

# Interações imagem-texto: uma análise de composições multimodais instrucionais

## *Image-text interaction: an analysis on multimodal instructional compositions*

---

Sâmia Alves Carvalho\*

Universidade Estadual do Ceará

Fortaleza - Ceará / Brasil

**RESUMO:** Neste estudo, analisamos composições multimodais instrucionais em um contexto específico de ensino a distância com o objetivo de lançar luzes sobre como os processos de interação entre os modos de expressão visual e escrito se delineiam no contexto escolhido. Para isso, escolhemos a perspectiva das relações de status e lógico-semânticas desenvolvidas por Martinec e Salway e levamos em consideração também as interações semióticas delineadas por Roland Barthes. Foi escolhida como objeto de análise uma disciplina de um curso de Licenciatura em Letras – modalidade a distância. A disciplina tem como objetivo principal o desenvolvimento das habilidades orais (audição e fala) e o desenvolvimento de aspectos da escrita ligados ao conteúdo estudado. Os resultados preliminares mostraram que a teoria de análise dessas interações proposta por esses autores tem muito a oferecer a nós, professores, na tarefa de ajudar os alunos no esclarecimento de como as interações entre os modos semióticos se constituem. Percebemos também que o potencial de contribuição para o aprendizado que pode ser oferecido por cada modo precisa ser mais bem compreendido para que possamos facilitar o processo de letramento de nossos aprendizes em composições multimodais.

**PALAVRAS-CHAVE:** Interação imagem-texto; Relações lógico-semânticas.

**ABSTRACT:** This study analyzes multimodal instructional compositions from the specific context of distance learning in an attempt to shed light on how the processes of interaction between modes of visual expression and writing are outlined in this context. We applied the perspective of the relations of status and logical semantics developed by Martinec and Salway, as well as the semiotic interactions developed by Roland Barthes. A class of a Language and Literature teaching degree course, distance mode, was chosen as subject of analysis. This class mainly aims to develop oral skills (listening and speaking) and writing aspects related to the studied content. Preliminary results showed

---

\* [scarvalho50@gmail.com](mailto:scarvalho50@gmail.com)

that the theory of analysis of these interactions proposed by those authors has much to offer to us, teachers, in the task of helping students to understand how the interactions between semiotic modes are shaped. We also see that the potential of contribution to learning that may be offered by each mode needs to be better understood, so that we can facilitate the process of literacy through the use of multimodal compositions.

KEYWORDS: Image-text interaction; Logical semantics relations.

## 1 Introdução

O desenvolvimento das tecnologias de informação (TI) tem ampliado notadamente as formas de comunicação da sociedade contemporânea, e um dos principais efeitos desse crescente uso do meio digital na comunicação é que esta tem se tornado cada vez mais multimodal, ou seja, as pessoas têm se expressado utilizando mais de um *modo*<sup>1</sup> *semiótico*<sup>2</sup>. No contexto educacional não é diferente. As TI têm proporcionado um rápido desenvolvimento da educação a distância que, por sua vez, tem modificado os conteúdos das aulas tradicionais e presenciais. Esses conteúdos, antes majoritariamente instanciados em livros impressos, agora são transformados em disciplinas virtuais compostas, na maioria das vezes, por composições multimodais instrucionais<sup>3</sup> (CMI), que chegam aos alunos por meio de ambientes virtuais de aprendizagem (AVA). Esses ambientes, de uma maneira geral, são

---

<sup>1</sup> Para esse trabalho consideramos modo semiótico de forma similar a Kress (2010, p. 79) quando o autor define modo como “um recurso formado social e culturalmente para a construção de significados. A imagem, a escrita, o layout, a música, o gesto, a fala, o filme, a trilha sonora e objetos em 3D são exemplos de modos usados na representação e na comunicação.” No original: “mode is a socially shaped and culturally given semiotic resource for making meaning. Image, writing, layout, music, gesture, speech, moving image, soundtrack and 3D objects are examples of mode used in representation and communication.”

<sup>2</sup> Todas as traduções contidas neste trabalho são de nossa autoria. O texto original será apresentado em nota de rodapé.

<sup>3</sup> Entendemos composição multimodal como texto digital ou impresso que expresse sentido através de mais de um modo semiótico (ANSTEY; BULL, 2006). Resolvemos acrescentar ao termo *composição multimodal* o termo *instrucional*, pois trataremos das composições multimodais que são utilizadas com o intuito de instruir ou ensinar algo. A *composição multimodal instrucional* é aqui definida, portanto, como a apresentação de material didático que utiliza palavras escritas e imagens estáticas com a intenção de fomentar o aprendizado.

constituídos por páginas da web que reúnem diversos modos semióticos, tais como a escrita, o layout, a imagem e o som. Esses modos interagem entre si para construir os significados pretendidos pelos professores autores das disciplinas, assim, a interação entre os modos é significativa para a construção dos sentidos. É, portanto, de fundamental importância entender as relações entre a linguagem verbal e os outros modos não linguísticos de comunicação, pois em textos multimodais complexos as fronteiras entre os modos esvanecem e eles se fundem em novas configurações.

