

ANÁLISE MULTIMODAL: NOÇÕES E PROCEDIMENTOS FUNDAMENTAIS

MULTIMODAL ANALYSIS: FUNDAMENTAL NOTIONS AND PROCEDURES

Theodoro Casalotti Farhat*
Paulo Roberto Gonçalves-Segundo**

RESUMO

Partindo da Teoria Sistêmico-Funcional e da proposta decomposicional de Bateman, Wildfeuer e Hiippala (2017), este artigo delinea noções e procedimentos que consideramos fundamentais para a análise de textos multimodais, particularmente os digitais, fornecendo uma proposta teórico-metodológica que permite ao analista um maior comprometimento com uma organização clara e replicável do procedimento analítico. Assim, em primeiro lugar introduzimos os conceitos fundamentais de realização e instanciação e suas implicações para o estudo da multimodalidade; depois, parte-se para a noção de “tela” e para a classificação multidimensional das materialidades semióticas; em seguida, explicitam-se os passos básicos para a fundamentação da análise de textos multimodais; e, finalmente, o modelo é exemplificado com a análise sumária da composição em telas de postagens de Facebook.

Palavras-chave: multimodalidade; análise multimodal; realização; instanciação; Facebook.

ABSTRACT

Drawing on Systemic Functional Theory and the decompositional model proposed by Bateman, Wildfeuer and Hiippala (2017), this paper outlines notions and procedures we consider fundamental for the analysis of multimodal texts, particularly digital ones, by providing a theoretical and methodological model that enables a greater commitment to a clear and replicable organization of the analytical procedure. We introduce the fundamental systemic functional concepts of realization and instantiation and their implications for the study of multimodality; next, we present the notion of ‘canvas’ and the multidimensional classification of semiotic materialities; then, we make explicit the basic steps for the substantiation of the analysis of multimodal texts; and, finally, we exemplify the model with a cursory analysis of the canvas composition of Facebook posts.

Keywords: multimodality; multimodal analysis; realization; instantiation; Facebook.

INTRODUÇÃO

O estudo do fenômeno semiótico conhecido como *multimodalidade* ganhou inegável importância, nos últimos trinta anos, no âmbito dos estudos do texto e do discurso, especialmente pela necessidade contemporânea de uma educação que dê conta não somente do letramento verbal, mas também dos multiletramentos (COPE; KALANTZIS, 2009). Como parte de uma área emergente, porém, pesquisas (especialmente aplicadas) sobre multimodalidade e multiletramentos podem carecer de um aparato teórico-metodológico que efetivamente oriente os procedimentos empíricos de análise e permita uma replicabilidade explícita, especialmente quando os objetos em investigação são textos digitais. Este artigo busca apresentar contribuições nessa direção. Para isso, lançamos, a partir de princípios da Teoria Sistêmico-Funcional (p. ex. HALLIDAY; MATTHIESSEN, 2014; MARTIN, 1992), da proposta teórico-metodológica de Bateman, Wildfeuer e Hiippala (2017) – doravante BWH – e de nossas próprias experiências analíticas, alguns procedimentos que consideramos fundamentais para qualquer investigação que tenha como objeto artefatos que resultam da instanciação de mais de um modo semiótico.¹

Não se tem a pretensão, aqui, de desenvolver reflexões minuciosas sobre todos os procedimentos e noções introduzidos – isso exigiria, no mínimo, o espaço de uma obra inteira. O que pretendemos é apresentar elementos teórico-metodológicos que amparem o pesquisador em um engajamento ativo e consciente com o objeto de análise, minimizando o risco das imprecisões e almejando o máximo de clareza.

O artigo é estruturado da seguinte forma: em primeiro lugar, introduzimos as noções de realização e instanciação como fundamentos da Teoria Sistêmico-Funcional, discutindo as implicações de tais conceitos para os estudos de multimodalidade; depois, especificamos a proposta de Bateman, Wildfeuer e Hiippala (2017) para a análise de textos multimodais, expondo dois importantes parâmetros para a análise: a classificação multidimensional

* Universidade de São Paulo (USP), São Paulo, SP, Brasil. farhat@usp.br.

Orcid: <https://orcid.org/0000-0002-9646-6301>

** Universidade de São Paulo (USP), São Paulo, SP, Brasil. paulosegundo@usp.br.

Orcid: <https://orcid.org/0000-0002-5592-8098>

1. Os resultados apresentados neste artigo provêm de pesquisa financiada pela Fundação de Amparo à Pesquisa do Estado de São Paulo (processo FAPESP n° 2021/03332-0).

e o procedimento de decomposição em telas; a partir disso, reelaboramos tais elementos em termos de uma série de indagações que devem ser feitas sobre o objeto de análise e propomos uma tipologia geral de articulações entre telas; então, o procedimento é sintetizado em termos de quatro passos que organizam o início da análise de textos multimodais; em seguida, tais parâmetros são empregados na análise decomposicional básica para textos de grande relevância na contemporaneidade: as postagens de Facebook; finalmente, concluímos o artigo com considerações sobre as limitações de nossa proposta e sugestões de possíveis caminhos na pesquisa sobre multimodalidade.

1. FUNDAMENTOS DA TEORIA SISTÊMICO-FUNCIONAL

1.1 Realização: do contexto à semiótica; da semiótica ao contexto

Embora a multimodalidade tenha recebido atenção de diversas abordagens e teorias, é inegável que a Teoria Sistêmico-Funcional e a Semiótica Social estão entre as abordagens que mais se debruçaram sobre o estudo das produções multimodais. Tal movimento, iniciado na década de 1980, consolidou-se nos anos 1990 com a publicação de dois livros: em 1990, a primeira versão da Gramática do Design Visual de Kress e van Leeuwen (cuja primeira edição internacional é de 1996); e, em 1994, a obra *The Language of Displayed Art*, de Michael O’Toole.

Ambos os trabalhos tinham como o objetivo fornecer categorias semânticas para a análise de imagens. Para isso, seguiram alguns princípios da Teoria Sistêmico-Funcional² (TSF) (HALLIDAY; MATTHIESSEN, 2014). Um dos fundamentos mais evidentes é o princípio de que modos semióticos são *metafuncionalmente organizados*. Isso significa que há algumas funções que são tão incontornáveis na atividade semiótica que acabam sendo incorporadas à própria organização do sistema semiótico. Halliday mostrou que isso era válido para a linguagem verbal, enquanto Kress, van Leeuwen e O’Toole indicaram que o mesmo ocorria com as imagens. Consideram-se três metafunções:

- A metafunção *ideacional* (ou representacional) abrange a ideia de que a linguagem envolve a construção e articulação de experiências.
- A metafunção *interpessoal* (ou interativa) abarca a noção de que utilizar um recurso semiótico é sempre negociar intersubjetivamente significados.
- A metafunção *textual* (ou composicional) engloba a concepção de que a linguagem dispõe de recursos especializados na composição de textos coesos e coerentes.

A TSF, porém, propõe ainda outro princípio de organização do sistema semiótico: a *hierarquia de realização*, que explicita de que maneira linguagem e sociedade se entrelaçam.

Em primeiro lugar, há uma divisão básica entre um plano de *conteúdo* e um plano de *expressão*. A expressão realiza o conteúdo. Na linguagem verbal há, ainda, uma segunda estratificação do plano de conteúdo – em semântica e lexicogramática; e do plano de expressão – em fonologia e fonética (HALLIDAY; MATTHIESSEN, 2014). Em outros modos semióticos, essas segundas divisões são mais passíveis de debate. A totalidade dos planos semióticos (expressão e conteúdo) servem, por sua vez, para a realização do *plano de contexto* (de situação ou de cultura, ver seção 1.2).

O contexto se realiza na linguagem; o plano de conteúdo se realiza na expressão. Em outras palavras, qualquer modo semiótico deve ser entendido em relação a como se organiza para funcionar em sociedade – daí a noção de metafunção –, estando o plano de conteúdo em íntimo contato com o contexto. O esquema geral de realização de modos semióticos e a hierarquia de realização específica para a linguagem verbal estão representados na Figura 1.³

2. Optamos por “Teoria Sistêmico-Funcional” (e não “Linguística Sistêmico-Funcional”) porque, como indicamos a seguir, as dimensões globais sistêmico-funcionais são relevantes para investigar qualquer entidade semiótica, não somente os textos verbais e seus sistemas.

3. O modelo de Martin (1992) propõe uma estratificação do plano de contexto que não adotamos. Entretanto, as considerações apresentadas neste artigo são, de modo geral, compatíveis com o modelo martiniano.

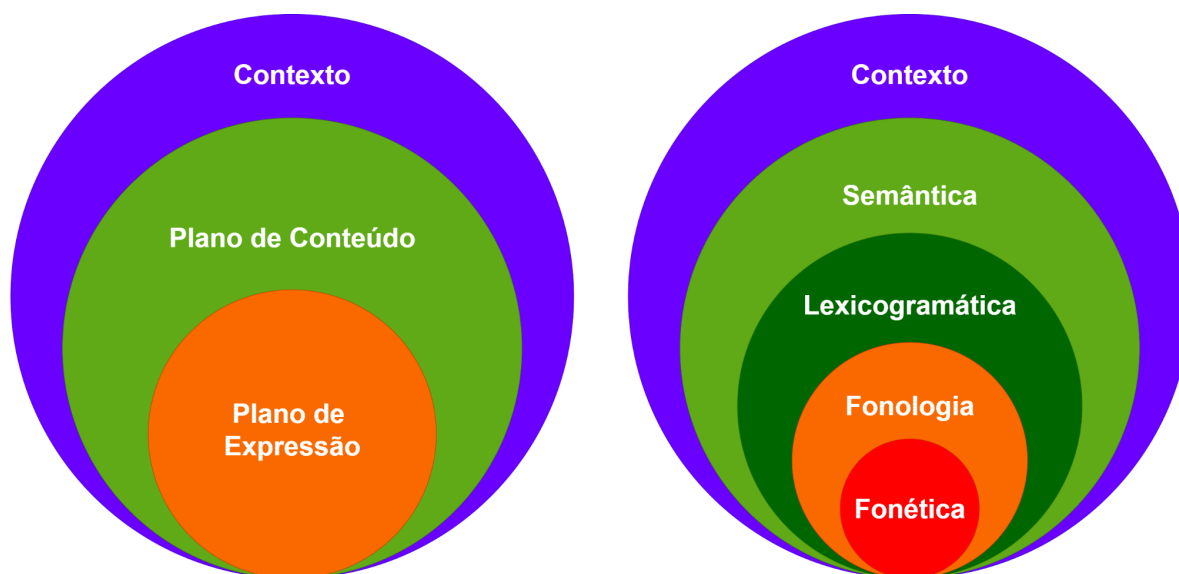


Figura 1. A hierarquia de realização na semiótica geral e na linguagem verbal

Fonte: elaboração própria.

