, fizemos o *download* do arquivo da plataforma de compartilhamento de vídeos *YouTube* em formato *.mp4* (extensão de arquivo de vídeo) para deixar o dado apto a uma transcrição. Deste modo, para anotação da correlação de diferentes sistemas semióticos das produções

multimodais em movimento, utilizamos o *software* ELAN (*Eudico Linguistic Annotator- Version 5.9*)<sup>9</sup>, ferramenta profissional que possibilita a criação de anotações complexas em arquivos digitais de áudio e vídeo. Este programa para anotações foi desenvolvido pelo Instituto de Psicolinguística *Max Planck* dos Países Baixos da Europa (*Max Planck Institute for Psycholinguistics, The Language Archive, Nijmegen, The Netherlands*) (Sloetjes; Wittenburg, 2008).

Conforme notamos na *figura 1*, a tela do *software* ELAN, em **a** temos as trilhas com os modos semióticos presentes no registro audiovisual selecionado (música, imagem, fala, escrita e som), sendo, quantitativamente, 10 anotações de trechos da música (da canção “Paciência”<sup>10</sup>), 40 imagens (estáticas e em movimento), 03 falas de pessoas públicas (de uma jornalista e do ex-presidente da república, Jair Messias Bolsonaro), 15 mensagens escritas (títulos e legendas) e 07 sons (instrumental e ruídos); em **b** encontram-se as anotações realizadas de acordo com o vídeo, que está representado em **c**.

**Figura 1:** A coordenação de recursos multimodais na tela do *software* ELAN (*Eudico Linguistic Annotator*).



Após a preparação do vídeo para exploração e investigação multimodal a partir do subsídio do ELAN, extraímos as anotações em *.eaf* (extensão de arquivo do *software* ELAN) para o formato texto, bem como as capturas de tela (uso da ferramenta

<sup>9</sup> Disponível em: <<https://archive.mpi.nl/tla/elan>>. Acesso em: 06 de jan. de 2023.

<sup>10</sup> Compositores: Carlos Eduardo Carneiro De Albuquerque Falcão / Oswaldo Lenine Macedo Pimentel. Letra de Paciência © Altafonte, Universal Music Publishing Group.

*LightShot*<sup>11</sup>) com as coocorrências dos diferentes modos semióticos que compõem a chamada televisiva aqui analisada, e, por fim, realizamos a análise multimodal levando em conta as categorias analíticas definidas, ou seja, as metafunções da linguagem representacional, interacional e composicional.

É importante salientar que a seleção do dado e sua anotação multimodal já fazem parte do empreendimento analítico, uma vez que o pesquisador, nesse tipo de estudo, necessita realizar recortes de acordo com seus propósitos investigativos.

O critério de seleção do vídeo para análise foi estabelecido com o pressuposto de que o registro audiovisual deveria conter uma mescla de distintos modos semióticos, tais como: sons, falas, imagens – estáticas e em movimento, gestos, escrita, dentre outros, com o fito de dar visibilidade à forma como os diferentes modos de linguagem são conjugados de forma a construir o significado em textos multimodais

A seguir, apresentamos inicialmente a análise descritiva desta produção multimodal em movimento, seguida da análise videográfica com base nas categorias analíticas selecionadas, a qual é acompanhada de recortes de trechos do registro audiovisual em capturas de tela.

#### **4. AMULTIMODALIDADE EM UMA CHAMADA TELEVISIVA JORNALÍSTICA**

A chamada de TV para a exibição da retrospectiva da TV Cultura<sup>12</sup> (Fundação Padre Anchieta)<sup>13</sup>, com os principais destaques do ano de 2020, foi exibida nos intervalos da programação da emissora no período de 14 a 31 de dezembro de 2020. A produção audiovisual consistia de uma divulgação das edições especiais do Jornal da Cultura que seriam exibidas nos dias 24 (primeira parte) e 31 de dezembro de 2020 (segunda parte) às 21h15<sup>14</sup>. De acordo com informação encontrada no site da rede midiática, "no ar de segunda a sábado, o Jornal da Cultura apresenta os principais fatos

<sup>11</sup> Disponível em: <<https://app.prtscr.com/pt-br/>>. Acesso em: 06 de jan. de 2023.

<sup>12</sup> Chamada televisiva editada sob o olhar do diretor de criação Henrique Bacana, que “desde 2013 é diretor de arte da TV Cultura e demais emissoras da Fundação Padre Anchieta, desenvolvendo campanhas institucionais, promos e identidades visuais do canal”. Disponível em: <<https://br.linkedin.com/in/henriquebacana>>. Acesso em: 06 de jan. de 2023.

<sup>13</sup> Disponível em: <<https://www.youtube.com/watch?v=NwZevQCf5T8>>. Acesso em: 07 de jan. de 2023. Copyright © 2020 TV Cultura. Todos os direitos reservados.