Para o propósito deste estudo foi considerado texto, conforme a visão de Halliday e Hasan (1989), toda linguagem que for funcional, ou seja, toda linguagem que estiver desempenhando uma função em algum contexto. “O texto deve ser considerado uma unidade semântica: não de forma, mas de significado” (HALLIDAY; HASAN, 1976, p. 2)<sup>4</sup>. Também consideramos a visão dos autores de que coesão é um conceito semântico, “que se refere às relações que existem dentro do texto e que o definem como texto” (HALLIDAY; HASAN, 1976, p. 4)<sup>5</sup>. Embora os autores estivessem se referindo à linguagem verbal, eles reconhecem que sua visão de texto permite ser estendida a outros modos de construção de significados. O texto pode ser “falado ou escrito ou, na verdade, qualquer outro meio de expressão que possamos pensar” (HALLIDAY; HASAN, 1989, p. 10)<sup>6</sup>. Consideramos texto visual ou imagem tudo aquilo que não for texto verbal e que estiver desempenhando uma função dentro de um contexto, como desenhos, fotos, figuras, diagramas, gráficos, entre outros. As nossas unidades de análise, as composições multimodais instrucionais, são, portanto, textos compostos por imagens estáticas e palavras escritas que formam unidades de sentido coesas e contextualizadas.

Nos AVA são utilizadas essas composições multimodais instrucionais de forma ainda pouco investigada. O ensino através do uso de CMI apresenta, portanto, um espaço de investigação convidativo a uma exploração mais profunda de como os modos semióticos visual e verbal interagem na construção de significados. Chegamos então à nossa questão de pesquisa:

---

<sup>4</sup> “A text is best regarded as a semantic unit: not of form but of meaning.”

<sup>5</sup> “The concept of cohesion is a semantic one; it refers to relations of meaning that exist within the text and that define it as a text.”

<sup>6</sup> “It (a text) may be either spoken or written, or indeed any other medium of expression that we like to think of”

como se dão as relações de construção de significado entre os modos visual e verbal em composições multimodais instrucionais? Ou seja, como os modos visual e verbal trabalham em conjunto para atingir um objetivo específico de ensino?

Neste artigo estamos especialmente interessadas em explorar as relações entre o texto escrito e as imagens estáticas no contexto de ensino de língua inglesa a distância. Partimos do pressuposto de que, se os modos são elaborados e configurados em contextos histórico, social e cultural específicos para desempenhar papéis em um dado momento, e a interação entre os modos faz parte do processo de composição do significado, seria então pertinente investigar esse contexto na tentativa de explorar alguns tipos de relações que podem ser encontradas em Composições Multimodais Instrucionais que utilizem o texto e a imagem em interação para construir significados.

Nosso objetivo é, portanto, descrever as formas como as relações entre os modos visual (imagem estática) e verbal (texto escrito) interagem na construção dos significados dos conteúdos apresentados em uma disciplina on-line de um curso de licenciatura em Letras (modalidade a distância). A intenção é tornar explícito como diferentes modos semióticos trabalham juntos para criar significados dentro das construções multimodais utilizadas pela professora-autora da disciplina, na tentativa de entender suas escolhas, e apresentar uma descrição desse entendimento à luz de um aporte teórico que apresentaremos a seguir.

## **2 A multimodalidade e as interações imagem-texto**

A multimodalidade, como descrita por Kress e Leeuwen (2001, p. 20), é “o uso de diversos recursos semióticos no desenho de um evento ou produto comunicativo”<sup>7</sup>. Entendemos recursos semióticos de forma similar a Leeuwen (2005, p. 285), quando o autor afirma que:

Recursos semióticos são ações, materiais e artefatos que usamos para fins comunicativos, junto com as formas pelas quais essas ações, materiais e artefatos podem ser organizados. Eles podem ser produzidos fisiologicamente – por exemplo, com o nosso aparelho vocal, os músculos que usamos para fazer expressões faciais e gestos – ou tecnologicamente – por exemplo, com caneta e tinta, ou hardware e *software*. Recursos semióticos têm um significado potencial

---

<sup>7</sup> “The use of several semiotic modes in the design of a semiotic product or event”.

que se baseia em seus usos anteriores, e em um conjunto de *affordances*.<sup>8</sup> (VAN LEEUWEN, 2005, p 285).

Modo, por sua vez, segundo Kress (2011 p. 54), “é um recurso social e culturalmente construído para produzir significado. A imagem, a escrita, o layout, a música, os gestos, a fala são exemplos de modos utilizados na representação e na comunicação”<sup>9</sup>. Ainda segundo o autor, os modos oferecem diferentes potenciais para a produção de significados, o que afeta a sua escolha no momento da comunicação. Dependendo de nossos interesses e propósitos, escolhemos os modos e/ou os combinamos para atingir nossos objetivos comunicativos em um fenômeno multimodal. Todos os modos, portanto, consistem em um conjunto de recursos semióticos. Dessa forma, “os modos e os recursos semióticos que uma pessoa usa (ou pode usar) dão forma à comunicação e ao significado”<sup>10</sup>. (JEWITT, 2011, p. 2).

Nessa perspectiva, o código verbal é, portanto, parte de um conjunto multimodal de expressão de significados. Os significados são produzidos, recebidos, distribuídos e reinterpretados de vários modos.

Ainda de acordo com Jewitt (2011), a comunicação sempre aconteceu por mais de um modo de expressão. Contudo, essa percepção da existência das composições comunicativas multimodais tornou-se mais intensa com o desenvolvimento de tecnologias de representação visual e com a expansão das diversas formas de as pessoas se comunicarem e interagirem (síncrona e assincronamente) e abriu espaço para a inserção da imagem em suas vidas diárias. A escrita e a fala parecem não ser mais suficientes para produzir sentidos nas interações que ocorrem para além do código verbal. Os significados são produzidos socialmente no sentido de que são regidos por

---

<sup>8</sup> “Semiotic resources are the actions, materials and artefacts we use for communicative purposes, whether produced physiologically – for example, with our vocal apparatus, the muscles we use to make facial expressions and gestures – or technologically – for example, with pen and ink, or computer hardware and software – together with the ways in which these resources can be organized. Semiotic resources have a meaning potential, based on their past uses, and a set of affordances”.