Nos planos de conteúdo e de contexto, há sistemas especializados segundo as metafunções, isto é, há sistemas específicos para significados e parâmetros ideacionais, interpessoais e textuais. Esses sistemas são sistemas de *escolha*: há um conjunto de alternativas de conteúdo e de expressão dentre as quais, dado um certo contexto, o falante (geralmente inconscientemente) faz escolhas. Por exemplo, na linguagem verbal, há a escolha semântica de qual elemento começa a oração: o sujeito, um adjunto, etc.; em uma imagem, de que elementos colocar do lado direito e quais dispor do lado esquerdo.

É por causa da relação de realização que, ao encontrarmos um artefato semiótico, conseguimos, a partir de sua expressão sensível, depreender certos significados; e, a partir de tais significados, depreender aspectos do contexto no qual tal artefato foi construído.

Assim, a relação sociosemiótica de realização nos permite compreender a dialética do “entrelaçamento” entre texto e situação social, entre sistema semiótico e cultura.

1.2 Instanciação: do potencial à instância; do sistema ao texto

Conforme o que explicamos até agora, podemos dizer que, estando diante de um texto, podemos, a partir dele, depreender um contexto, já que o texto realiza esse contexto. Porém, um texto – que definimos, aqui, como uma *instância semiótica* – não se relaciona a qualquer contexto; ele realiza um *contexto de situação*, que pode ser entendido, por sua vez, como uma instância cultural. Entretanto, há também um outro tipo de contexto: o contexto de cultura; este, porém, não é realizado por um texto, mas por um sistema semiótico – que é entendido como um potencial de significação (HALLIDAY; MATTHIESSEN, 2014).

A relação entre sistema semiótico (potencial de significação) e texto (instância semiótica) é a de *instanciação*. Um texto é resultado da instanciação de um potencial de significação; um contexto de situação é uma instância do contexto de cultura. No sistema (potencial), as escolhas estão “por fazer”; no texto (instância), elas estão feitas e concretizadas. Há ainda uma possibilidade intermediária, entre o potencial e a instância: os subpotenciais, em que as escolhas ainda estão “por fazer”, mas algumas são muito mais prováveis do que outras, devido à realização de um certo *tipo* de contexto. Dois grandes tipos de subpotencial recebem especial atenção: registros e gêneros. Não nos deteremos aqui sobre cada um de tais conceitos, mas basta explicar que, quando estudamos vários textos que realizam contextos de situação muito semelhantes, tentando chegar a generalizações acerca de tais textos, partimos do polo da instância para a região dos subpotenciais. Quando estudamos muitos subpotenciais e fazemos generalizações sobre eles, chegamos perto do polo do potencial.

A relação de instanciação pode ser combinada com a de realização, resultando na representação da Figura 2.

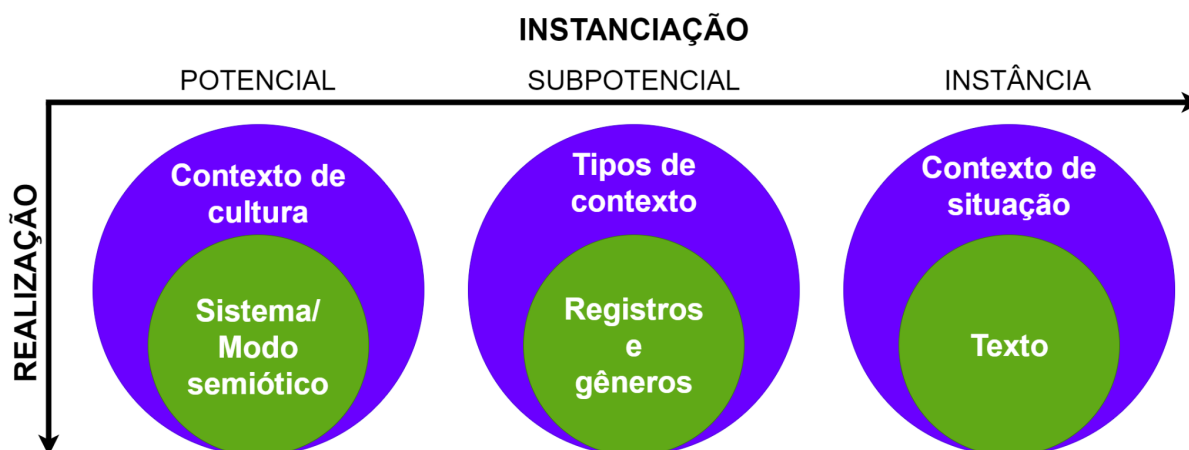


Figura 2. Instanciação e realização combinadas
Fonte: elaboração própria.

Assim, em uma perspectiva sistêmico-funcional, um texto é a instância de um potencial de significação e a realização de um contexto de situação. Em outras palavras, um texto é a ocorrência concreta de escolhas semioticamente relevantes que “ressoam” com elementos situacionais definidos culturalmente, configurando uma unidade de significação integrada internamente (por exemplo, por meio de laços coesivos) e contextualmente (o que envolve, por exemplo, relações de (in)adequação situacional, em que as escolhas semióticas se conformam – ou não – a expectativas culturalmente estabelecidas sobre modos de agir, ser e representar). Por conseguinte, o conceito permite abarcar artefatos semióticos bastante diversos, desde simples bilhetes até aulas inteiras.

Dessa forma, a partir dos conceitos de realização e, principalmente, de instanciação, conseguimos chegar a uma definição de base sistêmico-funcional para o fenômeno da multimodalidade: um texto é multimodal por *resultar da instanciação, em um mesmo contexto de situação, de dois ou mais sistemas semióticos distintos*. Cada sistema (modo) semiótico é uma realização específica do contexto de cultura; porém, em um dado contexto de situação, suas instâncias podem se misturar, se integrar e, portanto, resultar em uma *instância semioticamente heterogênea*. Assim, o texto multimodal é caracterizado pela articulação, em uma só instância, de elementos instanciados por sistemas semióticos diferentes.⁴

A articulação entre dois ou mais modos semióticos não é de pura “soma” comunicativa: a coocorrência de modos semióticos distintos permite a construção de significados que nenhum dos modos semióticos, isoladamente, poderia produzir. Cada um dos modos tem, em si, *affordances*⁵ e potenciais específicos; juntos, há algo novo, embora herdeiro dos modos em questão.

O analista de textos multimodais tem no centro de sua atenção a descrição e a explicação de como ocorrem essas articulações intersemióticas, procurando padrões – de modo que, após diversas análises, tenta fazer generalizações acerca das relações multimodais e, assim, busca a descrição de um *potencial de significação multimodal*. Em outras palavras: alguns modos semióticos coocorrem tão frequentemente que acabam desenvolvendo potenciais especializados na articulação multimodal; esse deve ser o caso, por exemplo, de sistemas paralinguísticos, como gestos e expressões faciais, que são modos semióticos não verbais, mas que certamente desenvolveram recursos especializados na sua instanciação conjunta com enunciados verbais. A Figura 3 explicita tal situação.

4. Pode-se questionar adequação da categoria “texto multimodal”, visto que é difícil imaginar instâncias semióticas em que somente um único modo esteja em uso, resultando em textos “puros”, isto é, “monomodais”. Porém, como o foco do procedimento apresentado neste artigo é justamente investigar questões de multimodalidade, empregaremos tal categoria (“texto multimodal”), por mais redundante que possa ser, para enfatizar exatamente a relação entre modalidades na configuração de um texto.

5. A noção de *affordance* tenta capturar a ideia de que a percepção de um dado objeto é uma atividade, não simplesmente um evento passivo – ao ser “visto em termos de *affordances*, um objeto é diretamente percebido em termos das possibilidades de ação que ele abre para um agente em um ambiente” (BWH, 2017, p. 90, tradução nossa).

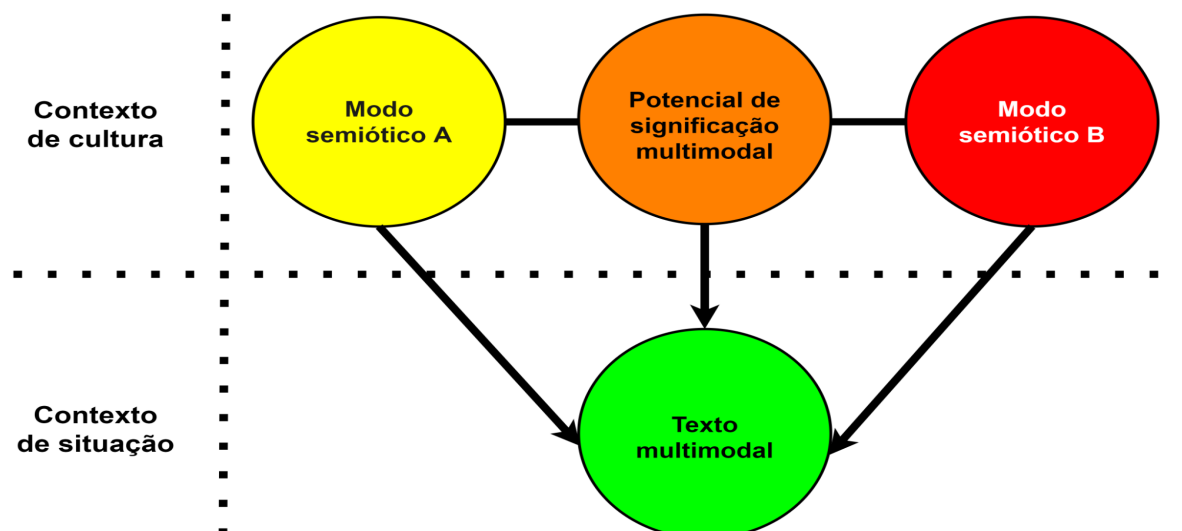


Figura 3. Instanciação multimodal

Fonte: Elaboração própria.

