

Multimodalidade e contexto: problemas, assunções e hipóteses

Multimodality and context: problems, assumptions and hypotheses

Theodoro C. FARHAT (USP)
theo.cfar@gmail.com

Recebido em: 31 de ago. de 2022.
Aceito em: 06 de nov. de 2022.

FARHAT, Theodoro C. Multimodalidade e contexto: problemas, assunções e hipóteses. **Entrepalavras**, Fortaleza, v. 12, n. 3, e2536, p. 38-61, set.-dez./2022. DOI: 10.22168/2237-6321-32536.

Resumo: Tomando como base a orientação paradigmática e sociossemiótica oferecida pela Teoria Sistêmico-Funcional, este artigo explora a seguinte questão: assumindo uma relação de constituição mútua entre texto e contexto de situação, como isso se dá nos artefatos multimodais? Por exemplo, uma imagem pode realizar os mesmos elementos culturais que um enunciado verbal (e vice-versa)? Para sugerir algumas vias de resposta, examinaremos como as dimensões globais de estratificação, instanciação e metafunção podem explicitar os elementos que estão em jogo quando se enquadram as relações entre linguagem e contexto, assim como as consequências de tais teorizações para a compreensão da multimodalidade como fenômeno sociossemiótico. A isso se segue uma ilustração analítica e, enfim, apresentamos considerações finais sobre questões de que pesquisas futuras podem se ocupar.

Palavras-chave: Estratificação. Realização. Parâmetro contextual.

Abstract: Drawing on the paradigmatic and sociosemiotic orientation provided by Systemic Functional Theory, this paper explores the following question: assuming a mutually constitutive relationship between text and context of situation, how does this play out in multimodal artifacts? For example, can an image realize the same cultural elements as a verbal instance (and vice versa)? To suggest possible answers, we will examine how the global dimensions of stratification, instantiation and metafunction can make explicit the elements that are at play when framing the relationship between language and context, as well as the consequences of such ideas for understanding multimodality as a sociosemiotic phenomenon. This is followed by an analytical illustration and, finally, by concluding remarks on issues that future research may address.

Keywords: Stratification. Realization. Contextual parameter.

Introdução

O fenômeno da multimodalidade é intrínseco à comunicação humana desde os seus primórdios: interagir multimodalmente é regra, não exceção (KRESS, 2010; BATEMAN *et al.*, 2017). Nos últimos trinta anos, porém, diversas disciplinas e teorias passaram a focalizar a integração e coordenação de diferentes modos semióticos como um fenômeno que merece atenção própria – houve uma espécie de “virada multimodal” na semiótica, na análise do discurso e em outras áreas. Isso implica tanto um enriquecimento epistemológico quanto uma complexificação exponencial de questões que, quando vistas sob uma ótica monomodal, já eram complexas. Uma dessas problemáticas, que consideramos fundamental, é a relação entre linguagem e contexto.

Tendo isso em vista, este artigo discutirá algumas questões que emergem quando se tenta responder a seguinte pergunta: considerando que há uma relação de constituição mútua entre texto e contexto de situação, como isso se dá no caso das produções multimodais, isto é, em artefatos semioticamente heterogêneos? Mais especificamente, uma vez que contexto e linguagem são vistos como redes de paradigmas (redes sistêmicas) associadas em termos de realização, é possível afirmar que dois modos semióticos distintos realizam os mesmos paradigmas contextuais?¹ Para sugerir alguns caminhos de respostas – e apontar outras questões –, exploraremos como teorizações sistêmico-funcionais podem explicitar os elementos que estão em jogo quando se enquadram as relações entre linguagem e contexto.

O artigo está organizado da seguinte forma: em primeiro lugar, introduzimos as três dimensões semióticas globais postuladas

¹ Pesquisas anteriores, como Matthiessen (2009) e Bowcher (2013), exploraram perguntas semelhantes, mas por vias distintas das que seguimos aqui.

na Teoria Sistêmico-Funcional (TSF): estratificação, metafunção e instanciação, que permitem explicitar como linguagem e contexto se articulam. Depois, apresentamos uma definição sistêmico-funcional para a multimodalidade e uma série de implicações de tal definição, o que possibilita uma discussão refinada dos problemas envolvidos na relação contexto-multimodalidade. A isso se segue uma ilustração analítica e, enfim, expomos nas considerações finais algumas questões de que pesquisas futuras poderão se ocupar.

Princípios sistêmico-funcionais: estratificação, metafunção e instanciação

A TSF é organizada em três dimensões semióticas globais: instanciação, estratificação e metafunção (HALLIDAY; MATTHIESSEN, 2014). A **instanciação** diz respeito à ideia de que é possível investigar a semiose em diferentes níveis de generalidade: o que é (uma **instância**), o que *pode ser* (um **sistema**, entendido como um **potencial** de significação) ou ainda o que *costuma ser* em dada configuração contextual (os **subpotenciais** – registros e gêneros). Assim, enquanto este artigo é um texto (uma instância) do potencial do português brasileiro (o sistema), ele também é resultado da instanciação do registro acadêmico (um subpotencial bastante geral) e do gênero “artigo científico” (um subpotencial mais específico).

Como já se pode entrever no parágrafo anterior, a TSF também estabelece que há uma relação sistemática entre linguagem e contexto. Isso é explicitado pela dimensão de **estratificação**, que compreende a noção de que há uma hierarquia de estratos semióticos: em primeiro lugar, uma divisão entre linguagem e contexto; em segundo, uma divisão entre conteúdo e expressão. Os estratos estão associados pela relação de **realização**: linguagem realiza contexto; expressão realiza conteúdo:

- contexto \searrow linguagem
- contexto \searrow (conteúdo \searrow expressão)

Na linguagem verbal, há ainda outras divisões: o conteúdo é estratificado em semântica e léxico-gramática, enquanto a expressão é estratificada em fonologia e fonética, de modo que temos o seguinte esquema realizacional:

- contexto \searrow ((semântica \searrow léxico-gramática) \searrow (fonologia \searrow fonética))

Cada estrato é descrito sistemicamente, isto é, em termos de redes de paradigmas que representam as opções naquele nível de

abstração simbólica: exemplos são a oposição entre distância social “íntima” e “impessoal” no contexto, entre ordens e perguntas na semântica, entre modo imperativo e indicativo na gramática e entre movimento tônico ascendente e descendente na fonologia.

O mais importante, aqui, é notar que, com esse esquema, o contexto – o estrato em que descrevemos elementos culturais semioticamente relevantes² – entra em uma relação de constituição mútua com a linguagem: a realização é ao mesmo tempo **ativação** – os elementos contextuais “ativam” elementos semióticos – e **construção** – os elementos linguageiros “constroem” elementos contextuais (HASAN, 2009). O exemplo mais claro está, talvez, no uso de apelidos: pode-se usar apelidos porque há uma intimidade entre os interlocutores (ativação) ou porque se pretende fabricar uma intimidade (construção). Em muitos casos, entretanto, os dois movimentos são simultâneos, de modo que a categoria geral de realização permanece bastante útil: apelidos realizam (constroem e/ou são ativados por) distância social reduzida.