<sup>9</sup> “Mode is a socially shaped and culturally given resource for making meaning. Image, writing, layout, music, gesture, speech, moving image, soundtrack are examples of modes used in representation and communication.”

<sup>10</sup> (...) the modes and semiotic resources a person chooses (or are permitted) to use shape communication and meaning.

regras em vigor no momento de sua produção, em um contexto específico, e, ainda, influenciados pelos interesses de quem os produz.

As afirmações do parágrafo anterior são corroboradas por alguns pressupostos apresentados pela pesquisadora Jewitt. O primeiro deles é que “a linguagem é parte de um conjunto multimodal”<sup>11</sup> e que a “representação e a comunicação sempre recorrem a uma multiplicidade de modos, com todos os modos tendo o potencial de contribuir igualmente para a construção do significado” (JEWITT, 2011, p. 14)<sup>12</sup>. O segundo é que “cada modo, em um conjunto multimodal, é entendido como realizador de um trabalho comunicativo diferente” (JEWITT, 2011, p. 14)<sup>13</sup>, ou seja, diferentes aspectos do significado são expressos por cada modo. O terceiro é que as “pessoas orquestram os significados através de sua seleção e configuração dos modos” (JEWITT, 2011, p. 14)<sup>14</sup>, assim a interação entre os modos é parte do processo de geração do significado.

Jewitt (2011, p. 25) afirma ainda que “cada modo (como se realiza em um contexto social particular) possui uma lógica específica e fornece potenciais comunicativos e representacionais diferentes, característica conhecida como lógica modal”<sup>15</sup>. Segundo ela, a imagem e outros modos não linguísticos assumem papéis específicos em um contexto específico em um dado momento, ou seja, a linguagem toma forma pelo uso que as pessoas fazem dela. Segundo Leeuwen (2011, p. 673), “as formas de comunicação visual são radicalmente diferentes das formas de comunicação da linguagem verbal, mas isso não significa que não possam exercer as mesmas funções comunicativas e, até certo ponto, expressar os mesmos significados”<sup>16</sup>.

---

<sup>11</sup> “Language is part of a multimodal ensemble”

<sup>12</sup> “(...) representation and communication always draw in a multiplicity of modes, all of which have the potential to contribute equally to meaning”.

<sup>13</sup> (...) each mode in a multimodal ensemble is understood as realizing different communicative work.

<sup>14</sup> “(...) people orchestrate meaning through their selection and configuration of modes.”

<sup>15</sup> “Each mode (as it is realized in a particular social context) possesses a specific logic and provides different communicational and representational potentials, referred to as modal logic.”

<sup>16</sup> “The forms of visual communication are radically different from those of language. But that does not mean that both cannot fulfill the same communicative functions, and, at some level, express the same meaning”.

Essas afirmações de Jewitt e Leeuwen nos remetem aos contextos sociais de produção dos significados e às funções sociais da linguagem idealizadas por Halliday (1978, 1985). Segundo Halliday, existem três funções presentes nos usos que fazemos da linguagem: a função ideacional, a função interpessoal e a função textual. Pela função ideacional, as pessoas constroem representações sobre o mundo, sobre suas ideias e experiências. A função interpessoal expressa as relações sociais entre os envolvidos no evento comunicativo. E pela função textual é expressa a organização dos elementos que compõem os significados.

Embora originalmente idealizadas para se aplicarem ao uso da linguagem verbal, as metafunções de Halliday foram observadas pelos autores da Gramática do Design Visual, Kress e Leeuwen (1996, 2006), como presentes em todos os modos de expressão de significado, não somente o verbal. Kress e Leeuwen (1996, 2006) fazem um paralelo dessas metafunções da linguagem verbal para as imagens e as renomeiam como função representacional, interativa e composicional, respectivamente. Dessa forma, ao construir significados por meio das composições multimodais instrucionais, estamos realizando de forma similar as três metafunções da linguagem: estamos representando o mundo, as ideias e as experiências que temos do mundo (função representacional); estamos estabelecendo relações sociais com os leitores/observadores e com a própria imagem/texto (função interacional); e estamos organizamos os elementos dentro da CMI de forma intencional (função composicional). O pressuposto é que, como a gramática da língua verbal, a gramática da comunicação visual pode ser descrita como um sistema de escolhas semântico-funcionais. Contudo, como destaca Leeuwen (2011, p. 673) “nem todos os significados que podem ser realizados verbalmente podem ser realizados visualmente e vice-versa, e que, mesmo quando os dois modos podem realizar os mesmos significados, cada um irá adicionar seu tom próprio e sua ressonância”<sup>17</sup>.