Existem níveis distintos de integração multimodal (MATTHIESSEN, 2009): enquanto a linguagem verbal e sistemas paralinguísticos podem ser altamente integrados, há graus intermediários, como os que ocorrem entre linguagem verbal escrita e imagens estáticas, e mais baixos, como (possivelmente) entre imagens estáticas e sons não verbais. Tais possibilidades devem, porém, ser testadas empiricamente.

Havendo ou não um alto grau de integração de multimodal que nos permita teorizar sobre um potencial de significação multimodal com sistemas especializados em relações intersemiótica, podemos propor – como fazem BWH (2017) – que qualquer modo tem, no seu plano de conteúdo, sistemas especializados em uma semântica discursiva análoga àquela que Martin (1992) propõe para a linguagem verbal. Assim, em qualquer situação em que dois ou mais modos são concomitantemente instanciados, certos elementos de suas respectivas semânticas já estariam preliminarmente dispostos a conexões textuais com elementos de outros sistemas semióticos, permitindo ligações multimodais. Isso explicaria por que é tão fácil que inovações multimodais surjam das mais diversas maneiras, sempre contornando as barreiras da monomodalidade.

Além disso, embora não tenhamos incluído na Figura 3 o nível dos subpotenciais – ou seja, de registros e gêneros –, partes dos potenciais de significação multimodais podem se especializar em certos tipos de contexto, de modo que, na investigação de um texto multimodal, é extremamente recomendável considerar qual registro ou gênero é instanciado (BATEMAN, 2008). Por exemplo, música e imagens dinâmicas são instanciadas tanto em filmes quanto em videoclipes, porém, é muito provável que, em cada contexto, os dois modos semióticos se articulem de maneira bastante distinta.

2. PARA UM PROCEDIMENTO DE ANÁLISE DE TEXTOS MULTIMODAIS

Nesta seção, apresentaremos noções e procedimentos qualitativos que, baseados em nossa experiência de análise, têm utilidade mais prática e imediata para aqueles que querem se aprofundar em abordagens empíricas. Os procedimentos se fundamentam em modelos discutidos por John A. Bateman, Janina Wildfeuer e Tuomo Hippala (BATEMAN, 2008; BWH, 2017; entre outros), mas resultam de nossas interpretações e apropriações de tais propostas, focando particularmente em textos digitais e sistematizando noções que se encontram em um estado mais difuso em BWH (2017). Tal texto, aliás, versa sobre diversos outros aspectos da investigação multimodal que não abordaremos aqui, especialmente em sua faceta quantitativa, que inclui considerações estatísticas, computacionais, etc. Para uma visão mais completa, sua leitura é, portanto, fortemente recomendada.

2.1 A classificação multidimensional: compreendendo semelhanças e diferenças entre materialidades semióticas

Na discussão acima, demos ênfase à noção de modo/sistema semiótico – isto é, uma rede de opções semânticas e expressivas que são culturalmente construídas e podem ser instanciadas em textos, constituindo um potencial de significação.⁶ Porém, tais possibilidades semânticas e expressivas dependem dos materiais que dão base à existência física do texto. Por exemplo, enquanto a modalidade oral da língua tem como material ondas transmitidas pelo ar, a modalidade escrita é manifestada visualmente, com sinais gráficos que chegam aos nossos olhos por meio da luz. Tais diferenças têm impacto relevante sobre o potencial de significação de cada modo e sobre as possibilidades de articulação entre diferentes modos. Pela sua materialidade visual, a linguagem escrita pode, por exemplo, articular-se mais facilmente com imagens estáticas do que a linguagem oral.

Tal dimensão material dos modos semióticos e dos textos que os instanciam é frequentemente conceptualizada a partir da noção de “mídia” ou “meio”. Porém, como tais termos já têm uma grande história de definições distintas e por vezes contraditórias, chegando em alguns casos a se confundir com as noções de “modo”, “sistema”, “expressão” e mesmo “gênero”, preferimos falar de *materialidade semiótica*.

Bateman, Wildfeuer e Hiippala (2017) fazem ainda outra distinção: por um lado, temos as “mídias” – ou, como preferimos, as materialidades semióticas; por outro, em uma materialidade, podemos encontrar uma série de *telas* (*canvases*) divisíveis em subtelas: a divisão é oportuna porque, quando tratamos pura e simplesmente da materialidade semiótica, abre-se a possibilidade de que os mais diversos aspectos materiais se “intrometam” na análise, o que pode levar a um caos analítico; com o conceito de tela, coloca-se em foco o fato de que “a materialidade deve ser sempre considerada em relação às demandas que lhe são feitas semioticamente, e não simplesmente a partir de suas propriedades físicas” (BWH, 2017, p. 104)⁷. Não se produz um texto meramente a partir de materiais, mas sim a partir de telas que, consigo, trazem toda uma história de usos, convenções, percepções socialmente construídas, etc.

Uma tela seria, portanto, não simplesmente a materialidade semiótica, mas a *semiose materializada*: quando se faz a análise de um texto em suas diversas telas e subtelas, o que está em jogo não é simplesmente a materialidade do texto, mas a complexidade do seu aproveitamento cultural, de modo que parte das distinções concretas possibilitadas pelo material passam a ser também distinções semióticas (semânticas, interacionais, etc.).

Portanto, as demandas feitas sobre as telas são demandas culturais. Para que haja, de fato, comunicação e significação, não basta um indivíduo que “lê” o mundo; são necessárias regularidades materiais ancoradas em uma cultura – sem uma cultura, as regularidades não se tornam possíveis, já que, para algo seja percebido como regular, deve haver algum acordo *intersubjetivo*, nunca puramente *subjetivo*. Com isso, preserva-se a premissa de que tanto o potencial quanto a instância são realizações de um contexto social (contexto de cultura no potencial, contexto de situação na instância).

Assim, quando o analista tem sob seu olhar uma instância semiótica, o primeiro passo é identificar as propriedades fundamentais – ou seja, aquelas que são mais relevantes em termos sociosemióticos – da materialidade das telas em questão. Para isso, BWH (2017) fornecem uma classificação multidimensional. As dimensões são agrupadas sob duas facetas; a primeira, mais “básica”, compreende as seguintes distinções:

- *Dinâmico versus estático*: se “o que é representado muda, com o tempo, independentemente de qualquer ação do leitor” (p. 104), trata-se de algo dinâmico; caso contrário, tem-se uma tela estática.
- *2D versus 3D*: tratando-se de algo plano, chato, temos duas dimensões; se a tela permitir extensões em uma terceira dimensão, há três.
- *Graus de transitoriedade* (ou transiência): neste caso, não há uma dicotomia, mas uma escala. Por exemplo, enquanto a língua oral desaparece instantaneamente (se não for gravada), marcas em uma lousa de aula geralmente permanecem por minutos ou horas; já uma tábua pétrea pode ter uma longevidade de milênios. Tais propriedades “influenciam os

6. Tal definição ecoa a de Kress (2010, p. 79; tradução nossa) – “Um modo é um recurso semiótico moldado socialmente e dado culturalmente para a construção de significado” –, mas explicita que há duas escalas de abstração em jogo: diferentes planos (realização) e diferentes graus de generalidade (instanciação).

7. Todas as traduções são nossas.

tipos de atividades de significação possíveis, porque os pesos colocados sobre a memória e a atenção são bastante diferentes” (p. 104).⁸

- *Participante versus observador*: se o consumidor sógnico (aquele que de alguma forma “lê” o texto) for completamente externo às representações, trata-se de um observador; se, entretanto, ele de algum modo também for representado no texto, tendo necessariamente uma perspectiva de primeira pessoa, terá um status de participante. Muitos *videogames*, por exemplo, são justamente caracterizados pelo papel de “participante” que o jogador assume.⁹

O segundo conjunto de dimensões proposto pelos autores é mais complexo. Trata-se, aqui, do “trabalho que os próprios ‘leitores’ têm de realizar para a criação do ‘texto’ com o qual estão engajados em um artefato ou uma performance” (BWH, 2017, p. 105). Para isso, recorre-se à noção de cibertexto: aquele “em que algum tipo de ciclo de ‘feedback’ entre o texto e o consumidor do texto é assumido” (BWH, 2017, p. 105). Mais precisamente, destacam-se casos “em que o tipo de atividade ou engajamento necessário por parte do leitor, visualizador, jogador, etc. ‘estende’ o meio e, assim, afeta os tipos de situações comunicativas possíveis” (BWH, 2017, p. 105).

As dimensões resultantes de tais noções estão ligadas ao grau de *ergodicidade* de um texto. Propõem-se quatro grandes níveis:

- *Linear* (ergodicidade baixa): espera-se que o texto seja lido em uma única ordem, sendo imutável e exigindo do leitor esforço relativamente baixo. Trata-se de um caso extremo e, assim, bastante raro. Exemplo: um texto escrito, preferencialmente curto, compondo solitariamente toda uma página – ressaltamos, porém, que, mesmo nesse caso, o leitor não é “obrigado” a seguir uma ordem linear, ainda que isso seja o esperado ou o preferível. Algo análogo pode ser dito, por exemplo, em relação a textos “puramente” musicais.
- *Micro-ergódico* (ergodicidade média-baixa): casos “em que a linearidade é relaxada e o texto ou a representação no mínimo fazem uso da dupla dimensionalidade da materialidade de seu suporte” (BWH, 2017, p. 107). Dessa maneira, “a ordem de consumo de qualquer representação não é mais fixada à própria apresentação material” (BWH, 2017, p. 107), o que proporciona novas possibilidades de significação. Toda representação pictórica está, no mínimo, nesse nível. Exemplos: uma pintura em um museu, um programa de televisão ao vivo.
- *Ergódico imutável* (ergodicidade média-alta): aqui, o trabalho do leitor é mais claro, sendo caracterizado por uma espécie de exploração de rede. Entretanto, esse trabalho não chega a alterar o texto de fato. Exemplos: revistas e jornais, cujas seções raramente são lidas em ordem, mas “exploradas” de acordo com os interesses do leitor – que, porém, tipicamente não “se intromete” em seu conteúdo.
- *Ergódico mutável* (ergodicidade alta): neste último nível, “o próprio leitor [...] pode alterar a organização ou o conteúdo do texto” (BWH, 2017, p. 108). Exemplos: conversação face-a-face, comentários de Facebook, YouTube, etc. Não se deve confundir com a noção de participante (ver acima), que envolve a representação de um interactante “dentro” do texto, como em muitos jogos digitais.