Notemos, ademais, que a relação de realização opera em todos os níveis de generalidade, isto é, nos diversos pontos ao longo da escala de instanciamento. Assim, focando nas relações entre linguagem e contexto, temos: 1. no polo do potencial, um contexto de cultura (uma rede de opções culturais semioticamente relevantes, descritas em termos de Campo, Relações e Modo – ver abaixo) é realizado pelo potencial de significação (uma rede de opções de conteúdo e de expressão); 2. no polo da instância, um contexto de situação (os elementos culturais concretos, como as atividades culturais empreendidas e os papéis sociais assumidos pelos interactantes) é realizado pelo texto (os elementos semióticos concretos); e, 3. na região intermediária dos subpotenciais, uma configuração cultural típica é realizada por escolhas semióticas típicas (por exemplo, os papéis agentivos de professor e aluno são realizados em gêneros como a aula e a prova, o que tem consequências semióticas importantes, como certos padrões de polidez). Ver Figura 1.

² Note-se, portanto, que, embora o sociossemiotista tenha como parte de seus objetivos a descrição de elementos culturais, seu trabalho não se confunde com o do antropólogo; mantém-se firmemente um ponto de vista *semiótico* sobre a cultura.

Figura 1 – Realização e instanciação



Fonte: elaboração própria.

Enfim, a última dimensão global, a das **metafunções**, explicita a orientação funcional da teoria: metafunções são funções que se generalizam e, assim, alcançam o nível do potencial, de modo que seu uso é quase inevitável – o caso (verbal) prototípico é o das orações, em que há uma estrutura para cada uma das metafunções; autores como Kress e van Leeuwen (2006) indicam como esse princípio pode ser expandido para outras semioses, como as imagens e a música. Propõem-se três metafunções: a interpessoal, que compreende os recursos semióticos especializados na encenação de papéis e relações; a ideacional, que se ocupa dos recursos de construção e articulação de experiências; e a textual, que lida com os recursos composicionais da linguagem – como um conjunto de elementos linguísticos é urdido para formar um texto.

Mais relevante ainda para os propósitos deste artigo, porém, é a hipótese de que, no processo de realização, os elementos culturais estão alinhados metafuncionalmente. Em outras palavras, há uma hipótese de **ressonância entre contexto e metafunções** (HASAN, 2014) – cada metafunção seria especializada na realização de um conjunto específico de opções culturais: sistemas ideacionais realizam opções de Campo³ (as possibilidades de ação em uma dada cultura); sistemas interpessoais realizam as Relações interactanciais (as relações e papéis potencialmente assumidos pelos interactantes); e sistemas textuais realizam Modo (possibilidades de contato – escrito, oral, digital, etc.).

Não temos espaço, aqui, para expor os principais sistemas verbais e pictóricos que utilizaremos na nossa ilustração analítica; as referências fundamentais são: para os elementos verbais, Halliday e

³ Seguimos, aqui, a convenção utilizada, por exemplo, por Bowcher (2014), de utilizar iniciais maiúsculas para os parâmetros contextuais.

Matthiessen (2014), Martin e Rose (2007) e Figueredo (2011); e, para os elementos pictóricos, a Gramática do Design Visual de Kress e van Leeuwen (2006).⁴ Como, entretanto, os sistemas contextuais que utilizaremos, propostos por Hasan (1999, 2014, 2020), são relativamente pouco conhecidos, apresentamos no Quadro 1 uma síntese (brevíssima) dos principais sistemas de cada parâmetro contextual:

Quadro 1 – Principais sistemas contextuais nas propostas de Hasan

Parâmetro	Sistema	Glosa
Campo	AÇÃO MATERIAL	Descreve ações não verbais que podem ter impacto sobre a ação verbal.
	AÇÃO VERBAL	Descreve o tipo de ação verbal (assemelha-se, em grande medida, à proposta topológica de Matthiessen (2015) – ver Farhat e Gonçalves–Segundo (<i>submetido</i>)).
	ESFERA DE AÇÃO	Descreve o domínio cultural em que a ação é realizada em termos de: 1. institucionalização (o grau de coerção explícita sobre a ação realizada); 2. especialização (a distância da ação realizada em relação ao “senso comum” da cultura).
Relações	PAPÉIS AGENTIVOS	Descreve os papéis assumidos pelos interactantes em termos das suas ações, como “professor–aluno” e “médico–paciente”.
	PAPÉIS TEXTUAIS	Descreve os papéis em termos da produção e da recepção do texto – enunciador e enunciatário (que pode ou não estar presente no momento de produção, ser uma pessoa específica ou uma categoria, etc.).
	PAPÉIS SOCIAIS	Descreve diversos elementos, incluindo: 1. status ((des)igualdade de poder); 2. distância social (intimidade, (im)ressoalidade)); 3. atributos identitários (“adquiridos” ou “imputados”): gênero, idade, capital simbólico e material, etc.
Modo	CONTATO MATERIAL	Descreve o contato em termos “sensíveis”: o “canal” do contato (oral, escrito, etc.).
	CONTATO SEMÂNTICO	Descreve o contato em termos “inteligíveis”: por exemplo, o potencial de troca de turnos.

Fonte: elaboração própria.

Assim, por exemplo, o texto deste artigo realiza as seguintes opções (há várias outras; a exposição é meramente didática):

⁴ Introduções acessíveis em português podem ser encontradas em Fuzer e Cabral (2014) (para a Gramática Sistêmico–Funcional) e Nascimento, Bezerra e Heberle (2011) (para a Gramática do Design Visual).

- Campo: AÇÃO VERBAL epistêmica; ESFERA DE AÇÃO especializada. Isso é evidenciado, por exemplo, pelo uso de terminologia especializada.
- Relações interactanciais: STATUS EPISTÊMICO do enunciatário: elevado. Em outras palavras, assume-se que o autor sabe algo que o leitor quer saber, o que é evidenciado pelo uso quase absoluto de orações declarativas.
- Modo: CONTATO MATERIAL escrito. Isso se evidencia em padrões coesivos e de estrutura temática (Tema-Rema) típicos da modalidade escrita da língua.

Como se pode inferir pelos parágrafos anteriores, esses desenvolvimentos teóricos da TSF – as dimensões semióticas globais e suas implicações – foram propostas originalmente com vistas à descrição da linguagem verbal. O que pretendemos fazer neste artigo é questionar: quais as implicações dessas teorizações para o estudo da multimodalidade em contexto? E, no sentido oposto: quais as implicações da multimodalidade para o estudo do contexto como estrato semiótico? Antes, porém, apresentaremos uma definição sistêmico-funcional da multimodalidade.

Definindo a multimodalidade: assunções e implicações

Muito já se discutiu sobre a natureza da multimodalidade – sugerimos fortemente a leitura de, por exemplo, Kress (2010) e Bateman et al. (2017). Em Farhat e Gonçalves-Segundo (2022, p. 438, *itálicos no original*), propusemos a seguinte definição:

um texto é multimodal por resultar da instancição, 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.