Roland Barthes (1977) elaborou alguns dos primeiros estudos sobre as relações entre texto e imagem. Segundo ele, as imagens relacionavam-se com os textos de três formas distintas. A primeira forma ele chamou de “ancoragem”. Nesse tipo de relação, o texto conota e direciona a leitura, servindo de apoio à imagem. Segundo Barthes (1977, p. 39),

---

<sup>17</sup> “(...) not all meanings that can be realized verbally can be also realized visually, and vice-versa, and that, even when two modes can realize the same meaning, each will add its own overtone and resonances.”

todas as imagens são polissêmicas, pois em uma cadeia flutuante de significados, o leitor pode escolher uns e ignorar outros. [...] Assim, cada sociedade desenvolve várias técnicas para fixar a cadeia flutuante de significados, de modo a combater o terror dos signos incertos; a mensagem linguística é uma destas técnicas.<sup>18</sup>

O autor alerta que a ancoragem pode ser ideológica e que, na verdade, essa é a sua principal função. O texto direciona o leitor pelos significados da imagem, guiando-o por essa cadeia de significados para que ele escolha alguns e evite outros. Nos casos de ancoragem, a linguagem tem claramente a função de elucidação, mas uma elucidação seletiva, aplicada não à totalidade da imagem, mas apenas a algumas partes.

A segunda forma foi denominada de “ilustração” e nela o processo é inverso: é a imagem que dá suporte à leitura do texto. Neste caso, o texto é o principal elemento e a imagem o interpreta em um contexto específico.

A última relação foi chamada de *relay*. Nesse caso, um modo apoia o outro, ou seja, eles são complementares, como exemplo, nas revistas em quadrinhos, e os modos dependem um do outro para construir o significado. As palavras, assim como as imagens, funcionam como fragmentos de um sintagma maior, digamos uma anedota ou uma história.

Martinec e Salway (2005), inspirados nos trabalhos de Barthes (1977) e nas metafunções da linguagem visual (KRESS; LEEUWEN, 1996, 2006), desenvolveram um sistema de relações semânticas entre texto e imagem que combinavam dois tipos de relações: as relações de status e as relações lógico-semânticas. Cada combinação imagem-texto foi descrita por uma seleção de escolhas semântico-funcionais. O foco dos autores era apresentar o sistema de relações semânticas entre os modos visual e verbal e especificar as suas possíveis realizações.

## 2.1 Relações de *status*

Segundo Martinec e Salway (2005), assim como as relações entre as frases em sentenças formadas por coordenação e subordinação, as relações

---

<sup>18</sup> “All images are polysemous; they imply, underlying their signifiers, a ‘floating chain’ of signifieds, the reader able to choose some and ignore others. (...) Hence, in every society various techniques are developed intended to fix the floating chain of signifieds in such a way as to counter the terror of uncertain signs; the linguistic message is one of these techniques.



entre a imagem e o texto escrito também podem ser consideradas como *iguais* ou *desiguais*. As relações de igualdade entre texto e imagem podem ser expressas de duas formas: eles podem ser *independentes* ou *complementares*.

O texto e a imagem são considerados independentes quando são apresentados em uma mesma composição multimodal, mas um não modifica o outro. Quando o texto e a imagem são independentes, eles não se combinam para formar um sintagma<sup>19</sup> composicional maior. A informação que eles fornecem existe em paralelo – cada um forma seus próprios processos. Por outro lado, quando a relação de status entre uma imagem e um texto é considerada complementar, eles se combinam para formar parte de um sintagma maior. Isso significa que desempenham o papel de participantes em um tipo de processo maior que eles mesmos, mas um não modifica o outro. Tanto nas relações de status independente quanto complementar uma imagem inteira se relaciona a um texto inteiro.

Em contraste, quando o status é de desigualdade, ou seja, um elemento é subordinado a outro, a imagem está relacionada com apenas uma parte do texto, ou o texto com apenas parte da imagem. Os processos através dos quais se dão essas relações são estabelecidos pelo que Martinec (1998) chama de “coesão componencial”. Esse termo trata-se na verdade da mesma coesão lexical, renomeada como coesão composicional e aplicada às composições multimodais. Segundo Martinec (1998, p. 162), “embora os princípios gerais desse tipo de coesão funcionem da mesma forma na ação e na linguagem verbal, as relações que se obtêm não serão entre itens lexicais, mas entre componentes.” (MARTINEC, 1998, p 162).<sup>20</sup> Entendemos por essa afirmação que os mesmos mecanismos que envolvem a repetição de uma unidade lexical ou a sua substituição por outras unidades lexicais que com ela mantêm relações semânticas de natureza hierárquica (*hiponímia*, *hiperonímia*), de inclusão (*holonímia*, *meronímia*) ou não hierárquica (*sinonímia*, *antonímia*), também podem ocorrer entre componentes de uma CMI. Quando um modo se relaciona com apenas parte do outro modo, os mecanismos pelos quais se dão essas relações entre as partes de um modo e de outro são regidos por esses mecanismos, chamados pelo autor de coesão componencial.

---

<sup>19</sup> Entendemos sintagma como um conjunto de elementos que compõem uma unidade significativa, mantendo entre si relações de dependência e de ordem.

<sup>20</sup> “Although the general principles of this kind of cohesion work the same in action as they do in language, the relations obtain between what will be called components rather than lexical items.”

Quando um texto é subordinado a uma imagem, o texto pode estar relacionado com apenas uma parte dela, mas esta não é a única possibilidade. A indicação mais confiável de subordinação do texto é a presença de dispositivos implícitos que precisam ser decodificados por referência a uma imagem.