<sup>14</sup> Disponível em: <[https://cultura.uol.com.br/noticias/15020\\_retrospectiva-2020-jornal-da-cultura-traz-destaques-do-ano-em-edicoes-especiais.html](https://cultura.uol.com.br/noticias/15020_retrospectiva-2020-jornal-da-cultura-traz-destaques-do-ano-em-edicoes-especiais.html)>. Acesso em: 06 de jan. de 2023.

do Brasil e do mundo, sempre com comentários de especialistas"<sup>15</sup>. Além de exibido na televisão, o telejornal é transmitido ao vivo nos canais oficiais do *YouTube* e *Facebook*.

O texto verbo-visual selecionado para análise tem uma duração de um minuto e é composto por uma mescla de modos semióticos com saliência para as imagens (fixas e em movimento), contendo fatos históricos contemporâneos que ocorreram naquele ano vigente, para a canção "Paciência" e para a coocorrência entre os diferentes elementos multimodais, tais como falas de pessoas públicas (de uma jornalista e do ex-Presidente da República), sons (ruídos de bomba e de tiro, instrumentais) e escritas (legendas e títulos).

A chamada jornalística é iniciada com o instrumental da canção "Paciência"<sup>16</sup>, uma imagem refletida com o ano "2020" na margem esquerda da tela (o número 20 na parte superior na cor branca, e outro número 20 invertido na cor vermelha clara que escurece ao longo dos primeiros segundos do vídeo) e a legenda "o ano em que o mundo parou." em caixa alta centralizada na cor preta e com fundo branco; o item lexical "parou" aparece destacado em negrito na cor vermelha. Esses três modos semióticos, isto é, a música, a imagem e a mensagem escrita, aparecem quase que simultaneamente e aciona no telespectador uma expectativa, ou, em termos cognitivos, um processo cognitivo inferencial, de que aquela chamada trataria acerca dos fatos que ocorreram ao longo daquele ano.

Imediatamente aparece uma imagem estática em preto e branco de um protesto e parte do número que representa o ano 2020, mais especificamente o número 20 invertido na cor vermelha, dá lugar às chamas de um fogo ateado por manifestantes em uma bandeira dos Estados Unidos. Após essa introdução, a música começa a ser cantada (instrumental com letra) e aparece as imagens do planeta Terra, na margem direita, e de um vírus, na margem esquerda, que se conectam e ficam divididas meio a meio, formando, logo em seguida, um círculo completo com um título escrito MUNDO (na cor branca e em caixa alta) X COVID19 (cor preta e em caixa alta) com um plano de fundo vermelho na margem da tela. Em seguida, aparece um mapa *mundi* com destaque para Wuhan, a vasta capital da província da China central e o epicentro da

<sup>15</sup> Disponível em: <<https://cultura.uol.com.br/programas/jornaldacultura/>>. Acesso em: 05 de jan. de 2023.

<sup>16</sup> LENINE. FALCÃO, D. (1999, fevereiro) *Paciência. Álbum Na Pressão*. Disponível em: <<https://goo.gl/KzKACk>>. Acesso em: 06 de jan. de 2023. Sony BMG Music Entertainment. São Paulo, SP, Brasil.

COVID-19<sup>17</sup>, que se amplia e mostra todo o globo com alguns pontos marcados em vermelho. Rapidamente esses círculos em vermelhos se ampliam em diferentes países em todos os continentes e o vocábulo "pandemia" (escrito em branco, em caixa alta, com plano de fundo vermelho escuro) aparece no centro da tela juntamente coma imagem de fundo de um homem de nacionalidade chinesa completamente vestido com uma roupa branca, utilizando máscara e ajustando os óculos industriais em seu rosto.

Logo depois, aparece a imagem de duas pessoas também totalmente vestidas de branco, usando botas pretas e luvas amarelas, desinfetando o chão de um espaço público. Aparece então a jornalista Karyn Bravo, âncora do Jornal da Cultura, anunciando “o primeiro caso de COVID-19 no Brasil e na América Latina”. Em seguida, surge uma sucessão de imagens que nos remete a fatos ocorridos durante aquele ano, como: protestos civis no Chile (manifestantes ateando fogo e segurando bandeiras do país), no Brasil (confronto com um bloqueio de policiais militares, em que é possível ver e ouvir um deles atirando contra um grupo de pessoas), no Líbano (com um manifestante segurando com as duas mãos a bandeira desta nação), o porto do Líbano sendo destruído junto com um ruído alto de bomba explodindo, uma urna eletrônica brasileira com um eleitor votando em um candidato para vereador representando as eleições municipais, três imagens do jogador argentino Maradona que faleceu naquele ano, a imagem de uma mulher colocando uma máscara hospitalar, de um respirador pulmonar, de um cemitério com covas feitas para enterrar as vítimas de COVID19 uma ao lado da outra para que ocupassem menos espaço e coubesse o maior número de pessoas mortas pelo vírus durante o pico da doença, do centro de uma cidade vazio com lojas fechadas, do então Presidente da República Jair Messias Bolsonaro ao lado de dois ex-Ministros da Saúde exonerados de seu governo (os médicos Luiz Henrique Mandetta e Nelson Teich) e do vigente até aquela data (general Eduardo Pazuello), seguido da fala do Presidente brasileiro para uma coletiva de imprensa afirmando enfaticamente: "o povo tem que trabalhar!", de um policial americano sufocando um homem negro no chão com o joelho esquerdo (caso George Floyd), da série de protestos *Black Lives Matter (Vidas Negras Importam)* nos Estados Unidos, dos protestos no Brasil, influenciados pelas manifestações ocorridas nos EUA, contra a rede de supermercados *Carrefour* após um homem negro ter sido assassinado por dois