Tentamos, com isso, utilizar as dimensões sistêmico-funcionais de instancição e estratificação como ferramentas teóricas, o que é especialmente útil para pensar as relações linguagem-contexto em termos explícitos. Seja como for, essa definição tem implicações e desdobramentos que precisam ser explicitados:

1. Todo modo semiótico realiza (ou seja, constrói e é ativado por) um estrato contextual – é inconcebível que um modo semiótico genuíno opere de uma maneira totalmente descontextualizada, mesmo que esse contexto seja muito distinto daquele em que operam os textos verbais.
2. Assim como o modo verbal, modos não verbais realizam, no polo da instância, um contexto de situação e, no polo do potencial, um contexto de cultura.
3. Há variados graus de integração entre diferentes modos semióticos, mas é presumível que, quanto mais “elevado” o estrato, mais fácil será a integração: por exemplo, o contexto de situação de uma imagem e o dos enunciados verbais que a acompanham devem ser aproximadamente equivalentes e provavelmente haverá ligações semântico-discursivas verbo-pictóricas nítidas, mas a associação entre a expressão da imagem e a da linguagem verbal será, em geral, mais indireta. Isso é ainda mais evidenciado quando a expressão de cada modo é acessada por sentidos distintos: a trilha sonora de um filme certamente pode ter laços “coesivos” com a imagem, mas isso se dará sobretudo por meio de associações semântico-discursivas (p. ex.: imagem “triste” – música “triste”).

O contexto dos textos multimodais: problemas e hipóteses

Partimos, então, dessas assunções, para enunciar as seguintes questões, que têm como foco especialmente os textos verbo-pictóricos, mas podem abarcar as instâncias multimodais como um todo: em que grau o contexto de situação de uma imagem é o mesmo de um enunciado verbal instanciado concomitantemente? E, no nível do potencial, o contexto de cultura realizado pelo sistema verbal é realizado da mesma forma pelo sistema pictórico?

Começando pela segunda pergunta, parece-nos razoável afirmar que, se dois modos semióticos são frequentemente instanciados em um só texto, eles fazem parte da mesma cultura. Em um primeiro momento, isso poderia servir de indício para argumentar que, no nível do potencial, não haveria contextos diferentes – os diferentes potenciais semióticos seriam simplesmente realizações distintas de um mesmo contexto de cultura.

Tal raciocínio talvez seria válido e até óbvio em uma reflexão sociológica ou antropológica sobre o que é uma cultura, mas devemos recordar que a descrição de um contexto de cultura (um potencial de situações) em uma perspectiva sociossemiótica não é equivalente à feita nas ciências sociais: para o sociossemiotista, a ênfase não está na cultura como um todo, mas nos seus elementos que, por estarem sistematicamente associados a padrões semióticos, podem explicar: (1) no polo da instância, as ativações e construções que ocorrem ao longo de um texto; (2) na região dos subpotenciais, padrões sociossemióticos probabilísticos, descrevendo o uso em tipos de contextos; (3) no polo do potencial, de que modo o sistema se organiza para realizar as potencialidades situacionais – evidentemente, nesse ponto a empreitada passa a ser mais especulativa.

O problema de tais concepções é que o que, em uma cultura, é semioticamente relevante para um modo não necessariamente o é para outro. Por exemplo, como lidar com o sistema de AÇÃO VERBAL para imagens? E com o CONTATO MATERIAL oral? Essas considerações sugerem que seria sensato postular diferentes contextos de cultura para diferentes modos semióticos. Isso, porém, entra em conflito com o raciocínio inicial de que modos diferentes podem operar em uma *mesma* cultura, embora também pareça sensato crer que deve haver ao menos alguns elementos comuns entre os diferentes potenciais de situação – caso contrário, as possibilidades de integração entre o contexto de situação de uma imagem e o de um enunciado verbal seriam muito menores do que a que, pela última assunção elencada na seção anterior, esperaríamos.

Assim, parece-nos desejável buscar uma saída conciliatória, em que se reconhecessem: 1. por um lado, no contexto de cultura, alguns elementos comuns a todos os modos semióticos; 2. e, por outro, outras variáveis, que caracterizariam a cultura como um fenômeno em si heterogêneo, tendo múltiplas possibilidades de realização (e, conseqüentemente, de instanciação).

A obra de Hasan (1999, 2014, 2020, etc.), embora não tome a multimodalidade como um tema central, fazendo somente alguns comentários esporádicos, pode fornecer uma primeira possibilidade de conciliação: a hipótese ARC: *Ação, Relações e Contato*. Essa hipótese surgiu da procura de um “quadro contextual que poderia ser aplicado a qualquer prática social em que a mediação cultural não seja limitada aos meios linguísticos ou fortemente dependente de ‘convenções

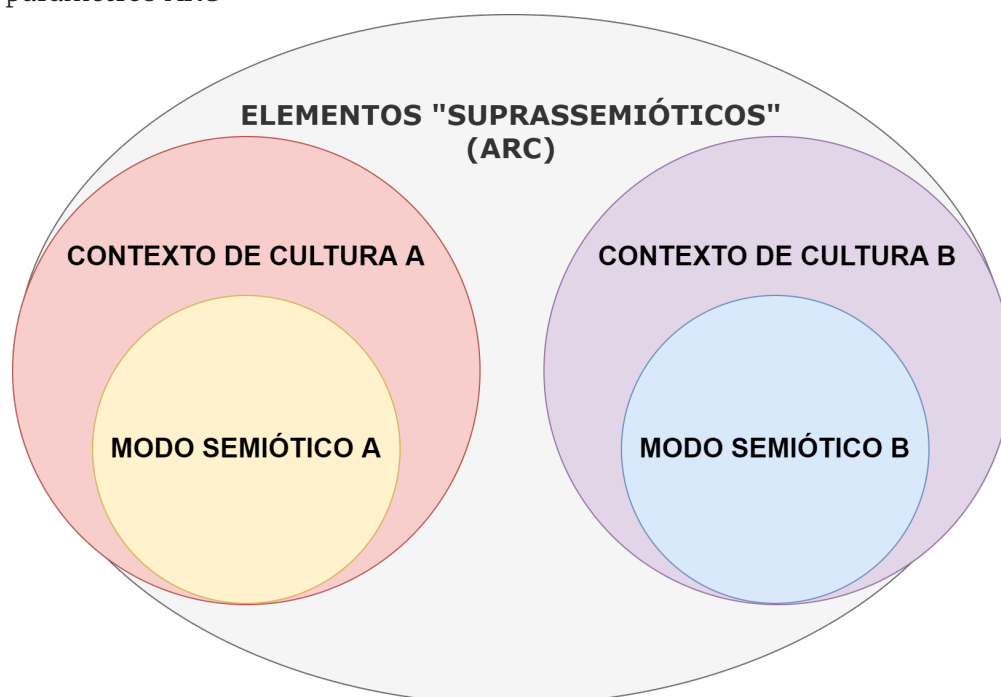
sociais’” (HASAN, 2014, p. 11-12; todas as traduções são nossas). Em termos bastante gerais, propõe-se que os três parâmetros contextuais (Campo, Modo e Relações) seriam, na realidade, versões linguisticamente específicas de três parâmetros relevantes para qualquer prática social:

Uma prática social acarreta alguma ação (atividade/fazer/ato) suficientemente significativa para atrair coengajamento. Sendo social, implica-se alguma relação entre os actantes, mesmo que ela surja simplesmente do ato de fazer algo junto [...]: na natureza das coisas, cada participante é posicionado socialmente (Bernstein, 1990); eles não podem deixar de ser ‘relacionados’. E deve haver alguma maneira de estabelecer contato entre os atores dessas ações, isto é, algum meio de possibilitar acesso ao que está acontecendo; sem isso, tanto a iniciação quanto o prosseguimento da atividade compartilhada seriam prejudicados. (HASAN, 2014, p. 12).