Martínez e Salway (2005, p. 351) resumem o status das relações texto-imagem no seguinte diagrama:

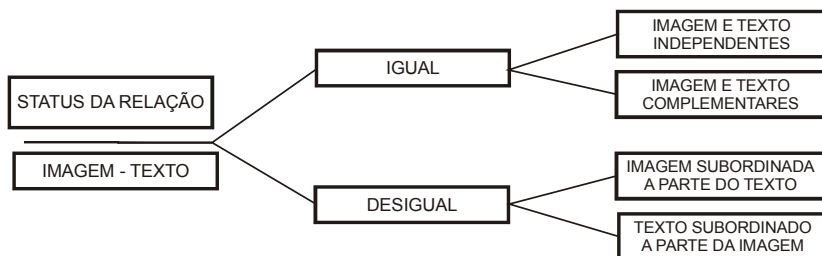


FIGURA 1 – Resumos das relações de *status* imagem-texto

Fonte: Traduzido e adaptado de Martínez e Salway (2005, p. 351)

Em suma, as relações entre os elementos se dão entre o texto inteiro e a imagem inteira, ou entre a imagem inteira e parte do texto, ou ainda entre o texto inteiro e parte da imagem. A maior unidade textual trabalhada pelos autores é um parágrafo, mas eles alertam que é possível que uma imagem se relacione com fragmentos textuais maiores. Por exemplo, se todas as sentenças independentes de um parágrafo estão relacionadas com a imagem pode-se dizer que a imagem se relaciona com o texto inteiro. Se algumas sentenças não se relacionam com a imagem, pode-se dizer que a imagem se relaciona com parte do texto. Essas relações entre partes dos modos se dão pela coesão componencial.

## 2.2 Relações lógico-semânticas

Para a construção de seu segundo subsistema de relações, Martínez e Salway (2005) se fundamentaram em Halliday (1985). Eles utilizaram os dois tipos de relações lógico-semânticas de expansão e de projeção, utilizados por Halliday na análise linguística, para dar forma ao sistema de relações lógico-semânticas entre um texto e uma imagem.

Os autores apontam que a principal diferença entre a expansão e a projeção é que a expansão lida com as relações entre eventos representados em experiências não linguísticas, enquanto a projeção lida com eventos que já foram apresentados. A projeção é, portanto, útil nos casos em que o conteúdo já foi apresentado pelo texto ou pela imagem e é reapresentado no outro modo. Os exemplos mais comuns dessa relação são os diagramas e as tirinhas em quadrinhos.

### **2.2.1 Expansão**

As relações lógico-semânticas de expansão se subdividem em três tipos: elaboração, extensão e aprimoramento.

Os autores identificaram dois tipos de elaboração: a exposição e a exemplificação. Na exposição, a imagem e o texto possuem o mesmo nível de generalidade, enquanto na exemplificação, o nível de generalidade é diferente. Na exemplificação pode-se encontrar o texto mais geral que a imagem ou o inverso.

A extensão é a relação entre texto e imagem, em que um dos dois modos acrescenta uma nova informação, ou seja, informação que vá além do que está expresso no outro modo.

A relação de aprimoramento se dá quando um modo qualifica o outro circunstancialmente. Foram identificadas pelos autores as circunstâncias de tempo, lugar e razão ou propósito.

### **2.2.2 Projeção**

Como já mencionado, a projeção lida com eventos que já foram representados em um modo e são rerepresentados no outro, como exemplo, nos diagramas, em que aspectos importantes do texto são escolhidos e reexpressos visualmente. Outro exemplo são as tirinhas em quadrinhos, em que muitas vezes o que é expresso na imagem é reexpresso nos diálogos. A projeção é subclassificada em dois tipos: locução e ideia. Uma locução é uma projeção do texto, geralmente por um processo verbal. Uma ideia é a projeção de sentido de um pensamento, na maioria das vezes, um processo mental. Nas tirinhas em quadrinhos a diferença é facilmente perceptível se observarmos que as locuções aparecem em balões de fala e as ideias em balões de pensamento.

A figura a seguir apresenta o quadro geral do sistema de relações de status e relações lógico-semânticas das interações entre os textos e as imagens proposto pelos autores.

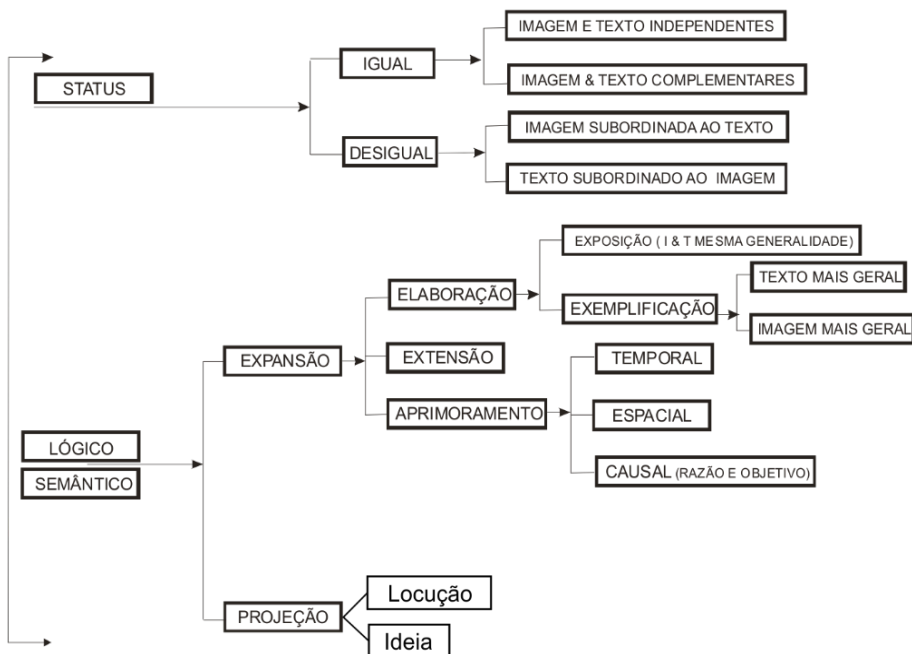


FIGURA 2 – Resumos das relações de *status* e lógico-semânticas entre imagem-

Fonte: traduzido e adaptado de Martinec e Salway (2005, p. 360)

Com esse resumo das interações imagem-texto de Martinec e Salway (2005) finalizamos a seção do aporte teórico deste trabalho. A seguir, apresentamos detalhes sobre a metodologia da pesquisa, bem como sobre a análise das composições observadas.