---

<sup>17</sup> Disponível em: <<https://www.brasildefato.com.br/2021/01/05/epicentro-da-covid-ha-um-ano-wuhan-e-hoje-a-cidade-mais-segura-relata-brasileiro>>. Acesso em: 05 de jan. de 2023.



seguranças, dos dois candidatos à eleição para presidência dos EUA, Donald Trump (Republicanos) e Joe Biden (Democratas), seguido do destaque para Biden em gesto de comemoração e a legenda "todo voto conta", das queimadas na floresta amazônica, de um contador com número de mortos pelo Coronavírus no Brasil aumentando ligeiramente em conjunto com fala do Presidente da República Jair Messias Bolsonaro dizendo: "quando muito, acometido de uma gripezinha" e uma legenda em caixa alta, que aparece simultaneamente no centro da tela, escrita: "você mentiu para nós", e, por fim, a imagem de um profissional de saúde aplicando a vacina no braço de uma outra pessoa.

Para finalizar o vídeo, reaparece na tela a imagem do ano 2020 em uma imagem refletida, assim como no início da chamada. No entanto, neste momento, o ano surge centralizado e não mais à esquerda, com o vocábulo em caixa alta "retrospectiva", aparecendo logo depois o logotipo da emissora TV Cultura (símbolo de uma pessoa de braços abertos na cor branca) e uma última mensagem grafada em um fundo preto, em caixa alta, nas cores vermelha e branca: "edição especial do Jornal da Cultura".

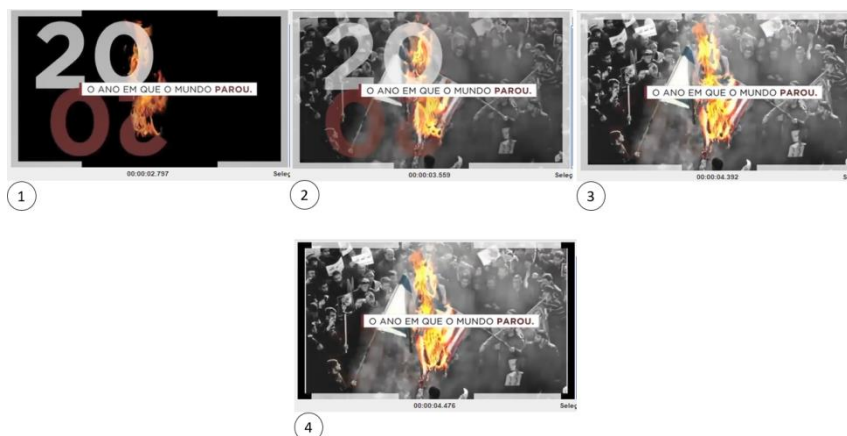
Após essa descrição verbal resumida do conteúdo presente na chamada televisiva, destacaremos 04 amostras do vídeo em que há coocorrências na articulação entre os modos semióticos, com o auxílio das anotações realizadas através da plataforma profissional para transcrição multimodal ELAN e das capturas de tela, com o objetivo de dar visibilidade a maneira como os diferentes modos de linguagem são conjugados de forma a produzir sentido nessa produção multimodal.

### **Amostra 01:**

Nesta primeira sequência imagética utilizada na introdução dos quatro primeiros segundos da chamada televisiva (*vide* figura 2), é possível perceber a chama de um fogo (1) substituir à ilustração do ano de 2020, em um plano de fundo preto, acompanhada de uma mensagem em caixa alta grafada em um fundo branco: "O ANO EM QUE O MUNDO **PAROU**.", com saliência para o verbo "parar" conjugado na terceira pessoa do singular do Pretérito Perfeito em cor vermelho falu.

Em menos de um segundo, o plano de fundo muda da cor preta para a imagem estática em preto e branco de manifestantes ao redor de uma bandeira dos Estados Unidos (sequência 2, 3 e 4), que está sendo queimada.