Assim, os três parâmetros gerais, para análise de qualquer prática social, seriam: Ação (em sua versão especificamente linguística, Campo); Relações (linguisticamente, *Tenor* – termo traduzido para o português justamente como Relações); e Contato (linguisticamente, o Modo). Assim como a prática verbal, práticas semióticas não verbais lidariam de maneira distinta com os três parâmetros, possivelmente os organizando sistemicamente (como no caso dos parâmetros contextuais linguísticos). Com isso, podemos nos comprometer com uma caracterização geral da cultura (um potencial geral de Ações, Relações e Contatos, possibilitando certas práticas sociais e impossibilitando outras) e, ao mesmo tempo, com o fato de que esses três parâmetros gerais têm relações diferentes com cada prática social específica.⁵ Ver Figura 2.

⁵ Essa ideia não é, na realidade, nova: Halliday (1975, p. 36), por exemplo, escreveu que “A própria cultura é um sistema semiótico, um sistema de significados ou informações que são codificados no potencial de comportamento de seus membros, incluindo seu potencial verbal – isto é, seu sistema linguístico. O sistema linguístico é somente uma forma de realização do sistema semiótico mais geral que constitui a cultura”.

Figura 2 – Modos semióticos, seus respectivos contextos de cultura e os parâmetros ARC



Fonte: elaboração própria.

Pode-se supor, porém, que, como partem dos mesmos parâmetros gerais, haverá inevitavelmente semelhanças muito grandes entre a organização, por exemplo, dos parâmetros de Ação^v (ação na prática verbal – Campo) e Açãoⁱ (ação na prática imagética). Essa semelhança básica seria o que permitiria a sensação de forte integração entre os contextos de situação e, de fato, seria extremamente desejável teorizar sobre quais aspectos de cada um dos parâmetros seriam os mais “essenciais”, porque talvez possamos encontrar aí um ponto nevrálgico da integração multimodal.

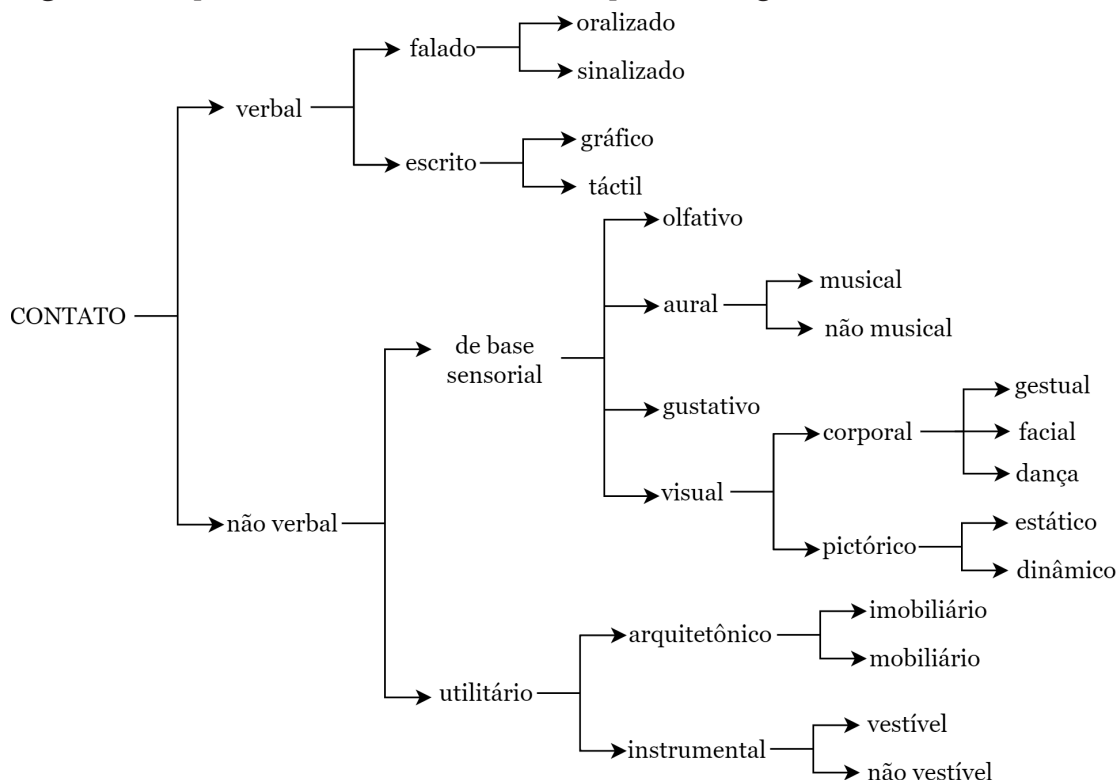
Por exemplo, em relação aos sistemas contextuais propostos por Hasan (1999, 2014, 2020), é provável que os sistemas de ESFERA DE AÇÃO, PAPÉIS AGENTIVOS, PAPÉIS TEXTUAIS e PAPÉIS SOCIAIS seriam, sem muita dificuldade, generalizados para outros modos semióticos. Entretanto, o mesmo não pode ser dito sobre os dois sistemas de ação (material e verbal), questão discutida por Bowcher (2013, 2014), e pelo sistema de CONTATO MATERIAL (que distingue, por exemplo, contato oral e contato gráfico com a linguagem verbal), o que torna o Modo um parâmetro bastante problemático em contextos multissemióticos. Tratem-se deste último problema.

Na literatura sistêmico-funcional e sociossemiótica, geralmente se usa o termo *mode* (modo) com dois sentidos distintos,

mas fortemente relacionados: o parâmetro contextual (Modo) e o modo semiótico, que é geralmente compreendido como um sistema semiótico: imagens estáticas, em movimento, sons musicais, não musicais, gestos, etc. Como vimos acima, o parâmetro de Modo, sendo um parâmetro contextual da linguagem verbal (isto é, do modo semiótico verbal), centra-se nas possibilidades de contato com tal linguagem. Entretanto, também se deve notar que Hasan (2014, 2020) inclui em suas hipóteses mais delicadas de descrição das variáveis de Modo algumas opções que envolvem modos não verbais. Embora certamente possa ser produtivo em um primeiro momento, isso nos parece potencialmente problemático, já que vai justamente contra a reflexão de Hasan sobre o fato de que Campo, Relações^v e Modo são parâmetros da linguagem *verbal*, casos mais específicos dos parâmetros gerais ARC.

Como tratar, então, do fato de que a presença de diferentes modos implica diferentes casos do parâmetro geral de Contato? Parece-nos que a resposta para essa pergunta pode estar em uma representação (inicial e altamente hipotética) desse parâmetro geral como uma rede sistêmica em que se apresentam as possibilidades gerais de Contato de uma cultura – os diferentes modos semióticos. Uma possibilidade para tal rede sistêmica está representada na Figura 3.

Figura 3 – Representando sistemicamente o parâmetro geral de Contato



Fonte: elaboração própria.