### 3 Natureza da Pesquisa e Metodologia

A pesquisa que propomos é de natureza qualitativa à medida que procura entender o seu objeto de estudo – as composições multimodais instrucionais – dentro de seu contexto específico e, a partir daí, situar a interpretação delas. Entendemos este estudo como qualitativo também por reunir algumas das características que Godoy (1995) apresenta como típicas da pesquisa qualitativa:

- o ambiente como fonte direta de dados e o pesquisador como
- a. instrumento fundamental;
  - b. o caráter descritivo;
  - c. enfoque indutivo.

Os dados do estudo foram levantados da disciplina Tópicos em Língua Inglesa: Práticas Oraís, que pertence à grade curricular de um curso de licenciatura em Letras: Língua Inglesa e suas Literaturas (modalidade a distância), ofertado por uma Universidade Federal do Brasil<sup>21</sup>. A disciplina é composta por quatro *lições*, hospedadas em um ambiente virtual de aprendizagem. Cada lição é composta por quatro páginas da web, denominadas introdução (*introduction*), pré-tarefa (*pre-task*), tarefa (*task*) e pós-tarefa (*post-task*). Na introdução, a professora situa o aprendiz acerca do que será estudado e apresenta em linhas gerais os conteúdos da lição. Na pré-tarefa, apresenta atividades para ativação do conhecimento prévio dos alunos acerca do tema e aponta os conteúdos que estes precisam saber para a execução da tarefa. Em seguida, propõe uma série de atividades para que esses conteúdos sejam desenvolvidos. Na página destinada à tarefa, a professora faz a proposição da tarefa com todos os elementos que devem constituí-la, orientando os alunos para desenvolvê-la. Na página destinada à pós-tarefa, os alunos são solicitados a refletirem sobre o que foi feito durante a tarefa e como poderiam melhorá-la.

Para a realização da pesquisa, primeiramente fizemos um levantamento das escolhas feitas pela professora e elaboradora da disciplina em relação aos recursos semióticos utilizados por ela no material didático. Depois, isolamos as composições multimodais instrucionais constituídas por imagens estáticas e todo o conteúdo verbal a elas relacionado presentes no material da disciplina. Por último, procuramos investigar como se processava a distribuição do significado entre as imagens estáticas e os textos verbais com base nas relações de status e lógico-semânticas apresentadas por Martinec e Salway (2005), que fornecem as bases teóricas para a análise das composições multimodais instrucionais.

---

<sup>21</sup> Não apresentaremos a universidade para preservar a identidade da professora participante da pesquisa.

#### 4 Análise das construções de sentido das interações texto-imagem

O conteúdo da disciplina analisada está organizado em quatro aulas, cada uma sendo prevista para ser estudada em três semanas. Todo o conteúdo da disciplina está hospedado em uma plataforma de ensino denominada Solar que foi especialmente desenvolvida para o ensino a distância. Foi feito um inventário dos modos de expressão de significados nos quais semioses variadas foram encontradas. Os diferentes tipos de CMI encontrados são apresentados na tabela a seguir:<sup>22</sup>

TABELA 1 – Modos encontrados em interação com outros modos

AULA	MODOS UTILIZADOS EM CONJUNTO				
AULA 1	6 – IMAGEM/TEXTO	---	1 TEXTO/SOM	---	---
AULA 2	5 – IMAGEM/TEXTO	3 – IMAGEM/SOM	---	---	---
AULA 3	1 – IMAGEM/TEXTO	---	---	--	1 – VÍDEO/IMAGEM/TEXTO
AULA 4	1 – IMAGEM/TEXTO	1 – IMAGEM/SOM	---	2 – VÍDEO/ TEXTO	---
TOTAL	<b>13 – IMAGEM/TEXTO</b>	<b>4 – IMAGEM - SOM</b>	<b>1 – TEXTO/SOM</b>	<b>2 – vídeo/ texto</b>	<b>1 – VÍDEO/IMAGEM/TEXTO</b>

Fonte: Elaborada pela autora.

Percebe-se na Tabela 1 que existe uma gama de composições multimodais que constroem sentido por meio de modos semióticos diversos no material didático on-line analisado. Foram encontradas interações entre diversos modos, tais como imagem/texto, imagem/som, texto/som, vídeo/texto, vídeo/imagem/texto. Observa-se ainda que as semioses mais comuns são entre imagens estáticas e texto verbal. Foram encontradas treze realizações desse tipo de CMI. Por restrições de espaço, limitamos nossa análise a essas interações uma vez que elas são as mais recorrentes no material. Foram encontrados quatro tipos diferentes de interação entre as treze composições multimodais instrucionais compostas por imagem estática e texto escrito. Analisamos uma CMI para cada tipo de interação encontrada. Primeiramente, discutimos os aspectos concernentes às relações de status e lógico-semânticas definidas por Martinec e Salway (2005). Em seguida, buscamos por relações entre as CMI e os objetivos didáticos de cada tarefa em que a CMI está situada. Achamos pertinente, antes da análise

<sup>22</sup> Utilizamos o termo vídeo quanto há imagens em movimento e o termo imagem quando esta for estática.

das composições, fazemos uma breve descrição do contexto em que cada CMI foi inserida, procurando revelar o provável objetivo da professora na utilização específica de cada composição.