**Figura 2:** Introdução da chamada televisiva para a retrospectiva 2020 do Jornal da Cultura (primeiros segundos do vídeo).



Como pontuamos, há alguns pontos de saliência na composição dessa sequência imagética em cada um dos modos semióticos destacados. Na imagem (1), temos o destaque para o ano de 2020 e a presença de chamas, bem como na mensagem escrita há uma ênfase para o vocábulo “parou”. Representativamente, é plausível inferir que a cadeia referencial (Marcuschi, 2008) construída nestes primeiros momentos do vídeo, que tem um caráter altamente argumentativo, classifica o participante representado “2020” como “o ano em que o mundo parou”. Deste modo, todos os recursos multimodais que são salientados ao longo da chamada posteriormente, corroboram a esta tese principal. Há, portanto, uma hierarquia dos elementos que se destacam em importância visualmente (Kress; Van Leeuwen, 1996, 2006). Nota-se ainda uma construção semântica conflitante entre o vocábulo “parar” e a imagem apresentada em seguida. Isto é, não existe exatamente uma ausência de ação, como é acionado verbalmente (“o ano em que o mundo parou”), mas a exibição de uma manifestação (aglomeração) em um ano pandêmico no qual a recomendação da Organização Mundial da Saúde (OMS) era a do isolamento social.

No encadeamento de imagens 2, 3 e 4, temos acesso à informação mais precisa de que a chama representa o fogo atado na bandeira dos EUA por manifestantes em um protesto. A fotografia naturalística selecionada para compor o plano de fundo da mensagem grafada verbalmente aparece nas cores preta e branca (monocromática). Compreendemos, assim, que uma foto em preto e branco transmite mais sensibilidade, estimulando emoções mais sutis no observador. Ademais, a fotografia dos manifestantes revela a intencionalidade dos produtores em destacar este acontecimento como um evento histórico relevante daquele ano, assim como geralmente aparecem em livros *Revista Gatilho*, Juiz de Fora, v. 25, p. 290-317, 2023 – ISSN: 1808-9461

clássicos e didáticos de História, visto que é salientada já na introdução da chamada. Apesar dessa escolha por uma imagem monocromática, o fogo e a bandeira estadunidense são mantidos nas cores originais, tendo este último elemento um evidente destaque nesse enquadre visual, com o propósito de acionar instantaneamente a memória do telespectador acerca do que está sendo representado ali.

No ano de 2020 houve uma série de manifestações após o assassinato de George Floyd, um homem afro-americano de 46 anos, em um ato extremamente truculento de um policial em Minneapolis, Minnesota, nos Estados Unidos, que foi gravado e compartilhado nas redes sociais. De acordo com o site de notícias *BBC*<sup>18</sup>, “um vídeo mostrando a prisão de Floyd viralizou após o episódio. Nele aparece um policial branco, Derek Chauvin, com o joelho sobre o pescoço de Floyd, enquanto este está algemado e de bruços no chão”.

Podemos notar que a seleção de modos semióticos nos frames destacados para análise na amostra 01 tem relação com as descrições acerca das *zonas de informação* do texto multimodal (Kress; Van Leeuwen, 1996, 2006). A imagem com o ano de 2020 aparece na margem esquerda da tela, contendo a informação já conhecida pelos leitores (o dado), à medida que do lado direito, mais centralizado, apresenta o elemento novo, isto é, a zona de maior destaque em relação à primeira com um maior teor informativo e descritivo: “o ano em que o mundo parou” (1).

Notamos ainda que o *framing* (o enquadramento ou a estruturação) exhibe uma continuidade das cores preta, branca e vermelha (2, 3, 4) que surgem inicialmente e são substituídas pelos elementos: fotografia da manifestação, fogo e bandeira, respectivamente. Além disso, há a presença de molduras nas margens da tela que, conforme se aproximam, mudam da cor branca acinzentada para a preta.

### **Amostra 02:**

Na *figura 3*, temos uma sequência de elementos visuais de 1 a 5 que representa a infecção do mundo pela COVID-19 (do inglês: *Corona Virus Disease 2019*). Em três segundos de vídeo, na margem esquerda, aparece a imagem de um vírus, proporcional ao tamanho do planeta Terra, que surge à direita. Esses dois elementos em formas arredondadas se juntam rapidamente (2, 3, 4) e trocam de posição; uma metade da

<sup>18</sup> Disponível em: <<https://www.bbc.com/portuguese/internacional-52868252>>. Acesso em: 07 de jan. de 2023.

Terra aparece neste momento na margem direita, e outra metade do vírus no canto esquerdo (5) com ênfase para o título MUNDO X PANDEMIA, que é grafado em caixa alta, nas cores branca (MUNDO) e preta (X COVID19). Nesta mensagem, podemos ressaltar a saliência do símbolo X que representa o item lexical de etimologia latina *versus*, indicando relações de oposição.