Conforme explicitaremos e exemplificaremos na sequência, para que a análise possa ser iniciada, devemos ter em mente tanto as dimensões anteriormente elencadas quanto conhecimentos específicos sobre as demandas que são feitas da materialidade semiótica em questão. Cada instância deve ser observada em suas especificidades semióticas; não basta uma rotulação mecânica das categorias que aqui fornecemos, embora elas sejam muito importantes. Um passo nessa direção está no procedimento de “decomposição” introduzido a seguir.

2.2 O procedimento de decomposição em telas

Bateman, Wildfeuer e Hiippala (2017) propõem um “navegador multimodal” (*multimodal navigator*) para “fazer uma ligação entre as categorias teóricas apresentadas [...] e decisões analíticas específicas sobre como abordar e,

8. Os exemplos em questão evidenciam a complexa relação entre a “materialidade semiótica” e as atividades de significação realizadas em tal materialidade por meio das telas. Sem dúvida tais articulações merecem refinamentos teóricos mais aprofundados, possivelmente entrando em contato com a noção de “suporte”.

9. A questão da “participação” do leitor no texto pode ser complexificada por considerações identitárias, visto que atributos sociais do leitor podem ser também representados textualmente, o que tem consequências relevantes (por exemplo, a experiência de uma mulher lendo textos sobre mulheres será distinta da de um homem, já que este não “participa” da representação). Esse tipo de participação pode ser compreendido como um grau intermediário, entre a participação “direta” (como a de avatares em jogos) e a não participação. Agradecemos a um dos pareceristas por chamar nossa atenção a tal aspecto.

então, lidar com qualquer fenômeno multimodal específico” (BWH, 2017, p. 211). O navegador toma, assim, as dimensões apresentadas na seção 2.1 como fundamentais e indica um caminho pelo qual se pode chegar a uma base sólida para a compreensão das maneiras pelas quais as telas se organizam na instância em questão, fornecendo um “mapa” básico sobre o qual considerações analíticas mais sofisticadas podem florescer explícita e empiricamente.

O primeiro passo é tomar o texto multimodal em questão, seja qual for sua complexidade, como uma primeira tela fundamental, incluindo todo “um ‘espaço’ de possibilidades para percepção e ação” (p. 214). Assim, haverá uma tela primária inevitavelmente heterogênea – a heterogeneidade semiótica é traço definidor dos textos multimodais.

Em seguida, então, deve-se “decompor os ‘locais’ particulares em que ocorre significado multimodal ‘por dentro’ da situação ou artefato geral que você está observando” (p. 211). Dividindo progressivamente a tela primária, chega-se a subtelas que tenderão a uma maior homogeneidade; porém, como mostraremos adiante, isso não significa desconsiderar a heterogeneidade do texto multimodal, mas explicitá-la e apontar os aspectos mais fundamentais da articulação entre diferentes modos.

Considerando textos multimodais extremamente complexos, como uma aula presencial, recomenda-se decidir “adiantadamente quais aspectos da situação serão o foco principal do estudo” (p. 212) – por exemplo, interação professor-aluno ou professor-lousa. Evidentemente, isso obriga o analista a tomar decisões “difíceis”, porque alguns aspectos do texto (e, conseqüentemente, algumas de suas telas) são ignorados, ao menos temporariamente. O problema é que, em casos de alta complexidade, tentar investigar todos os aspectos da articulação é, a não ser que se tenha uma equipe relativamente grande de pesquisadores, muito arriscado: algumas facetas podem acabar sendo ignoradas, enquanto outras podem ser analisadas com menos rigor do que o necessário. É sempre mais vantajoso ser explícito em relação a quais aspectos do texto estão em foco e quais são, por ora, deixados de lado.

No exemplo fornecido pelos autores, uma aula de ensino básico pode ser dividida nas seguintes grandes subtelas: interação professor-lousa, professor-apresentação de slides, professor-aluno(s), aluno-livro didático, aluno-aluno, etc. Cada tela deve ser propriamente classificada segundo as dimensões sugeridas anteriormente: a materialidade de uma interação face-a-face é muito distinta daquela que se tem no processo de escrita de uma palavra em uma lousa – há, no mínimo, distinções de dimensões (3D/2D), transitoriedade e ergodicidade.

2.3 Explicitando a decomposição: como começar a indagar o texto

De fato, as próprias dimensões apresentadas na seção 2.1 são um fator decisivo para a divisão da tela primária em telas secundárias, terciárias, etc. Assim, para organizar o procedimento de decomposição multimodal, sugerimos que o analista faça, no mínimo, as seguintes perguntas:

- Algum aspecto do artefato em observação é *temporalmente* organizado? Isto é, há uma natureza dinâmica em pelo menos alguma parte do texto que motive, por exemplo, a divisão entre telas estáticas e dinâmicas? Por exemplo, nota-se com clareza a distinção, em uma página de YouTube, entre o “vídeo” ao qual se pode assistir e outros elementos visuais, como as “*thumbnails*” de outros vídeos.
- Há diferenças no uso do *espaço* em termos de bidimensionalidade e tridimensionalidade? Esse fator seria extremamente pertinente, por exemplo, para distinguir os significados proxêmicos em uma sala de aula (por exemplo, provenientes da organização das carteiras).
- Há diferentes graus de *transitoriedade* em partes distintas do artefato? Por exemplo, no complexo caso da sala de aula, as expressões faciais de alunos e professores se transformam rapidamente, enquanto o que é escrito numa lousa tem geralmente uma fixação maior, mas ainda assim menor do que as anotações dos alunos em seus cadernos.
- Identificam-se diferentes níveis de *ergodicidade*? Se sim, os mesmos participantes têm acesso à mesma ergodicidade nas mesmas telas? Veremos na próxima seção que tal fator é decisivo, por exemplo, para a identificação e caracterização de diferentes subtelas de postagens de Facebook (e, muito provavelmente, de diversas plataformas digitais).

Com o refinamento das subtelas, chega-se a detalhes que permitem ao analista recorrer de maneira mais direta à literatura especializada que pode lhe informar sobre o objeto em questão: a multimodalidade, por suas complexidades inescapáveis, obriga o pesquisador a ter seu horizonte teórico-metodológico aberto a articulações com outras disciplinas e teorias. Caso contrário, corre-se o risco de análises demasiadamente fechadas, ignorando resultados anteriores de outras áreas que, porém, são pertinentes para o estudo da multimodalidade como fenômeno em si.

Por exemplo, um estudo sobre a multimodalidade no cinema poderá partir da Teoria Sistêmico-Funcional, mas não pode ignorar toda a tradição dos estudos do cinema; no caso de interações face-a-face ou de histórias em quadrinhos, é imprescindível recorrer às décadas de pesquisas em psicologia sobre as expressões faciais, recursos indispensáveis em tais instâncias.

2.4 Para uma tipologia das articulações entre telas

Após a divisão em telas, explicita-se de que maneira os diferentes modos semióticos estão distribuídos. Como diferentes modos estão associados a diferentes materialidades semióticas, geralmente a relação entre diferentes telas e diferentes modos será bastante “óbvia”; porém, por vezes o mesmo modo semiótico aparecerá em diferentes telas. Por exemplo, a modalidade escrita da língua geralmente aparecerá, numa aula, tanto na lousa quanto no caderno dos alunos e em seus livros didáticos: trata-se de três (possíveis) telas distintas. Essa possibilidade será explorada em termos de *harmonia modal* – ver abaixo.

Tendo a clareza sobre de que maneira o artefato em análise está organizado em telas e subtelas, cada uma sendo descrita segundo suas características materiais semioticamente relevantes e, por conseguinte, em termos de quais modos semióticos são típicos de cada tela, estamos finalmente em uma posição para prever de que maneira tal configuração semiótico-material do texto “induz” certas articulações multimodais. Tais articulações são extremamente variáveis, dada a infinidade de possibilidades de combinações multimodais. Porém, sugerimos que ao menos três grandes possibilidades de articulação, em um nível global, são possíveis:

- articulação por *harmonia modal*: diz respeito aos casos em que o mesmo modo semiótico é mobilizado em duas telas distintas, o que pode potencializar sua articulação;
- articulação por *proximidade estrutural*: quando uma tela é dividida em subtelas, essas subtelas “naturalmente” estarão em íntimo contato, o que incita certas articulações;
- articulação por *dependência ergódica*: em alguns casos, especialmente quando há altos graus de interação em um texto, uma ou mais telas podem resultar, direta ou indiretamente, da ergodicidade de outras telas, o que leva a certas articulações entre a tela “geradora” e a tela “gerada”.

Essas três possibilidades, aqui apresentadas de maneira bastante geral, serão concretamente aplicadas na seção 3 para a descrição básica de postagens de Facebook.

2.5 Observar, selecionar, decompor e investigar: quatro passos para a análise de textos multimodais

Em resumo, propomos quatro passos iniciais para a análise de qualquer texto multimodal:

1. *Observar*. Em primeiro lugar, o pesquisador deve *observar a complexidade* básica do texto em questão. Isso significa entendê-lo, em primeiro lugar, como uma tela primária em que diversas materialidades semióticas estão em funcionamento.
2. *Selecionar*. Caso a complexidade seja muito alta, deve-se *selecionar um foco* em um ou mais aspectos do texto em detrimento de outros – tal decisão deve ser explícita, sendo tomada preferencialmente antes do início da análise.
3. *Decompor*. Deve-se então *decompor as telas* em foco em subtelas secundárias, terciárias, etc. Para isso, aspectos da materialidade semiótica devem ser colocados em foco, especialmente a temporalidade (estático/dinâmico), espacialidade (2D/3D), a transitoriedade e a ergodicidade (graus distintos e diferenças de distribuição de acordo com participantes). Caso necessário, fatores peculiares ao texto em questão também podem ser levados em consideração para a divisão em subtelas.
4. *Investigar*. Chega-se, finalmente, ao âmago da pesquisa multimodal: *investigam-se as articulações* entre telas distintas (e, conseqüentemente, entre modos semióticos diferentes).