#### **4.1. Análise das composições multimodais instrucionais - Texto Verbal/Texto Visual**

##### **CMI 1 – Relações de status complementares com relações lógico-semânticas sobrepostas (expansão e extensão).**

Contexto A CMI 1 foi utilizada em uma atividade de pré-tarefa (Figura 3) cujo propósito era fazer o aluno refletir a respeito das qualidades que são características de um bom professor, para, posteriormente, escrever um texto sobre esse assunto.



FIGURA 3 – Pré-tarefa (CMI 1)

Fonte: Plataforma Solar

Na CMI 1 observa-se que o status da relação entre o texto e a imagem é de igualdade e, portanto, eles são complementares. A imagem inteira se relaciona com o texto todo. Ambos estão interagindo para a composição de um sintagma maior: o conjunto dos dois. Os significados ocorrem por combinação entre a imagem e o texto, portanto, qualquer um dos elementos utilizados em isolamento não construiria os mesmos sentidos que a composição dos dois constrói.

A imagem apresenta uma coruja usando um capelo e uma ponteira. Esses elementos da imagem adicionam alguns sentidos, diferentes daqueles expressos pelo texto verbal. Como cada modo é construído em contextos histórico, social e cultural específicos, para construirmos os sentidos da CMI 1, devemos considerar que a figura da coruja em nossa sociedade está associada ao magistério, pois ao pé da letra, o pedagogo é aquele que conduz a criança, e as corujas são reconhecidas pelo cuidado extremo que têm com seus filhotes. As corujas são associadas ao magistério também devido à sua inteligência, astúcia e sensibilidade, que são algumas das características que identificam um bom professor. A imagem da coruja funciona, portanto, como símbolo que reúne as qualidades do bom professor. O capelo, chapéu usado nas solenidades de formatura, e a ponteira, utilizada para indicar informações importantes nos quadros-negros, são elementos que reforçam essa associação.

O texto escrito, por sua vez, é apresentado por sentenças que especificam algumas qualidades associadas ao bom professor. As sentenças apontam para significados mais singulares que aqueles construídos pela imagem.

Quanto à relação lógico-semântica, acreditamos que a composição não se encaixe em nenhuma das relações apresentadas por Martinec e Salway (2005), mas em uma combinação de duas. Entendemos que há mais de um processo de relação co-ocorrendo entre a imagem e o texto. Existe uma relação de *expansão*, na qual o nível de generalidade entre texto e imagem é diferente. A imagem é mais geral que o texto, pois carrega vários sentidos culturalmente construídos. O texto escrito composto por algumas sentenças inseridas em balões de pensamento especifica ou exemplifica alguns desses sentidos. Existe também uma relação de projeção de sentido ou ideia. A imagem e o texto se relacionam novamente. A coruja (símbolo do bom professor), mediante os balões de pensamento, convida a um processo mental. Convida à reflexão sobre as qualidades de um bom professor, que é objetivo da atividade proposta pela professora da disciplina.



Percebemos que na CMI 1 há relações sobrepostas entre texto e imagem – relações de expansão e projeção. Os sentidos construídos pelos dois modos nessas relações contribuem para o objetivo da atividade que é acionar o conhecimento prévio acerca do que é um bom professor e convidar à reflexão sobre esse tema.

## CMI 2 – Imagem e texto em relações desiguais (projeções de significados)

Contexto A Figura 4 foi utilizada pela professora em uma atividade de pré-tarefa com o intuito de que os alunos conseguissem fazer a associação entre a palavra *advice* (“conselho”) e as palavras ou expressões que poderiam ser consideradas sinônimas. O objetivo dessa atividade é ensinar uma função: como dar conselhos na língua inglesa.



FIGURA 4 – Como dar conselhos em inglês (CMI 2)

Fonte: Plataforma Solar

Na CMI 2, o status relativo da imagem com o texto é de desigualdade – imagem subordinada ao texto, ou seja, a imagem se relaciona apenas com partes do texto. A imagem não se relaciona com todas as sentenças independentes. Por exemplo, a imagem se relaciona com a sentença “*Você deveria sorrir*” (*You should smile*), pois a garota aparece sorrindo. Relaciona-se também com a sentença “*Talvez você não devesse parecer tão ansiosa*” (*Maybe you could try not to be too anxious*), na imagem a garota parece não estar ansiosa. Há ainda relação com a sentença “*Você deve se vestir bem*” e a menina parece bem vestida (*You should dress nicely*). E, por último, a imagem se relaciona com parte

da sentença “*Você deveria ser legal e amigável com ele*” (*You should be nice and friendly with him*). Contudo, não há na imagem sinal de relações com as sentenças: “*Você deveria tentar não ser tão direta*” (*but maybe you could try not to be so direct*); “*Talvez você devesse convidá-lo para ir ao cinema*” (*Maybe you could invite him to see a movie*); “*Pode ser uma boa ideia falar sobre você e dar a chance a ele de falar também*” (*It might be a good idea to talk about yourself and also give him the chance to talk*).

Na imagem do balão os dois aparecem juntos, sorrindo e de mãos dadas, evidenciando a situação idealizada pela menina. Fora do balão, os dois aparecem separados. Ela aparece sorrindo com olhar apaixonado e ele parece de braços abertos, com semblante assustado ou surpreso. Os elementos fora do balão mostram distanciamento do garoto em relação à Bernice.