No plano da metafunção interacional, busca-se estabelecer o elo imaginário (KRESS; van LEEUWEN, 1996, 2006) entre a coisa representada (o produto) e produtor de sentido (o participante interativo, isto é, o observador). Assim, o leitor é levado a interpretar o imbricamento entre os dois elementos arredondados como uma relação de enfiamento, assim como em uma disputa, um embate, após o estabelecimento da representação narrativa intencionada por meio da justaposição entre o mundo e a COVID-19.

**Figura 3:** Elementos imagéticos e a metafunção composicional.

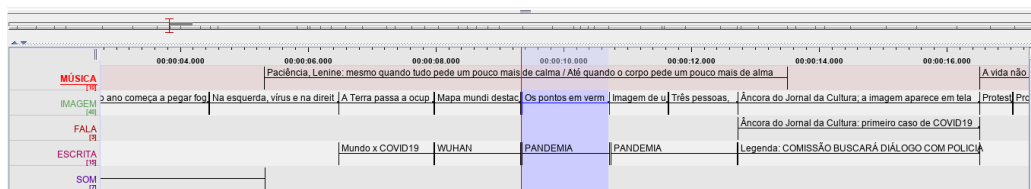


Abaixo, na *figura 4*, é possível observar minuciosamente a articulação entre os modos semióticos dos segundos 06 ao 16. No momento em que a música interpretada por Lenine começa a ser cantada (voz e instrumental), o encadeamento de imagens que vimos acima (*figura 3*) é articulado coordenadamente. Nesta ocasião do vídeo, é possível construir o sentido da correlação entre a música, a imagem e a escrita de que a partir daquele instante, em que o mundo todo foi infectado pelo Coronavírus, seria o período “[...] quando tudo pede um pouco mais de calma / Até quando o corpo pede um pouco mais de alma”.

É importante salientar que esta produção de sentido só incide caso realizemos a leitura simultânea dos modos semióticos mobilizados, visto que cada recurso

desempenha funções específicas, mas que juntos possibilitam a interpretação intencionada pelo *designer* (*The New London Group*, 1996) da produção multimodal em movimento em questão.

**Figura 4:** Correlação entre os diferentes modos semióticos



### Amostra 03:

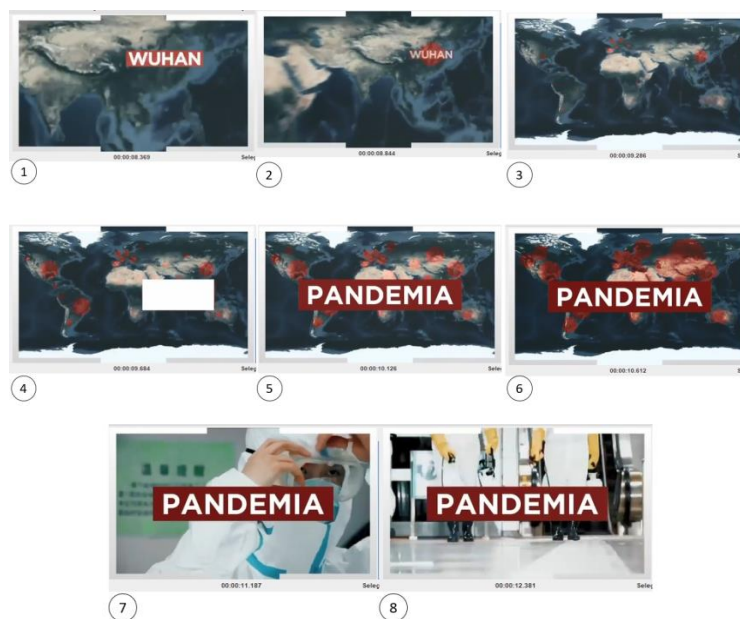
Dando continuidade a sequência de imagens mobilizada para abordar o Coronavírus, temos na *figura 5* elementos verbo-visuais que se articulam na produção de sentido de modo quase que independente.

Em (1), observamos uma aproximação da província da China Central, WUHAN, que ganha um distanciamento em (2) e surge acompanhada de um sinal vermelho ao fundo do letreiro, além do acesso ao mapa *mundi*, que aparece amplamente a partir deste *frame*. Logo após, o ponto vermelho que destaca Wuhan passa a surgir também em outros países ao redor do globo terrestre (3) e vai se alastrando cada vez mais (4, 5, 6) até que aparece centralizado na tela o item lexical “PANDEMIA” (5, 6).