Após o passo quatro, a pesquisa toma, de acordo com seu objeto, rumos distintos, o que impede generalizações sobre quais etapas são recomendadas no prosseguimento da análise; porém, com um fundamento como o proposto, esperamos que qualquer empreitada esteja em posição propícia à colheita de resultados que possibilitem, após uma

quantidade suficiente de análises individuais, alcançar a descrição de padrões multimodais que se voltem a responder algumas perguntas importantes, associadas à hierarquia de realização e à escala de instanciação:

- No polo da instância, como diferentes recursos semióticos, ao se integrarem, operam na realização de elementos contextuais, refratando-os ou construindo-os?
- Na região dos subpotenciais, de que modo a integração multimodal se padroniza, caracterizando multissemioticamente registros e gêneros? Existem ressonâncias entre tais padrões semióticos e elementos culturais típicos?
- No polo do potencial, como diferentes sistemas semióticos operam na realização da cultura, isto é, são moldados por ela, mas também a constroem? Isso inclui o modo como a semiose é colocada a serviço de ideologias, naturalizando relações e representações.

Dessas perguntas pode surgir uma miríade de questionamentos menores, mas consideramos altamente recomendável sempre enquadrá-las em termos de instanciação e realização (e, quando possível, metafunção), visto que tais dimensões fornecem um quadro teórico geral que pode ser extremamente útil na organização da pesquisa.

3 UM EXEMPLO: O CASO DAS POSTAGENS DE FACEBOOK

Nesta seção, exemplificamos o procedimento delineado acima com um caso específico de texto multimodal: postagens de Facebook. Esperamos, com isso, explicitar como as categorias e os passos propostos podem fundamentar análises de textos multimodais, permitindo replicabilidade e comparações – algo fundamental para partir do polo da instância em direção ao subpotencial (registros e gêneros) e, enfim, possivelmente chegar ao polo do potencial multimodal.

3.1 Primeiros passos: as possibilidades de subtelas de postagens de Facebook

Conforme explicitamos acima, o primeiro passo de qualquer análise de texto multimodal é observar o artefato em questão como uma grande tela primária. No caso do Facebook, trata-se de uma tela altamente heterogênea: na tela inicial para a versão para telefones celulares, por exemplo, encontramos o nome da rede social, uma “lupa” para pesquisas na rede, um ícone do aplicativo de conversas Messenger, um espaço em que o usuário é convidado a responder a “O que você está pensando?”, entre outros diversos elementos. Parece-nos, porém, que uma subtela extremamente relevante para qualquer estudo sobre textos e interações no Facebook é o que se convencionou chamar de “postagem”. É essa subtela, portanto, que selecionamos para a análise, embora certamente as outras telas e suas subtelas também mereçam atenção analítica.

Temos, então, nossa tela primária: a postagem tomada como um todo. Um exemplo de postagem é mostrado na Figura 4.



Figura 4. Exemplo de postagem de Facebook

Fonte: <<https://www.facebook.com/estadao/posts/5356244907723861>>. Acesso em 13 de abril de 2021.

Um primeiro aspecto, bastante evidente, é o fato de que toda e qualquer postagem será eminentemente visual: com imagens (dinâmicas ou estáticas) ou somente com linguagem verbal escrita, sempre haverá alguma presença de elementos visuais para que haja uma postagem. Por isso, um dos fatores para a distinção entre as telas de uma postagem é a espacialidade da postagem como um todo, considerando elementos de composição visual especialmente em termos de sua posição (acima e abaixo) e de sua molduragem – ou seja, a maneira como os elementos são visualmente representados como conjuntos ou disjuntos (KRESS, VAN LEEUWEN, 2006). Embora consideremos tais critérios bastante claros, ainda carecemos de, por exemplo, pesquisas com rastreamento ocular no Facebook que tenham como foco justamente a distinção visual de telas, fornecendo fundamentos mais empíricos.

Além de questões de espacialidade, plataformas de redes sociais são caracterizadas pela alta interatividade (cf. BWH, 2017, p. 355). Nas dimensões elencadas acima, isso significa que plataformas como o Facebook têm, em grande parte de suas telas, um nível alto de ergodicidade (ou seja, há telas ergódicas mutáveis). Entretanto, como apontamos anteriormente, não devemos considerar somente o nível de ergodicidade para a distinção de telas, mas também como a ergodicidade é distribuída: quais e quantos usuários têm acesso a qual nível de ergodicidade em cada tela. É possível que telas distintas, mesmo com o mesmo grau de ergodicidade, tenham autores distintos. Ademais, mesmo distinções entre níveis baixos de ergodicidade (sem mutabilidade) são relevantes para, por exemplo, distinguir telas especializadas em texto verbal (escrito) daquelas em que se instanciam imagens – embora ambos sejam visuais, cada modo semiótico tende a caminhos de leitura muito distintos.

A partir dos critérios gerais de composição visual e de ergodicidade (níveis e distribuição), já se pode propor uma divisão da tela de postagem em suas primeiras subtelas: *tela-base*, *comentários*, *interface 1* e *interface 2*: ver figura 5.

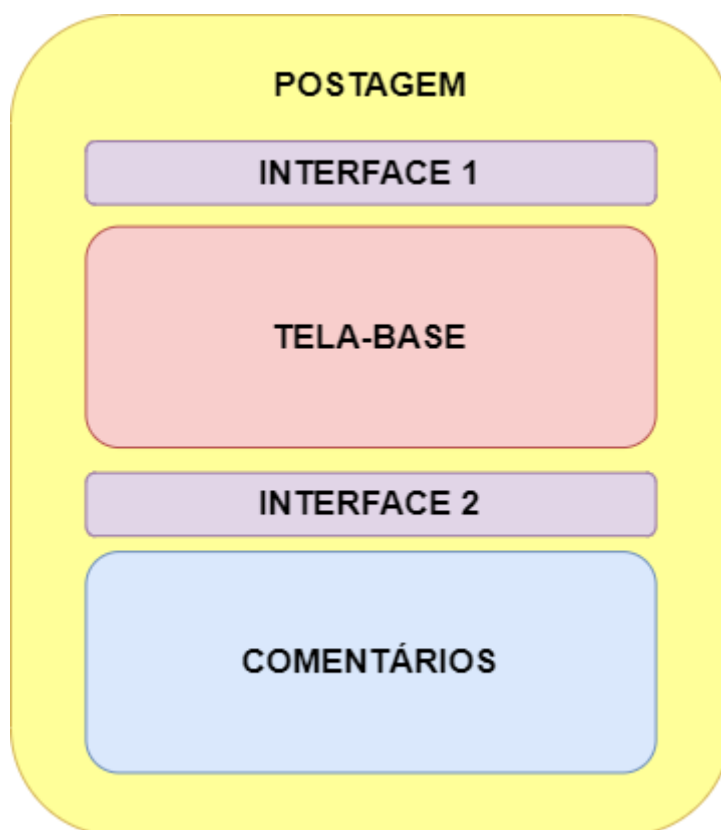


Figura 5. Decomposição de uma postagem em suas subtelas iniciais

Fonte: Elaboração própria.

A tela-base, como sua denominação indica, tem uma importância “basilar”, salientada pela sua posição superior na composição visual, para qualquer postagem: é nela que encontramos os enunciados instanciados pelo “postador”, o “autor original” da postagem. Em termos de ergodicidade, trata-se de uma tela mutável para o postador, mas não mutável para outros usuários – dependendo de suas subtelas, que são variáveis, a tela-base pode ser, para seus leitores, linear, micro-ergódica, ou ergódica imutável.

A principal subtela da tela-base é denominada *matriz*: trata-se do conteúdo interpretado pelos leitores como “original”, sem edições posteriores. Há, no âmbito da matriz, diversas possibilidades de configuração em subtelas. Caso se trate de uma matriz monomodal (por exemplo, somente uma imagem ou somente texto escrito), é possível considerar que a matriz não tem subtelas; porém, quando há mais de um modo, é altamente recomendável dividir a matriz, com ao menos uma subtela para cada modo. Assim, em geral, a matriz é geralmente encontrada em uma das três opções a seguir:

- Matriz *monomodal*: caso relativamente raro em que somente um modo semiótico é instanciado na matriz, o que pode impedir a divisão da matriz em subtelas.
- Matriz *verbo-imagética sem link*: ocorre quando o modo verbal escrito e o modo imagético (estático ou dinâmico) são instanciados na matriz sem que isso ocorra devido à presença de um *link* de internet. Divide-se em matriz imagética e matriz verbal. Se o modo imagético for estático (imagens), há a possibilidade de haver várias imagens (um “carrossel”), o que justifica a divisão da matriz imagética em mais subtelas (uma para cada imagem); caso seja dinâmico, quase sempre só haverá um vídeo ou um GIF, mas é muito provável que o vídeo ou o GIF demande uma divisão em subtelas internas a si. Existe ainda a possibilidade de matriz imagética com imagens estáticas e vídeo(s), o que também exigiria mais subtelas (ao menos uma para cada imagem e cada vídeo).
- Matriz *verbo-imagética com link*: o conteúdo da matriz é resultado, ao menos em parte, do compartilhamento de um *link* que leva a outro espaço digital (por exemplo, na Figura 4, a uma notícia no site do jornal *O Estado de S. Paulo*). Nesse caso, o que geralmente se tem é a divisão da matriz nas seguintes subtelas: matriz autoral (enunciados escritos que inferimos serem “provenientes” do postador) e matriz-*link* (enunciados provenientes da presença do link); esta

última tende a ser verbo-pictórica (caso da Figura 4) e, por isso, pode ser dividida em matriz-*link* verbal e matriz-*link* imagética.

Outra subtela que pode estar presente na tela-base, ainda que seja mais rara, restringindo-se a alguns espaços específicos do Facebook, como certos “grupos” (cf. FARHAT; GONÇALVES-SEGUNDO, 2021a, 2021b), são os “edits”. Como sua denominação indica, são edições feitas à tela-base – porém, mais especificamente, “edits” são edições explicitadas como tais, o que leva à sua distinção da matriz (que, como vimos, é a parte da tela-base interpretada como “original”, em que possíveis edições não são explicitadas). Por resultarem de alterações posteriores à publicação inicial da postagem, os “edits” são estabelecidos, em termos de materialidade semiótica, pela alta ergodicidade (na perspectiva do postador, que pode alterar a tela-base, criando “edits”) e pela temporalidade (na perspectiva do leitor, que, ao encontrar um “edit” infere uma diferença temporal entre o momento de publicação dos “edits” e o da matriz).