Essa representação é somente um esboço e há diversas questões potencialmente polêmicas em relação a ela. Por exemplo, omite-se a conhecida continuidade entre os gestos não verbais e os gestos verbais (isto é, as línguas de sinais); além disso, como a divisão mais fundamental é entre “verbal” e “não verbal”, pode-se acusar o sistema de um certo “verbocentrismo”.⁶ Não temos o espaço, aqui, para abordar tais problemáticas, mas pretendemos simplesmente apontar que talvez seja possível fazer descrições como essa, indo além das práticas semióticas específicas. Isso certamente seria algo produtivo para os estudos sobre multimodalidade, mas é algo que ainda só está começando – para outras propostas para a (sub)classificação dos modos semióticos, ver Kalantzis e Cope (2020) e Forceville (2020).

Para o propósito prático da análise de (con)textos multimodais, sugerimos que se adotem os seguintes procedimentos: os modos semióticos podem ser analisados como realizações dos três parâmetros gerais (ARC); como as descrições delicadas que temos desses parâmetros se referem à sua dimensão verbal (oferecidas por Hasan (1999, 2014, 2020), Martin (1992), Matthiessen (2015), entre outros), é dela que o analista partirá, mas sempre com o cuidado de deixar claro que tais descrições foram propostas exclusivamente para enunciados verbais. Com sorte, com mais pesquisas empíricas, será possível propor variáveis contextuais específicas para modos não verbais ou, idealmente, explicitar quais opções são “suprassemióticas”, ultrapassando as barreiras da realização monomodal.

Uma ilustração analítica

Para ilustrar as questões discutidas anteriormente, apresentamos nesta seção uma breve ilustração. Trata-se da análise de uma postagem de Facebook instanciada em uma comunidade que se volta ideacionalmente à tradução e à interpretação. Utilizamos, para isso, o modelo decomposicional elaborado em Farhat e Gonçalves-Segundo (2022), que resulta na divisão da postagem nas seguintes “telas”:⁷

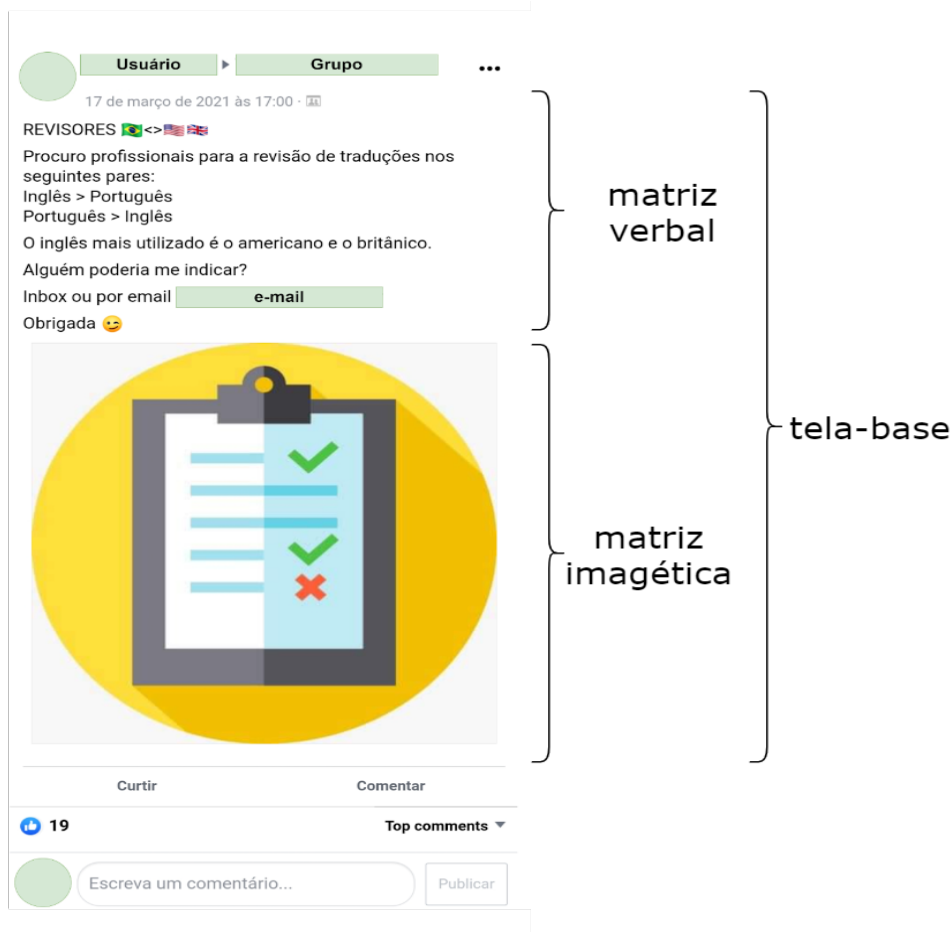
⁶ Além disso, há diversos outros pontos que mereceriam discussão: por exemplo, algumas oposições, por serem puramente negativas (“musical” X “não musical”; “vestível” X “não vestível”), assumem uma primazia da opção não negada. Essas discussões, entretanto, deverão ser realizadas em trabalhos futuros, em que cada “passo” do sistema será discutido em seus pormenores.

⁷ Esta decomposição é relativamente simplificada. Para os procedimentos que levam a ela e para uma versão mais refinada, recomendamos a leitura de Farhat e Gonçalves-Segundo (2022).

1. Tela-base: a seção nuclear da postagem, associada ao seu “postador”. É divisível em:
 - a. Matriz verbal: a seção verbal original da tela-base.
 - b. Matriz imagética: a seção pictórica original da tela-base.
2. Comentários: seção da postagem que se abre a outros enunciadores (os “comentadores”). É divisível em diversas unidades que, por sua vez, podem ser divididas em unidades menores, se recebem respostas.

A tela-base da postagem a ser analisada está representada na Figura 4 (elementos que poderiam identificar os usuários foram eliminados):

Figura 4 – A tela-base de uma postagem e suas subtelas.



Fonte: *corpus* coletado pelo autor e elaboração própria.⁸

A análise a seguir provém de um projeto em que investigamos padrões sociossemióticos – textuais e contextuais – em dois grupos de

⁸ Não apresentamos o endereço da postagem original para preservar a identidade dos usuários.

Facebook. Para isso, adotamos o seguinte procedimento: 1. em primeiro lugar, realizamos a coleta do *corpus* (oito postagens de cada grupo), com as condições de que ao menos metade do *corpus* de cada grupo teria tela-base verbo-pictórica (garantindo a investigação de questões multimodais) e todas as postagens teriam de apresentar ao menos vinte e cinco comentários (para que houvesse material semiótico suficiente nos comentários para que investigássemos interações minimamente complexas); 2. depois, cada postagem passou por uma análise textual fundamentada pela Linguística Sistêmico-Funcional: para elementos verbais, fizemos análises léxico-gramaticais (HALLIDAY; MATTHIESSEN, 2014; FIGUEREDO, 2011) e semântico-discursivas (MARTIN; ROSE, 2007); para elementos pictóricos, utilizamos a Gramática do Design Visual (KRESS; VAN LEEUWEN, 2006); 3. em seguida, realizamos uma análise contextual da postagem com base nos trabalhos de Hasan (1999, 2014, 2020); 4. enfim, chegamos a resultados globais: generalizações sobre padrões sociossemióticos nos grupos – por exemplo, verificou-se que um grupo realizava consistentemente uma ESFERA DE AÇÃO mais “institucionalizada” – isto é, havia práticas explicitamente coercivas por parte dos administradores da comunidade sobre que tipo de atividade discursiva era permitida e incentivada (para outros resultados, ver Farhat e Gonçalves-Segundo (*submetido*)).