Percebemos que a mancha vermelha, que aumenta incessantemente, representa o vírus se espalhando ao redor do mundo de uma forma rápida; no vídeo isto acontece em um intervalo temporal de menos de três segundos (de 8.369s a 10.612s), representando o modo como a COVID-19 se espalhou por todo o mundo em pouco menos de três meses, gerando assim uma pandemia global. Em (7), temos acesso no plano de fundo da palavra “pandemia”, que continua fixa e centralizada na tela (*saliência*), a um homem de nacionalidade chinesa completamente vestido com uma roupa branca e de máscara, ajustando os óculos industriais em seu rosto. Já em (8), aparecem dois profissionais da limpeza desinfetando um espaço institucional.

Em suma, este conjunto imagético representa narrativamente (Kress; Van Leeuwen, 1996, 2006) o epicentro da pandemia nos primeiros *frames* (1 e 2), sua propagação ao redor do mundo (3, 4, 5) e algumas de suas consequências para a humanidade (6 e 7).

**Figura 5:** Elementos verbo-visuais na produção de sentido.

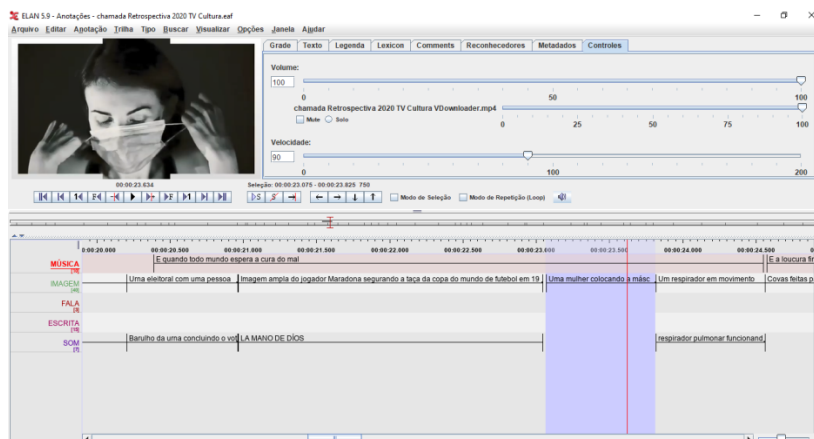


**Amostra 04:**

Na *figura 6*, ressaltamos a articulação entre dois modos semióticos: a música e a imagem. No momento em que o trecho da canção de Lenine diz: “E quando todo mundo espera a cura do mal”, a imagem móvel monocromática em preto e branco de uma mulher jovem colocando uma máscara hospitalar é exibida.

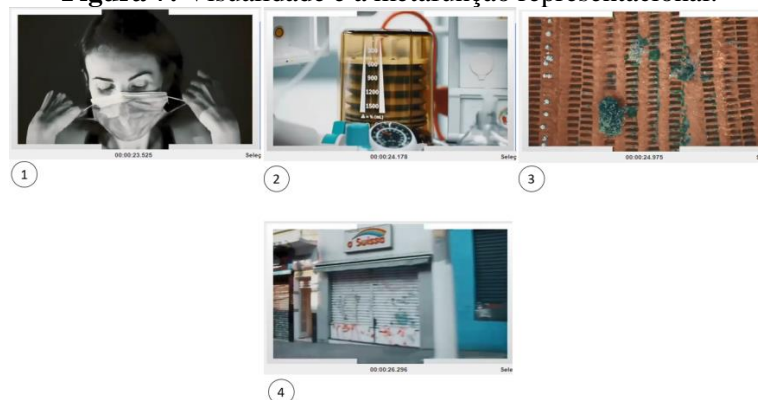
É possível observar que, neste trecho da chamada, acontece uma busca por identificação entre o participante representado (que constitui o assunto/problema da comunicação), isto é, a mulher no vídeo, e o participante interativo (o que visualiza a imagem), o telespectador. A música corrobora a esta produção de sentido, uma vez que, até aquele momento, no Brasil não havia iniciada a vacinação contra o Coronavírus. Há, portanto, uma intencionalidade por parte do produtor em criar um elo de empatia emotivo entre o espectador a partir da mobilização dos modos apresentados.

**Figura 6:** Articulação entre os modos semióticos verbo-visuais.



Nas imagens naturalísticas abaixo (*figura 7*), há um conjunto de fotografias que ilustra trecho da música: “É quando todo mundo espera a cura do mal”. Ou seja, a jovem mulher colocando a máscara hospitalar, como dissertamos anteriormente, a pessoa que está internada no hospital utilizando um respirador pulmonar (2), como podemos inferir, as vítimas mortas pela COVID-19 e, por extensão de leitura, as pessoas que perderam algum conhecido ou ente querido (amigo, vizinho, membro da família) por conta do vírus (3) e, por fim, o centro de uma cidade completamente vazio, representando o comércio que precisou ser fechado para evitar aglomeração e conter o vírus nos primeiros meses de isolamento social (4).