Os “edits” são geralmente indicados pela construção “Edit: X”, em que “X” é o enunciado acrescido à tela-base. Na grande maioria dos casos, “edits” são verbais; porém, nada impede que um postador faça uma edição pictórica, acrescentando, por exemplo, uma imagem à tela-base – porém, para ser um “edit”, deve ficar evidente para os leitores que a imagem adicionada não faz parte da matriz. Na Figura 6, por exemplo, veem-se dois “edits”: um verbal, marcado por “Edit”, e um imagético, marcado pela posição à direita, o que indica ser algo “novo”¹⁰ (KRESS; VAN LEEUWEN, 2006), e pela coesão multimodal com o “edit” verbal, especialmente com a construção “Fotinha dela”.

Minha mulher tá em isolamento por covid e minha cachorra não sai de perto da porta e fica chorando pra entrar o dia todo e me indicaram a não deixar ela entrar de jeito nenhum, to com mt dó cara 😞

Edit: pessoalllll, minha mulher refez o teste hoje!!! Ela tá curada, minha cachorrinha vai poder voltar a dormir com a gente. Cara, eu to tão feliz!!!!
Fotinha dela aí toda feliz lambendo tudo 😄



👍 Curtir

💬 Comentar

👍👍👍 2,1 mil

Figura 6. Exemplo de postagem com dois “edits” (verbal e pictórico)

Fonte: <https://www.facebook.com/groups/LDRV12>. Acesso em julho de 2021.

10. Nas palavras de Kress e van Leeuwen (2006, p. 180-181), “Se algo é Dado, tal elemento é apresentado como algo que o leitor já conhece, como um ponto de partida da mensagem que é familiar e estabelecido. Se algo é Novo, tal elemento é apresentado como algo que ainda não é conhecido, ou talvez que ainda não esteja em acordo com o leitor, e, portanto, como algo a que o leitor deve prestar atenção especial”.

Assim, em resumo, temos as duas subtelas da tela-base: matriz e (opcionalmente) “edits”. Caso não haja “edits”, sugerimos que se mantenha a categoria de matriz, já que, embora de certa forma ela fique, nesses casos, em redundância com a tela-base, a matriz indica o “núcleo duro” da postagem e reconhece o potencial ergódico da tela-base.

Passemos agora a um constituinte “oposto” à tela-base: os comentários. Sua característica distintiva é sua mutabilidade “aberta”: ao contrário da tela-base, os comentários são caracterizados por sua abertura a uma grande quantidade de usuários – caso a postagem seja totalmente pública, qualquer usuário do Facebook pode modificar a tela de comentários. Essa mutabilidade se dá basicamente de três formas: o usuário pode acrescentar um novo comentário, responder a um comentário e “reagir” a um comentário.

Cada comentário único na tela geral de comentários pode ser entendido como uma subtela da tela de comentários, que chamamos de “unidade-C”; caso uma unidade-C receba respostas, pode-se fazer uma distinção entre o comentário-matriz e a subtela de respostas. Se houver diferenças de modo semiótico, também é possível dividi-lo em comentário-matriz imagético, comentário-matriz verbal, resposta verbal, resposta imagética, etc. Desse modo, dá-se conta tanto do fato de que diferentes participantes gerarão diferentes telas quanto do princípio de que modos semióticos diferentes operam diversamente.

Finalmente, junto à tela-base e aos comentários, temos duas telas de interface. Embora relacionadas, elas apresentam graus diferentes de ergodicidade e estão separadas visualmente, o que justifica sua desunião. As interfaces da postagem apresentada na Figura 4 são exibidas na Figura 7.



Figura 7. As interfaces de uma postagem

Fonte: Adaptado de <<https://www.facebook.com/estadao/posts/5356244907723861>>. Acesso em 13 de abril de 2021.

As interfaces têm em comum o fato de que, em geral, há baixos níveis de ergodicidade, independentemente do participante: no caso da interface 1, há uma tela micro-ergódica; na interface 2, há possibilidades bastante limitadas de mutabilidade da tela.

A interface 1, no topo da postagem, apresenta alguns elementos. Em primeiro lugar, à esquerda, em formato circular, vemos a “foto de perfil” do postador, que geralmente é, de fato, uma representação da pessoa ou da página que fez a publicação. Em segundo lugar, temos o nome do postador, que, no caso de pessoas físicas, exige um certo “realismo”, apresentando, em geral, seu nome jurídico (uma exigência do Facebook) e, com frequência, seu sobrenome.

Abaixo de tais elementos, vê-se a data de publicação do texto (ou, se a postagem for recente, há quanto tempo foi postada); trata-se de uma “âncora temporal”. Finalmente, ao lado do marcador de tempo, há um pequeno símbolo indicando o nível de publicidade da postagem: em postagens públicas, há um pequeno globo terrestre; em postagens restritas a “amigos”, vemos três figuras humanas, etc.

Ainda pode haver, além da “âncora temporal”, uma “âncora espacial”: é o que ocorre quando encontramos, no “feed de notícias”, postagens provenientes de grupos. Há, no espaço habitualmente ocupado pela “foto de perfil” do postador, dois elementos sobrepostos: um maior, ao fundo, em que se vê a imagem de “capa” do grupo; e outro, menor, em que se vê a foto de perfil do postador. Analogamente, há, no espaço usualmente reservado para o nome do postador, o nome do grupo, enquanto a denominação do postador fica abaixo, menos saliente. Ver Figura 8.

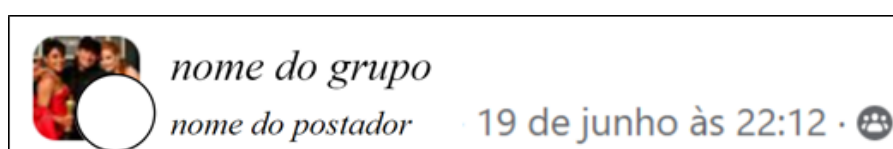


Figura 8. Interface 1 em postagens de grupos

Fonte: Elaboração própria.

Essa configuração, ao dar maior saliência à identificação do grupo e menor destaque à identidade do postador, parece supor que, na interação em grupos, importa menos a identidade específica dos usuários do que a identidade geral de “membro do grupo X”. Essa diferença de enquadramento pode ter efeitos diretos na maneira como um usuário, ao encontrar uma postagem de grupo em seu *feed*, a interpreta, já que há uma mudança no modo como a identidade do postador é apresentada.

A interface 1 tem, portanto, a importante função de orientar o leitor em relação a elementos fundamentais da enunciação (BENVENISTE, 2006): o enunciador (imagem-nome), o tempo (data/horário), o espaço (os modos de indicar o grupo) e os enunciatários (o indicador do nível de privacidade/publicidade da postagem). Assim, essa tela é um “conjunto multimodal” (*multimodal ensemble*; JEWITT; BEZEMER; O’HALLORAN, 2016) que parece servir de compensação para a “descorporificação” do discurso digital, ancorando a interação e oferecendo uma introdução elementar à postagem. Pode-se dizer, ainda, que há, entre a representação do enunciador (imagem-nome) e a tela-base, uma relação de projeção verbo-pictórica: o postador está “dizendo” a postagem.

A interface 2, por sua vez, é encontrada abaixo da tela-base e acima dos comentários; a da postagem mostrada na Figura 4 está em destaque na Figura 9:

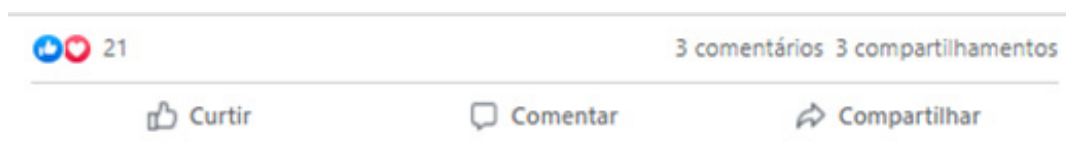


Figura 9. Exemplo de interface 2

Fonte: <<https://www.facebook.com/estado/posts/5356244907723861>>. Acesso em 13 de abril de 2021.

Há, visualmente, uma separação entre a parte superior dessa interface e a inferior, o que (junto a níveis distintos de ergodicidade) autoriza uma divisão em duas subtelas. Veremos que a superior é, porém, ergodicamente dependente da inferior – e é por isso que começamos a discussão por essa última. Encontram-se, na região inferior, três “botões”: à esquerda, há um pictograma com os contornos de uma mão fazendo um sinal de “positivo” e o verbo “curtir”; ao centro, vemos um pictograma de “balão” acompanhado pelo verbo “comentar”; à direita, uma seta curva e o verbo “compartilhar”. A não finitude dos verbos, ao não ancorar deiticamente o processo, sugere uma abertura ergódica que é definidora das funcionalidades da (sub)interface: ao tocar virtualmente o “joinha” de “curtir”, abre-se um subtela em forma de balão com diversas possibilidades de reação: uma mão com o mesmo sinal, sobre a qual,

quando colocamos o dedo/cursor, vemos o verbo “curtir”, um coração, e quatro emojis – com expressões de riso, surpresa, tristeza e raiva.

Assim, a interface fornece uma paleta de reações “primárias”, não verbais e de menor esforço necessário do que, por exemplo, publicar um comentário. As seis opções, particularmente as quatro faces, refletem emoções consideradas “básicas” pela literatura especializada: por exemplo, com base em expressões faciais, Ekman, Friesen e Ellsworth (1982) listam como básicas raiva, nojo, medo, alegria, tristeza e surpresa. Como não há de fato uma mudança da interface, mas sim um “caminho” a ser seguido, podemos classificar a subtela como ergódica imutável, indicando sua interatividade relativa.

A segunda parte da subtela inferior, com o balão de “comentar”, liga-se diretamente à tela de comentários: quando clicamos sobre ela, abre-se uma janela em que são vistos os últimos comentários (ou os “destacados”, de acordo, por exemplo, com o número de reações e respostas recebidas). Dessa maneira, a interface, com sua interatividade limitada, funciona como via de acesso a outra tela – esta, sim, com possibilidades de interação muito maiores.