Análise textual

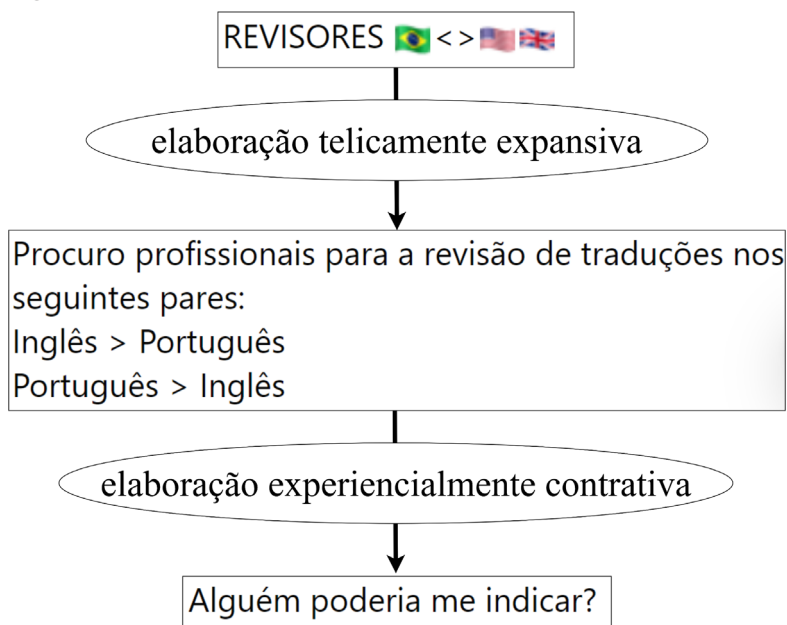
Começamos com a matriz verbal, que é iniciada por um segmento que, integrando signos linguísticos, emojis e sinais matemáticos, sintetiza o tópico e o propósito do texto: a busca por revisores de traduções de textos do inglês para o português e vice-versa. Os segmentos em seguida (*Procuro [...] O inglês [...]*) explicitam verbalmente tais informações, o que evidencia como o uso de modos não verbais pode ser útil para esse trabalho de “resumo” – note-se, porém, que as bandeiras nacionais, assim como sinais de “maior” e “menor” significando “tradução”, são tão convencionais quanto palavras, o que os distancia de outros signos visuais.

Em *Procuro profissionais para a revisão de traduções nos seguintes pares*, há claramente uma configuração de demanda, o que é explicitado em *Alguém poderia me indicar?*, uma oração interrogativa realizando indiretamente uma ordem (além da incongruência, note-se também o uso de modulação muito baixa (*poder* + pretérito imperfeito) como

estratégia de polidez,⁹ diminuindo o peso da demanda sobre os enunciatários).

Assim, nesses três primeiros segmentos, verificam-se dois movimentos de elaboração: do primeiro para o segundo, tem-se uma elaboração *telicamente expansiva* – o que antes fora sintetizado em seis signos agora é elaborado em uma cadeia muito mais complexa, explicitando o propósito da postagem; do segundo para o terceiro, porém, a elaboração é *experiencialmente contrativa*: deixam-se elípticos os elementos originalmente introduzidos no primeiro segmento (e repetidos no segundo) para se isolar somente a razão de ser da postagem: demandar indicações. Ver Figura 5.

Figura 5 – Movimentos de elaboração na matriz verbal



Fonte: elaboração própria.

Essa “dança” elaborativa é provavelmente motivada por fatores diversos, mas um motivo central pode ser a polidez: assim como a incongruência no terceiro segmento apontada acima (e o *poderia*), o espraimento semântico dissolveria, textualmente, o peso cultural da demanda. Além disso, a polidez “cerimonial” também associaria a postagem a um profissionalismo por parte do postador, construindo não só um atributo social positivo para si (principalmente em termos de capital simbólico (BOURDIEU, 2007)), mas também um grau alto de

⁹ Utilizamos aqui a concepção de Brown e Levinson (1987) de polidez – um ato é “polido” para manter as “faces” dos interactantes: a face positiva (o desejo de ser apreciado, sentir-se afiliado) e a negativa (o desejo de ser desimpedido, sentir-se em poder).

especialização de esfera (o que pode afastar “amadores”). Em conjunto, esses elementos são claramente vantajosos para a demanda realizada.

O quarto segmento é de natureza mais “concentrada”: o postador dá as instruções sobre como responder à demanda – por “inbox” (mensagem privada do Facebook) ou e-mail. Tal explicitação também pode contribuir para os elementos indicados acima: contraste-se a matriz em questão com, por exemplo, “Alguém pode me indicar revisores de tradução entre inglês e português?”. Além de possíveis ambiguidades que, na postagem real, são resolvidas no segundo segmento (por exemplo, o fato de que as traduções englobam as duas “direções”), essa matriz provavelmente também estaria sob maior risco de problemas de polidez e não construiria o mesmo Campo e as mesmas Relações que a da postagem, especialmente em termos de especialização (na ESFERA DE AÇÃO) e capital simbólico (nos PAPÉIS SOCIAIS).

Enfim, o último segmento, *Obrigada* 😊, além de sinalizar um “fechamento”, está em clara consonância com a polidez verificada anteriormente. O destaque, aqui, é o uso do emoji “piscando”: ao inscrever no texto uma atitude positiva, o signo rompe com o que se poderia interpretar como uma “frieza” que o profissionalismo e a especialização dos segmentos anteriores frequentemente carregam e, assim, talvez tente minimizar sutilmente a distância social, o que é estrategicamente oportuno para demandas como a que é feita. Tal uso, porém, pode trazer riscos, caso alguém decida interpretar “ironicamente” o emoji como um sinal de pessoalização excessiva, de “cumplicidade”. Evidencia-se, portanto, a tensão entre dois elementos comumente interpretados como positivos: o distanciamento profissional “objetivo” (com associações com “mérito”, “esforço”, etc.) e a aproximação afetiva subjetiva. O emoji, com a abstração de expressões fundamentalmente atitudinais, pode ser um bom mediador para esses dois pesos.

Passando para a matriz imagética, trata-se de uma representação bastante abstrata de uma prancheta em que há, aparentemente, uma lista de tarefas, o que é indicado por signos de “feito” e “não feito”.¹⁰ Nos termos representacionais (isto é, ideacionais) da Gramática do Design Visual de Kress e van Leeuwen (2006), teríamos aqui uma estrutura tipicamente analítica, dada a completa ausência de vetores: são os Atributos de um

¹⁰ Há, porém, uma outra possibilidade de interpretação – a imagem pode ser lida como uma representação do ato de *revisar*: o que interpretamos como “feito” seria, na realidade, um signo “correto”, enquanto o “não feito” seria um “incorreto”. A atividade de revisão envolveria a verificação dessas (in)correções. Nesse sentido, haveria maior “coesão” entre o verbal e o pictórico.