**Figura 7:** Visualidade e a metafunção representacional.



É importante pontuar que as imagens da *figura 6* aparecem com focos de câmeras distintos, produzindo, conseqüentemente, efeitos de sentido variados.

Em (1), temos uma imagem com o foco de câmera aproximado, criando uma afinidade com o observador. Isto é, há uma reconhecimento com a ação realizada na imagem móvel exibida, haja vista que o telespectador passou por aquele mesmo período, tendo

que realizar os mesmos protocolos de autocuidado para evitar o contágio e a transmissão do Coronavírus.

Já em (2), há uma imagem aproximada de um respirador pulmonar que apresenta um conceito sobre quem ou o que é representado, ou seja, os pacientes internados em casos severos que precisavam do aparelho para respirar, criando no observador uma empatia com aquele acontecimento. Na seleção desse elemento imagético, é possível relacionar o respirador pulmonar com o que intitulamos linguisticamente de metonímia, uma figura de linguagem caracterizada pela substituição de um termo por outro, havendo entre eles algum tipo de ligação. Noutros termos, o aparelho médico é substituído pela imagem dos pacientes internados.

Em (3), há uma imagem de um cemitério que foi construído para enterrar o maior número de pessoas possíveis que faleceram por COVID-19, para isso foram construídas covas pequenas e ligadas uma ao lado da outra. É dado a este elemento visual um foco de câmera extenso, no qual o observador tem total domínio (absoluto e pleno) de tudo o que está sendo mostrado (participante onisciente). O efeito de sentido que podemos compreender a partir da observação desta imagem é a de repulsa, pois há uma impessoalidade (lugar vazio, frio, melancólico) e falta de humanização, que nos remete diretamente ao sentimento de perda no ambiente holisticamente destacado.

Por fim, em (4), observa-se imagem em movimento do centro de uma cidade vazia filmado, provavelmente, de dentro de um automóvel. Neste *frame*, vemos um centro comercial atípico em um dia útil com uma estruturação que coloca o participante interativo como se estivesse inserido naquele ambiente, que é produzido com a intenção de criar no espectador uma sensação de estranheza e desconforto.

Como podemos perceber, essas imagens geram diferentes sentimentos no observador, ora de reconhecimento (1) e empatia (2), noutra de pesar (3) e incômodo (4). O enquadre composicional do foco de câmera é essencial para que estes sentimentos sejam experienciados de modo catártico em quem testemunha de tão de perto as consequências da pandemia. Esta sucessão imagética aciona contiguamente a memória afetiva recente do espectador com o intuito de sensibilizá-lo acerca daquelas ações apresentadas narrativamente.

## **5. CONSIDERAÇÕES FINAIS**



Conforme observamos minuciosamente ao logo do empreendimento metodológico e analítico apresentado nas 04 (quatro) amostras, a chamada televisiva de cunho jornalístico analisada neste estudo é um texto audiovisual bastante produtivo que possibilitou dar visibilidade a maneira como os diferentes modos de linguagem são conjugados de forma a construir o significado em textos multimodais.

Para evidenciar as estratégias de construção de sentido mobilizadas coordenadamente na chamada televisiva jornalística das edições especiais do Jornal da Cultura para a retrospectiva 2020, selecionamos, fundamentalmente, categorias analíticas da obra *AGramática do Design Visual* (Kress; Van Leeuwen, 1996, 2006), isto é, as metafunções da linguagem representacional, interacional e composicional, em seus componentes formais e funcionais.

Com o empreendimento analítico à luz das categorias de análise visual, foi possível observar minimamente a relação direta entre os estudos da multimodalidade e os de caráter linguísticos, corroborando assim à teoria sociosemiótica da verbo-visualidade de Kress e van Leeuwen (1996, 2006), compreendendo, então, como se organizam as modalidades selecionadas que se configuram no objeto de estudo. Deste modo, relacionamos os modos semióticos com o valor social que trazem, observamos em que medida os diferentes signos se articulam, explicitamos os possíveis efeitos de sentido da integração de elementos de ordem verbo-visual, nos atentando para estes recursos, para o contexto sócio-histórico e de que modo todos esses recursos ajudam na obtenção do propósito comunicativo do gênero.

Com todos os aspectos analisados, portanto, podemos afirmar que a chamada televisiva é um texto multimodal que agrega várias linguagens que juntas são processadas semioticamente pelo telespectador/leitor.

Por fim, é válido ressaltar que, por conta da limitação de páginas e do caráter sintético do gênero artigo acadêmico, não foi possível desenvolver a apreciação de todos os trechos da produção multimodal em movimento desejados, o que poderia oferecer, evidentemente, uma ampliação da análise do *corpus*.