Enfim, a última seção da subtela inferior da interface 2, com a opção de “compartilhar”, aparece quando o usuário pode “compartilhar” uma postagem em seu perfil – processo enunciativo complexo em termos de autoria, já que há, então, um comprometimento potencialmente ambíguo entre o “compartilhador” e aquilo que é compartilhado. Não nos deteremos nessa questão, mas é necessário indicar que postagens “compartilhantes”, por sua funcionalidade de “reapresentação”, apresentam uma decomposição em telas distinta da decomposição de postagens “não compartilhantes”: em postagens compartilhantes, cria-se uma “macrotela de moldura” que enquadra a postagem compartilhada como tal – ou seja, *reproduzida* pelo usuário, não propriamente *enunciada* por ele.

Passemos à parte superior da interface 2. Comparada com a inferior, ela é relativamente simples: à esquerda, vê-se o resultado de interações com o botão de “curtir” – a interface indica quais e quantas “reações” (da “paleta de emoções”) a postagem recebeu; à direita, há o número de comentários e de compartilhamentos da postagem.

Desse modo, a interface 2, em sua totalidade, tem duas grandes funções: em primeiro lugar, oferecer possibilidades de interação, com as reações, o botão de passagem para a tela de comentários e a possibilidade de compartilhamento; em segundo lugar, apresentar uma “moldura interacional” da postagem, indicando, a partir da subtela superior, de que modo a postagem foi recebida – quantas e quais reações, quantos comentários e quantos compartilhamentos a postagem recebeu: isso pode funcionar como índice da reputação do postador ou do sucesso da postagem. Em última instância, aqui haveria uma estratégia de captação para a leitura da postagem: “se outros se interessaram, eu também vou”.

Assim, as interfaces 1 e 2, embora não componham uma única tela, parecem existir para fornecer certas indicações rápidas ao leitor sobre de que maneira a postagem em questão se ancora enunciativamente. Em outras palavras, com as interfaces, o leitor rapidamente aprende quando, “onde”, por quem e para quem a postagem foi publicada (interface 1), e também recebe índices preliminares sobre como (e quantos) leitores anteriores com ela interagiram (interface 2).

Terminamos, aqui, a exposição primária de nossa proposta de decomposição em telas de postagens (não compartilhantes) de Facebook. Há, certamente, diversos aspectos que foram deixados de lado – nossa exposição não se pretende exaustiva –, mas estudos que se centrem exaustivamente em postagens de Facebook provavelmente terão mais considerações a fazer sobre as telas apresentadas, podendo propor subtelas mais delicadas do que as aqui sumariamente indicadas.¹¹

3.2 As articulações fundamentais entre as subtelas de postagens de Facebook

Agora, em consonância com o procedimento analítico proposto, apresentaremos algumas articulações básicas entre telas de postagens de Facebook. Como indicamos, há diversas configurações internas possíveis para a tela-base: com ou sem “edits”, matriz monomodal ou verbo-imagética, etc. Por isso, focaremos em um caso específico: postagens com “edits”, matriz verbal, matriz pictórica estática e comentários. Essa configuração pode ser vista na

11. Aqui se entrevê o problema do grau de *granularidade* da decomposição: até que ponto deve-se decompor um texto? Trata-se de uma questão cuja resposta depende ao menos de dois fatores: 1. o foco da análise – por exemplo, em uma pesquisa que investiga como usuários “reagem” a postagens, uma decomposição mais detalhada da interface 2 seria imprescindível; 2. a complexidade do objeto: por definição, um texto complexo exigirá mais passos de decomposição para que se alcance um grau satisfatório de explicitação da organização da materialidade semiótica.

Figura 10, que mostra uma postagem de tal tipo e sua decomposição. Por questões de espaço, não nos deteremos em articulações com as interfaces e com as unidades-C da tela de comentários, mas ambas certamente mereceriam atenção minuciosa em um estudo aprofundado sobre postagens de Facebook.

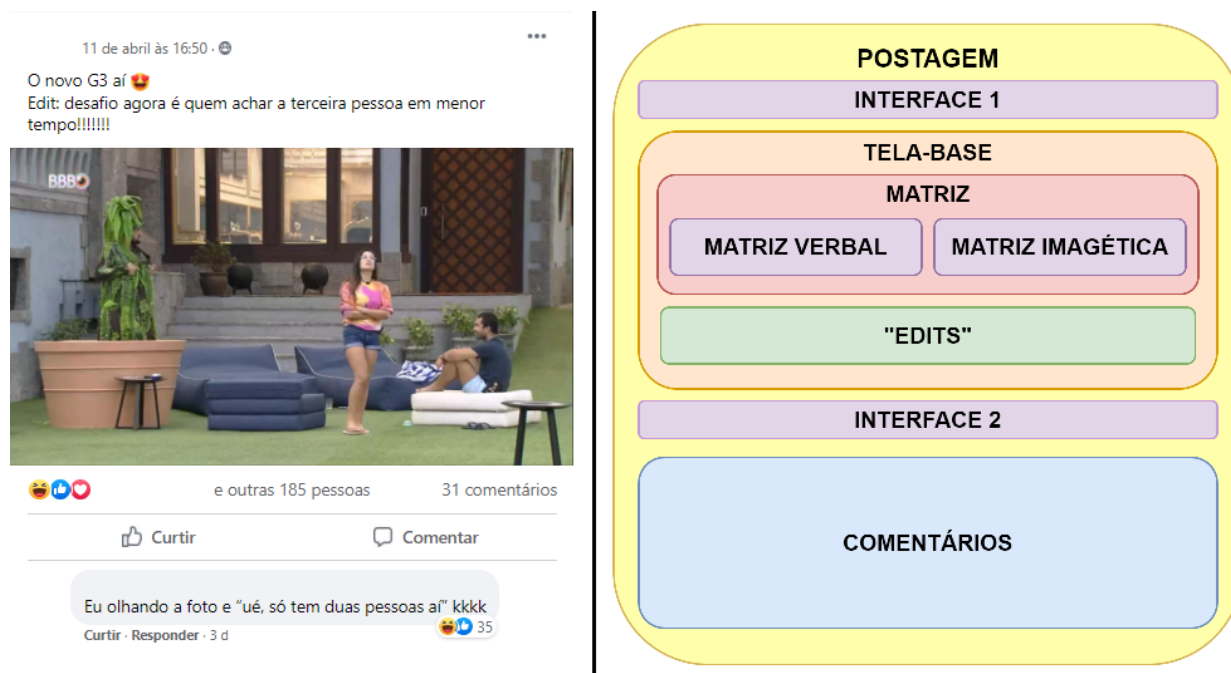


Figura 10. A decomposição básica de uma postagem com matriz verbo-imagética e "edits"

Fonte: postagem adaptada de <www.facebook.com/groups/270581714352692/>. Acesso em 14 de abril de 2021. Diagrama de elaboração própria; usuários foram anonimizados.

Na totalidade da postagem, a relação mais generalizada é a de *proximidade estrutural*: a partir do momento em que duas subtelas compõem, juntas, uma tela, existe proximidade estrutural. De fato, sem proximidade estrutural, seria impossível fazer qualquer uma das considerações acima sobre a divisão da postagem em subtelas. Porém, por mais trivial que seja, a relação de proximidade estrutural ganha importância ao possibilitar a união de subtelas que, superficialmente, podem parecer muito distintas. No nosso caso, a proximidade estrutural é muito importante para, por exemplo, dar fundamento à articulação entre matriz verbal e matriz imagética – é essa proximidade, ao tratar os dois constituintes como partes de um mesmo todo (a matriz), que indicará, por exemplo, que, na postagem na Figura 10, o adjunto "aí" na matriz verbal aponta para a matriz imagética.

Analogamente, toda a tela da matriz (ou seja, a matriz verbal unida à imagética) está em proximidade estrutural com a tela de "edits", já que matriz e "edits" formam, juntos, a tela-base. Tal relação indica que, embora "edits" e matriz sejam nitidamente diferentes pela percepção que se tem dos seus distintos momentos de publicação, percebe-se também certa homogeneidade entre as duas telas porque, basicamente, atribui-se ao postador a responsabilidade por ambas.

Porém, um rápido exame da postagem na Figura apontará para uma relação mais evidente entre matriz e "edits": tanto a matriz verbal quanto os "edits" são instâncias verbais (escritas) – salvo os raros casos de "edit" imagético; por isso, entre matriz verbal e "edits" há *harmonia modal*. Note-se que, por haver tantas articulações básicas (proximidade estrutural e harmonia modal) entre as telas, foi necessária a criação de uma norma (nas comunidades que usam "edits"): quando se instancia um "edit", deve haver uma indicação explícita; caso contrário, a tela de "edit" é interpretada como parte da matriz. Algo análogo ocorre quando há "edits" imagéticos – nesses casos, outras convenções são necessárias para evidenciar que a imagem é um "edit" (por exemplo, indicações composicionais, já que o "edit" imagético costuma aparecer à direita, na posição especializada em apresentar informações "novas" – cf. KRESS; VAN LEEUWEN, 2006).

Passemos agora a duas outras articulações básicas, que associam tela-base e comentários. A relação mais pertinente entre comentários e tela-base se dá especialmente por uma questão de *dependência ergódica*: tipicamente, após ler a matriz, o leitor pode assumir um papel ativo na modificação da tela em que tal atividade lhe é permitida

– os comentários. Mais especificamente, classificamos tal dependência ergódica como *de primeira via*, pois é essa relação que vai promover, na realidade, a própria ocorrência de comentários. Evidentemente, é possível que nem todos os comentários sejam resultado de alguma reação à matriz – e, portanto, não se trata de uma dependência total; porém, tanto a nossa intuição (enquanto usuários de Facebook) como estudos empíricos anteriores (FARHAT; GONÇALVES-SEGUNDO, 2021a, 2021b) indicam uma relação sistemática entre comentários e matriz cujo fundamento é, provavelmente, a dependência ergódica de primeira via.

Entre comentários e “edits”, há outra relação de dependência ergódica. Porém, diferentemente da dependência ergódica de primeira via, há entre “edits” e comentários *dependência ergódica de segunda via*: presume-se que o postador, que pode editar a tela-base, após ler os comentários, julgue conveniente fazer uma edição explicitamente marcada como tal, o que gera um ou mais “edits”. Entretanto, aqui a relação é menos “causal” do que entre matriz e comentários, já que sem dúvida o postador pode fazer um “edit” em reação a um fenômeno, por exemplo, exterior à interação na postagem. Seja como for, nossas análises têm mostrado que a presença de “edits” é geralmente motivada pelos comentários e, assim, esta relação merece destaque. Também cabe ressaltar que, aqui, a dependência ergódica é bidirecional: análises anteriores (FARHAT; GONÇALVES-SEGUNDO, 2021a, 2021b) sugerem que a presença de “edits” pode alterar significativamente a composição dos comentários – que reagem à tela-base como um todo, não somente à matriz.