Portador que estão em jogo. Seu papel, porém, parece ir muito além de uma mera representação: o alto grau de abstração sugere que, além de possíveis objetivos de “decoração” e saliência visual que a imagem muito claramente pode ter (note-se que ela está “subordinada” ao texto verbal – a imagem é, em princípio, “prescindível”), a matriz imagética estaria em consonância com os esforços de exibição de profissionalismo “sério” e ao mesmo tempo “convidativo” (o que pode ser depreendido pela modalidade interpessoal (no sentido de Kress e van Leeuwen (2006)) e, mais especificamente, pelo uso diversificado de cores; imagine-se a mesma imagem, mas em preto e branco).

Assim, em síntese, a tela-base, como um todo, mostra uma orquestração bastante hábil em sua construção de significados: o “núcleo demandante” da postagem é rodeado por significados interpessoais que, em clara realização prosódica (eles se “espalham” em diversos segmentos e níveis – HALLIDAY, 1979), evidenciam a complexidade da articulação entre variáveis contextuais (principalmente capital simbólico de “profissionalismo” e alta especialização), polidez e os diferentes modos semióticos.

Vejamos como os comentários reagiram a tal “orquestra”. Os comentários primários – isto é, aqueles que não são respostas a outros comentários – se distribuem em duas categorias:

1. “Autoindicação” (83%): enunciados em que o comentador responde à demanda enunciada na matriz indicando a si mesmo como disponível ao trabalho. Isso se dá das seguintes formas:
 - a. Indicação de que as instruções na penúltima seção da matriz foram ou serão seguidas – o comentador anuncia (ou vai anunciar) seu interesse “inbox” ou por e-mail (39% do total). Exemplos: C1 – *Vou mandar email*; C3 – *Te enviei inbox!*; C8 – *Inbox*; C14 – *Olá, [marcação do postador]! Eu trabalho com esse par e já fiz revisão em inglês. Vou te mandar um e-mail.*
 - b. Anúncio de interesse, mas sem indicar que as instruções da penúltima seção da matriz foram ou serão seguidas; em alguns casos, o comentador pede que o postador o contate, invertendo as posições estabelecidas na matriz (44% do total). Exemplos: C6 – *Estou disponível*; C10 – *Eu sou tradutora e revisora. Pode me contatar por mensagem.*; C16 – *Para ingles britanico tenho interesse. Sou nativo do Reino Unido. Contato [número de telefone].*

2. “Heteroindicação” por meio de marcação (17%): nesses casos, o comentador indica outro usuário que poderia ter interesse no trabalho. Há, portanto, duas faces nesses comentários: por um lado, trata-se de respostas à demanda da matriz (*Alguém poderia me indicar?*); por outro, o comentador também está “chamando” um usuário conhecido para um trabalho em que pode ter interesse (e, se tiver, poderá seguir as instruções na matriz).

A interação entre matriz, comentários e respostas a comentários pode ser analisada mais detalhadamente em termos do sistema de NEGOCIAÇÃO proposto por Martin (1992); para os propósitos deste artigo, entretanto, basta apontar que houve movimentos interacionais que, de modo geral, faziam um “ajuste télico”, isto é, a ação de comentadores e postador é coordenada para chegar ao fim estabelecido na postagem: que os interessados no trabalho enviem uma mensagem manifestando seu interesse. Tais movimentos são, entretanto, quase exclusivamente verbais.

Análise contextual

Feita a análise textual, podemos executar uma análise contextual. Idealmente, isso significaria passar detalhadamente por todos os sistemas e subsistemas apresentados na seção 1; isso, entretanto, demanda um espaço de que não dispomos. Assim, sugerimos a seguinte (breve e parcial) análise contextual, focando em alguns sistemas dos parâmetros de Campo e Relações:

Quadro 2 – Uma (breve e parcial) análise contextual

Parâmetro	Sistema	Opções realizadas
Campo	AÇÃO VERBAL	Uma ação de “incitação” (ver Farhat e Gonçalves-Segundo, <i>submetido</i>) – o postador procura levar os leitores a manifestarem interesse de natureza profissional.
	ESFERA DE AÇÃO	Institucionalizada (o grupo em questão é explicitamente regrado, o que se reflete em práticas relativamente homogêneas) e especializada (note-se, por exemplo, que o postador dá como certo que haveria profissionais para a revisão de traduções no grupo, além de utilizar termos como “par” (de idiomas)).

Relações	PAPÉIS AGENTIVOS	Postador: “ofertante-contratante”. Comentadores: “interessados-candidatos”/ “indicadores”.
	PAPÉIS SOCIAIS	STATUS: desigual, embora haja certas ambiguidades – o “contratante” tem o poder de oferecer um trabalho, mas deve-se mostrar polido por sua posição “demandante” (isto é, a oferta tem uma faceta de pedido – um ato de ameaça à face, nos termos de Brown e Levinson (1987)). DISTÂNCIA SOCIAL: social-consultiva, nos termos de Hall (1966) – um grau entre o totalmente público e o claramente pessoal. ATRIBUTOS: destaque para a construção de capital simbólico na natureza “cerimonial” da oferta e em alguns dos comentários, especialmente os de “autoindicação” (p. ex. C17: <i>Sou tradutora juramentada</i>).

Fonte: elaboração própria.

A realização multimodal de um contexto: algumas possibilidades

Podemos, enfim, voltar à questão central deste artigo: considerando que há uma relação de constituição mútua entre texto e contexto de situação, como isso se dá no caso das produções multimodais? Para a postagem analisada, a resposta parece ser relativamente simples: como indicamos anteriormente, trata-se de um texto primariamente verbal – nos termos de Matthiessen (2009), a “divisão do trabalho semiótico” é assimétrica, priorizando a verbalidade. Isso tem claras implicações contextuais: na matriz, a construção de uma ação de demanda profissional (Campo) polida (Relações) se dá primariamente por vias verbais, mas utiliza “subordinadamente” elementos pictóricos (emojis e a matriz imagética) para dar sustentação a isso. Em outras palavras, embora os modos não verbais sejam contextualmente relevantes, operando claramente na realização de opções contextuais, tal realização se mistura à realização verbal, que tem uma força semiótica mais fundamental (nesta instância).

Isso parece indicar que, para avançarmos nos estudos da multimodalidade-em-contexto, será necessário, em primeiro lugar, dar atenção ao fato de que, ao encararmos um artefato multimodal, deveremos observar o grau de “prescindibilidade” de cada modo: pense-se por exemplo, na diferença entre um videoclipe e um filme – ambos usam imagens dinâmicas e música, mas o “peso” de cada modo é muito distinto em cada um, o que tem claras implicações contextuais:

as atividades empreendidas são distintas e mesmo as relações que se estabelecem com o público tem pontos divergentes. De fato, talvez seja necessário postular a “divisão do trabalho semiótico” como um parâmetro cultural em si (no nível ARC), explicitando a importância de tal dimensão como um fator na organização da heterogeneidade sociossemiótica de uma cultura.

Outro caminho de investigação estaria em hipotetizar uma “escala de nuclearidade contextual” análoga à escala de “imprescindibilidade” dos modos semióticos em uma instância. Em outras palavras, pode-se propor que, assim como em diferentes textos há diferentes “pesos” para os modos semióticos, também no nível contextual poderia haver diversos graus de (im)prescindibilidade para diferentes elementos contextuais – o que teria implicações para a investigação multimodal do contexto, visto que, como discutimos ao longo deste artigo, modos diferentes podem realizar diferentemente contextos (aparentemente) semelhantes. Por exemplo, algumas interações (como a explorada na ilustração analítica) têm um claro foco “acional”, buscando resultados “práticos”, o que levaria a um núcleo no parâmetro de Campo; outras, porém, têm um foco claramente interpessoal, posicionando o “núcleo contextual” nas Relações interactanciais.