## Television spot analysis: multimodality and metafunctions

### ABSTRACT:

This paper aims to analyze the categories of the three metafunctions present in the work "Reading Images: The Grammar of Visual Design", by Kress and van Leeuwen (1996, 2006), the representational, interactive and compositional ones, from the perspective of multimodality (KRESS, 2010; VAN LEEUWEN, 2011; LIMA-LOPES, 2012; RAMOS; CARMELINO, 2014; SILVA, 2016; KNOLL; FUZER, 2019), in a television spot. Thus, the study has the purpose of relating the sign modes with the social value they bring, in addition to observing the extent to which the different signs are articulated, explaining the possible effects of meaning from the integration of verbal-visual elements, paying attention to these elements, to the socio-historical context and to the communicative purpose of the genre under discussion. To this end, it is proposed to analyze the TV spot for the special editions of the 2020 retrospective of *Jornal da Cultura* from the notes of the correlation of different semiotic systems of multimodal productions in motion using the software ELAN (SLOETJES; WITTENBURG, 2008). The results showed that the television spot is a multimodal text that aggregates several languages that together are semiotically processed by the viewer to produce meaning. This research concludes, therefore, that the TVspot analyzed is a very productive audiovisual text that made it possible to provide visibility to the way in which the different modes of language are combined in order to build meaning in multimodal texts, allowing to thoroughly observe the direct relationship between studies of multimodality and those of a linguistic nature, corroborating the Social Semiotic theory of verb-visibility.

**KEYWORDS:**The grammar of visual design. Metafunctions. Multimodality. Social Semiotics. Television spot.

**REFERÊNCIAS:**

ALMEIDA, D. B. L. Do Texto às Imagens: As Novas Fronteiras do Letramento Visual. In: Pereira, Regina Celi. Rocca, Pilar (Orgs.). (Org.). **Linguística Aplicada: um caminho com diferentes acessos**. 1ed. São Paulo: Contexto p. 173-202, 2009.

ARAÚJO, A. D. **Gêneros Multimodais: Mapeando Pesquisas no Brasil**. Revista do Programa de Pós-Graduação em Linguística Aplicada da UECE. Volume 3. No 5., p. 13-24, 2011.

BRITO, R. C. L.; PIMENTA; S. M. O. A Gramática de Design Visual. In: LIMA; C. H. P. L.; PIMENTA, S. M. O.; AZEVEDO; A. M. T. (Orgs.). **Incursões semióticas: teoria e prática de gramática Sistêmico Funcional, Multimodalidade, Semiótica Social e Análise Crítica do Discurso: Livre Expressão Editora: Rio de Janeiro**, p. 87-116, 2009.

EGGINS, S. **An introduction to systemic functional grammar**. London: Printer Publishers, 1994.

HALLIDAY, M. A. K.; MATHIESSEN, C. M. I. M. **An introduction to Functional Grammar**. 3rd. London: Edward Arnold, [1985, 1994], 2004.

KNOLL, G.; FUZER, C. Análise de infográficos da esfera publicitária: multimodalidade e metafunção composicional. **Alfa**, São Paulo, v. 63, n.3, p.583-608, 2019.

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**. London and New York: Routledge, 1996, 2006.

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.

LIMA-LOPES, R. E. **Sociossemiótica da produção audiovisual: uma proposta metodológica para análise multimodal da comunicação em vídeo**. 2012. 266 p. Tese (doutorado) - Universidade Estadual de Campinas, Instituto de Estudos da Linguagem, Campinas, SP. Disponível em: <[http://www.repositorio.unicamp.br/handle/REP\\_OSIP/269330](http://www.repositorio.unicamp.br/handle/REP_OSIP/269330)>. Acesso em: 04 de jan. de 2023.

MARCUSCHI, L. A. **Produção textual, análise de gêneros e compreensão**. São Paulo: Parábola, 2008.

RAMOS, P.; CARMELINO, A. C. **Contribuições para leitura de gêneros multimodais no ensino**. Intersecções. Edição 12, Ano 7. Número 1, p.61-77, maio/2014.

SILVA, M. Z. V. **O letramento multimodal crítico no ensino fundamental:** investigando a relação entre a abordagem do livro didático de língua inglesa e a prática docente. Tese de doutorado. Universidade Estadual do Ceará, 329f, 2016.

SLOETJES, H; WITTENBURG, P. **Annotation by category - ELAN and ISO DCR.** In: Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC), 2008.

THE NEW LONDON GROUP. A Pedagogy of Multiliteracies: Designing Social Futures. In: COPE, B.; KALANTZIS, M. (eds). **Multiliteracies** – Literacy learning and the design of social futures. London: Routledge, 1996.

VAN LEEUWEN, T. Multimodality. In: SIMPSON, J. (Editor). **The Routledge Handbook of Applied Linguistics.** London and New York: Routledge, p. 668-682, 2011.