Em síntese, as articulações fundamentais exploradas nesta seção estão representadas na Figura 11.¹²

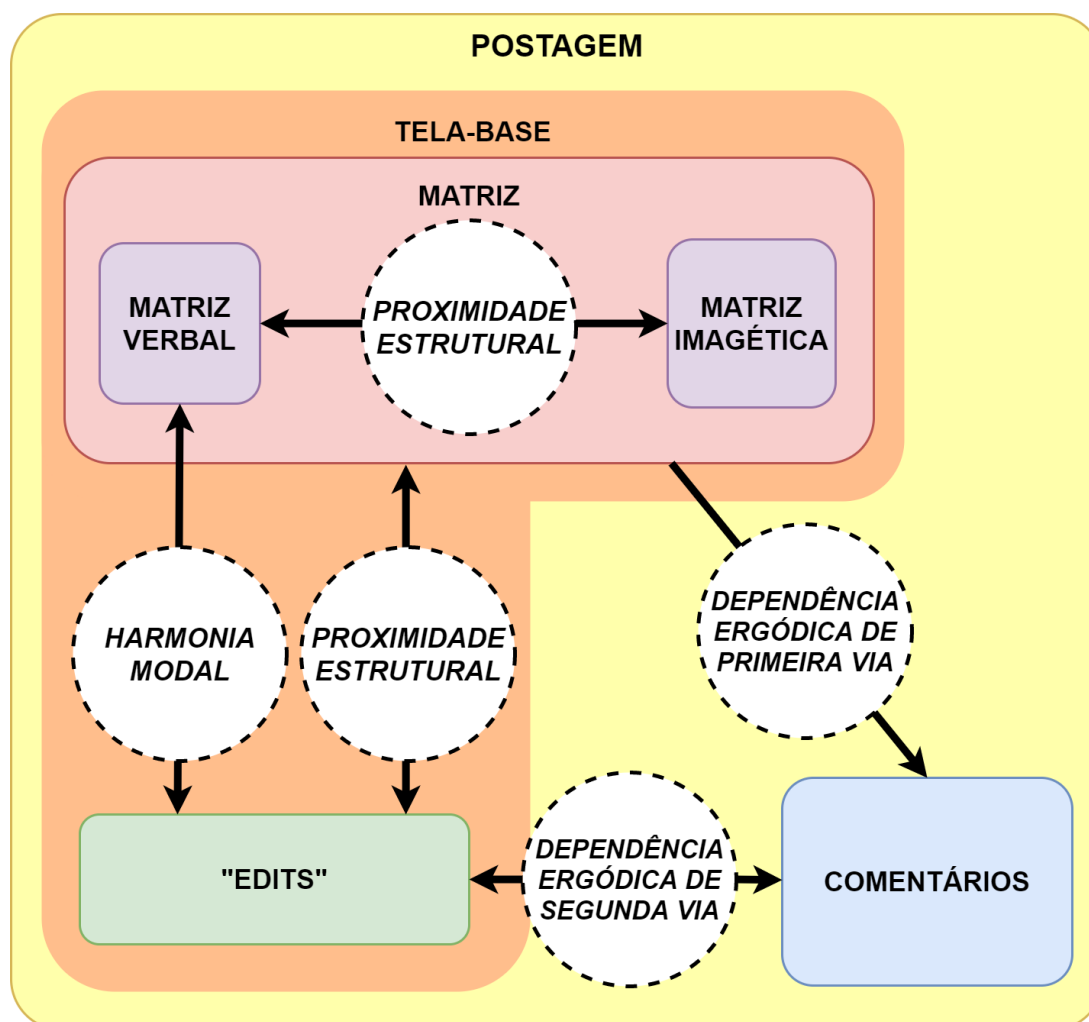


Figura 11. Algumas relações básicas em postagens com matriz verbo-imagética e “edits”

Fonte: elaboração própria.

12. Uma análise mais detalhada explicitaria mais telas e articulações. Por exemplo, entre um comentário (uma “unidade-C”) e suas respostas, há dependência ergódica de primeira via.

Assim como em nossa discussão sobre as telas de uma postagem de Facebook, nossas considerações sobre as articulações básicas entre as telas são sumárias, exemplificando o procedimento exposto na seção 2. Em uma pesquisa mais detalhada sobre tais relações, certamente outras articulações seriam exploradas, como as diversas proximidades estruturais a que não demos atenção e a maneira como as interfaces interagem com as demais telas.

Após o passo de número quatro, iniciam-se considerações analíticas que estão para além dos procedimentos aqui apresentados, mas que utilizam seus resultados como base. No caso de postagens de Facebook, exemplos de análises comparativas fundamentadas em um modelo semelhante ao proposto aqui são encontrados em Farhat e Gonçalves-Segundo (2021a, 2021b).

CONSIDERAÇÕES FINAIS

Neste artigo, apresentamos noções e procedimentos que consideramos produtivos para pesquisadores que queiram se engajar na análise de textos multimodais. Esperamos que tais propostas facilitem a organização do complexo trabalho que é a análise multimodal, proporcionando bases teoricamente motivadas para a investigação empírica de como potenciais semióticos distintos são instanciados em uma mesma situação, articulando-se em telas que fazem da materialidade física uma materialidade cultural. Com tais bases, que protegem o pesquisador do risco de vagueza e organizam o procedimento analítico, facilita-se o processo tão necessário de aplicação dos achados teóricos na prática educacional, o que exige sempre o máximo de transparência.

Deve-se ressaltar, porém, que o modelo proposto não deve ser tomado como dogma, mas como recomendação: sua adoção deve ser crítica, observando sua pertinência e utilidade e, sempre que possível, aprimorando-o. Por exemplo, como os estudos de multimodalidade estão ligados, historicamente, às questões da visualidade, pode ser necessário elaborar novos parâmetros para textos instanciados somente em telas sonoras. Ademais, mesmo textos visuais, mas diferentes de postagens de Facebook, terão especificidades de que não tratamos aqui – sugerimos, por isso, a leitura dos diversos exemplos de análise que Bateman, Wildfeuer e Hiippala (2017) oferecem em sua obra.

Enfim, é evidente que este texto cobre parte pequena dos desenvolvimentos recentes de uma área em expansão. Por isso, além de Bateman, Wildfeuer e Hiippala (2017), recomendamos fortemente a leitura de obras como Kress (2010), Forceville (2020), Jewitt, Bezemer e O'Halloran (2016), além dos capítulos em Norris e Maier (2014). Para uma apresentação de pesquisas sobre multimodalidade no Brasil, ver Gualberto e Santos (2019). Essas obras refletem a diversidade de abordagens, posicionamentos teórico-metodológicos e debates do campo; se os estudos de multimodalidade se firmarem como uma disciplina, ela certamente não será homogênea.

REFERÊNCIAS

- BATEMAN, J. A. (2008). *Multimodality and Genre: A Foundation for the Systematic Analysis of Multimodal Documents*. New York: Palgrave.
- BATEMAN, J. A.; WILDFEUER, J.; HIIPPALA, T. (2017). *Multimodality – Foundations, Research and Analysis: A Problem-Oriented Introduction*. Berlim: de Gruyter.
- BENVENISTE, E. (1974). *Problemas de linguística geral II*. 2a ed. Campinas: Pontes, 2006.
- COPE, B.; KALANTZIS, M. (2009). "Multiliteracies": New literacies, new learning. *Pedagogies: An international journal*, v. 4, n. 3, p. 164-195.
- EKMAN, P.; FRIESEN, W. V.; ELLSWORTH, P. (1982). What emotion categories or dimensions can observers judge from facial behavior? In: EKMAN, P. (ed.) *Emotion in the human face*. Nova York: Cambridge University Press, p. 39-55.
- FARHAT, T. C.; GONÇALVES-SEGUNDO, P. R. "Manas, preciso de ajuda": análise de pedidos de ajuda multimodais de um grupo de Facebook. *Texto Livre: Linguagem e Tecnologia*, v. 14, n. 1, p. 1-19, 2021a.
- FARHAT, T. C.; GONÇALVES-SEGUNDO, P. R. Identidades em comunhão: estratégias multimodais de individuação em um grupo de Facebook. *Texto Digital*, v. 17, n. 2, p. 35-71, 2021b.

- FORCEVILLE, C. (2020). *Visual and Multimodal Communication: Applying the Relevance Principle*. New York: Oxford University Press.
- GUALBERTO, C. L.; SANTOS, Z. B. (2019). Multimodalidade no contexto brasileiro: um estado de arte. *DELTA: Documentação de Estudos em Linguística Teórica e Aplicada*, v. 35, p. 1-30.
- HALLIDAY, M. A. K.; MATTHIESSEN, C. M. I. M. (2014). *Introduction to Functional Grammar*. 4a ed. New York/London: Routledge.
- JEWITT, C.; BEZEMER, J.; O'HALLORAN, K. (2016). *Introducing multimodality*. London: Routledge.
- KRESS, G. (2010). *Multimodality: A social semiotic approach to contemporary communication*. Oxon: Routledge.
- KRESS, G.; VAN LEEUWEN, T. (1996). *Reading Images: The Grammar of Visual Design*. 2a ed. London/New York: Routledge, 2006.
- MARTIN, J. R. (1992). *English Text: system and structure*. Amsterdam: John Benjamins.
- MATTHIESSEN, C. M. I. M. (2009). Multisemiosis and context-based register typology. In: VENTOLA, E.; GUIJARRO, A. J. M. (org.). *The World Told and the World Shown: Multisemiotic Issues*. London: Palgrave Macmillan, p. 11-38.
- NORRIS, S.; MAIER, C. D. (org.). (2014). *Interactions, Images and Texts: A Reader in Multimodality*. Berlin: de Gruyter.
- O'TOOLE, M. (1994). *The Language of Displayed Art*. London: Leicester University Press.

Recebido: 14/8/2021

Aceito: 22/6/2022

Publicado: 14/7/2022