Embora tal ideia esteja em parte em dissonância com o princípio de “não hierarquização” das metafunções, ela pode ser bastante produtiva para a investigação da multimodalidade em contexto: por exemplo, poderíamos esperar que diferentes “pesos” contextuais estejam em correlação com diferentes “pesos” semióticos. Além disso, é possível que certos modos sejam “especializados” na realização de alguns parâmetros contextuais.¹¹ Se tais hipóteses se confirmarem, haveria aqui também um caminho promissor para o estudo e a classificação de subpotenciais multimodais: a “nuclearidade” contextual pode servir de fator primário na distinção de diferentes práticas multimodais tipificadas.

Considerações finais

O propósito deste artigo foi dar atenção a uma questão que, embora intrinsecamente complexa, merece ser tratada como um tópico

¹¹ Pensamos, aqui, principalmente na (aparente) proeminência relacional dada por modos como as expressões faciais, que somente realizariam elementos acionais marginalmente.

de pesquisa central. Para isso, empregamos as dimensões globais da Teoria Sistêmico-Funcional e as propostas contextuais de Hasan como pontos de partida para considerações teórico-metodológicas que, embora somente iniciais, podem fornecer um primeiro fundamento para pesquisas posteriores.

Mais especificamente, com as considerações apresentadas na seção anterior, podemos sugerir a seguinte hipótese: a “divisão do trabalho semiótico” – isto é, o grau de “(im)prescindibilidade” de cada modo semiótico em um dado texto – tem importantes repercussões na “divisão do trabalho contextual”: quanto menos prescindível, mais fortemente um modo operará na realização de elementos culturais; por outro lado, se prescindível, espera-se que o modo terá contribuições contextuais “subordinadas” às do modo mais central. Pesquisas posteriores deverão testar essa hipótese empiricamente.

A prioridade, agora, é articular reflexões de diversas teorias e disciplinas sobre linguagem e cultura em modelos que, embora eternamente fadados à incompletude – não podemos esperar uma “teoria sociosemiótica de tudo” – fornecerão alguma base para pesquisas empíricas que levem a sério as tentativas de formalização das relações entre semiose e cultura. Trata-se de um empreendimento inevitavelmente coletivo: antes de tudo, a pesquisa deve partir de uma comunidade de pesquisadores que estejam abertos à experimentação teórico-metodológica – e que, do mesmo modo como superamos os limites da monomodalidade, ultrapassem as fronteiras da “monoteoricidade” e da “monodisciplinaridade”.

Referências

BATEMAN, J. A.; WILDFEUER, J.; HIIPPALA, T. **Multimodality – Foundations, Research and Analysis**: A Problem-Oriented Introduction. Berlim: Walter de Gruyter, 2017.

BOURDIEU, P. **A distinção**: crítica social do julgamento. Porto Alegre: Zouk, 2007.

BOWCHER, W. L. Material action as choice in field. In: FONTAINE, L.; BARTLETT, T.; O'GRADY, G. (eds.). **Systemic Functional Linguistics**: Exploring Choice. Cambridge: Cambridge University Press, 2013. p. 318–341.

BOWCHER, W. L. Issues for Developing Unified Systems for Contextual Field and Mode. **Functions of Language**, v. 21, n. 2, p. 176–209, 2014.

BROWN, P.; LEVINSON, S. C. **Politeness**: Some universals in language usage. New York: Cambridge University Press, 1987.

FARHAT, T. C.; GONÇALVES-SEGUNDO, P. R. Análise multimodal: noções e procedimentos fundamentais. **Trabalhos em Linguística Aplicada**, v. 61, p. 435-454, 2022.

FARHAT, T. C.; GONÇALVES-SEGUNDO, P. R. Entre o falar e o fazer: AÇÃO VERBAL e AÇÃO MATERIAL como parâmetros contextuais. Submetido para publicação.

FIGUEREDO, G. P. **Introdução ao perfil metafuncional do português brasileiro**: contribuições para os estudos multilíngues. 2011. Tese (Doutorado em Linguística Aplicada) – Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2011.

FORCEVILLE, C. **Visual and Multimodal Communication**: Applying the Relevance Principle. Oxford: Oxford University Press, 2020.

FUZER, C.; CABRAL, S. R. S. **Introdução à gramática sistêmico-funcional em língua portuguesa**. Campinas: Mercado de Letras, 2014.

HALL, E. T. **The Hidden Dimension**. New York: Doubleday, 1966.

HALLIDAY, M. A. K. **Learning how to mean**: Explorations in the Development of Language. London: Edward Arnold, 1975.

HALLIDAY, M. A. K. Modes of meaning and modes of expression: types of grammatical structure and their determination by different semantic functions. In: ALLERTON, D. J.; CARNEY, E.; HOLDCROFT, D. (eds). **Function and context in linguistic analysis**: a Festschrift for William Haas. Cambridge: Cambridge University Press, 1979. p. 57-79

HALLIDAY, M. A. K.; MATTHIESSEN, C. M. I. M. **Introduction to Functional Grammar**. 4a ed. New York/London: Routledge, 2014.

HASAN, R. Speaking with reference to context. In: GHADESSY, M. (ed.) **Text and Context in Functional Linguistics**. Amsterdam: John Benjamins, 1999. p. 219-328.

HASAN, R. The place of context in a systemic functional model. In: HALLIDAY, M. A. K.; WEBSTER, J. (eds.) **Continuum Companion to Systemic Functional Linguistics**. London: Continuum International, 2009. p. 166-189.

HASAN, R. Towards a paradigmatic description of context: systems, metafunctions, and semantics. **Functional Linguistics**, v. 1, n. 1, p. 1-54, 2014.

HASAN, R. Tenor: Rethinking interactant relations. **Language, Context and Text**, v. 2, n. 2, p. 213-333, 2020.

KALANTZIS, M.; COPE, B. **Adding Sense**: Context and Interest in a Grammar of Multimodal Meaning. Cambridge: Cambridge University Press, 2020.

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. 2a ed. London/New York: Routledge, 2006.

MARTIN, J. R. **English Text: System and Structure**. Philadelphia: John Benjamins, 1992.

MARTIN, J. R.; ROSE, D. **Working with Discourse: Meaning beyond the clause**. 2a ed. Continuum: London, 2007.

MATTHIESSEN, C. M. I. M. Multisemiosis and context-based register typology: Registerial variation in the complementarity of semiotic systems. In: VENTOLA, E.; GUIJARRO, A. J. M. (eds.). **The World Told and the World Shown: Multisemiotic Issues**. Palgrave Macmillan, London, 2009. p. 11-38.

MATTHIESSEN, C. M. I. M. Modelling context and register: The long-term project of registerial cartography. **Letras**, v. 25, n. 50, p. 15-90, 2015.

NASCIMENTO, R.; BEZERRA, F.; HEBERLE, V. Multiletramentos: iniciação à análise de imagens. **Linguagem & Ensino**, v. 14, n. 2, p. 529-552, jul./dez. 2011.