Ao considerarmos o contexto sócio-histórico-cultural da CMI 2, percebemos certo reforço da imagem da mulher como uma pessoa que se tomar a iniciativa em um relacionamento precisa ter certos cuidados, “não parecer muito ansiosa”, “tentar não ser muito direta”, “vestir-se bem”. O homem pode assustar-se com mulheres que os cortejam, como no caso da imagem em que o garoto parece assustado com a expressão de amor da garota.

Ao analisarmos os sentidos construídos pela CMI 2, sob a perspectiva das relações lógico-semânticas, verificamos a existência de uma projeção de significado (ideia). Embora o balão de pensamento não esteja se relacionando especificamente com os conselhos, relaciona-se com o texto que vem logo abaixo da imagem: *Bernice está apaixonada por Paul* (*Bernice is in love with Paul*) – na imagem do balão de pensamento, os dois estão juntos e felizes, uma projeção de sentido, um processo mental. A imagem dá suporte à leitura do texto. O texto é o principal elemento e a imagem o interpreta em alguns elementos específicos. Observa-se também que, ao contrário da CMI 1, em que texto e imagem alinhavam-se para atingir o objetivo da atividade proposta, que era a ativação do conhecimento prévio e a geração de reflexão acerca do tema, na CMI 2 apenas o texto verbal se relaciona diretamente com o objetivo didático. Talvez isso possa ser explicado pelo fato de o objetivo da atividade ser basicamente linguístico, ou seja, o reconhecimento das diferentes escolhas lexicais que podemos fazer para cumprir a função de dar um conselho a alguém na língua inglesa.

### CMI 3 - Imagem e texto em relações de igualdade (complementariedade).

Contexto A professora monta uma história para os alunos lerem e interagirem com a história mediante respostas que podem ser confirmadas ou não por meio das imagens. A atividade se baseia no contexto de um homem apaixonado por uma mulher e que quer conquistá-la. A instrução para o aluno é que leia o texto e em alguns pontos da história escolha, entre três opções, o conselho que ele achar que o homem apaixonado deveria seguir. Em seguida, o aluno deve clicar em um botão e a imagem do que de fato o homem apaixonado fez aparece ao lado. O objetivo da atividade é que o aluno interaja com a CMI e tente identificar quais opções o personagem escolheu e, assim, de forma indutiva, entre em contato com as formas de dar um conselho na língua inglesa, conforme mostra a Figura 5.

John é um homem que está perdidamente apaixonado por Carla. Ele quer se aproximar dela, mas ele não tem ideia de como alcançar seu objetivo. Ele tenta diferentes estratégias para atrair o seu amor. Você vai acompanhar John nessa jornada de amor, sinta-se à vontade para dar alguns conselhos. Dê seus conselhos a John escolhendo entre as opções a, b ou c. Em seguida, clique para ver se John seguiu seu conselho. Não fique desapontado se ele não seguir o seu conselho, continue tentando ajudá-lo até o final de sua jornada de amor.

John está apaixonado por Carla. O que ele deve fazer?

1. Que tal convidá-la para ver um filme?
2. Que tal visitá-la sem avisar?
3. Que tal implorar por seu amor?

Clique aqui para ver se John seguiu seu conselho→



John implorou pelo amor de Carla, mas ela nem ligou. O que ele deve fazer agora?

1. Que tal dar a ela um anel?
2. Que tal ir para casa e pensar um pouco mais sobre isso?
3. Que tal dar flores a ela?

Imagem 2

Clique aqui para ver se John seguiu seu conselho→



John deu um anel a Carla, mas ela nem ligou. O que ele deve fazer agora?

1. Talvez fosse uma boa ideia desistir.
2. Talvez fosse uma boa ideia dar a ela outro presente.
3. Talvez fosse uma boa ideia convidá-la para jantar.

Imagem 3

Clique aqui para ver se John seguiu seu conselho→



John deu a Carla um outro presente e, de novo, ela nem ligou. O que ele deve fazer agora?

1. Talvez fosse uma boa ideia esquecê-la.
2. Talvez fosse uma boa ideia esperar um pouco.
3. Talvez fosse uma boa ideia dar a ela um presente ainda maior.

Clique aqui para ver se John seguiu seu conselho→




Imagem 4

John foi a casa de Carla com uma caixa enorme nas mãos, ela nem ligou. O que ele deve fazer agora?

1. Talvez ele devesse mostrar um vestido belíssimo dentro da caixa.
2. Talvez ele devesse abrir a caixa imediatamente.
3. Talvez ele devesse dizer a ela que o presente enorme é para outra garota.

Clique aqui para ver se John seguiu seu conselho→




Imagem 5

John abriu a caixa imediatamente e deu a ela um casaco de pele, mas a Carla nem ligou. O que ele deve fazer agora?

1. Que tal visitá-la de novo?
2. Talvez fosse uma boa ideia procurar outra namorada.
3. Talvez ele devesse ir fazer terapia.

Clique aqui para ver se John seguiu seu conselho→




Imagem 6

John decidiu ir a casa dela de novo. Ele foi de de carro. Era um carro bonito. O que ele deve fazer agora?

1. Que tal dar a ela o carro bonito?
2. Talvez fosse uma boa ideia convidá-la para dar uma volta no carro.
3. Talvez devesse mostrar a sua nova namorada dentro do carro.

Clique aqui para ver se John seguiu seu conselho→




Imagem 7

Ao final da história Carla está interessada em John quando o vê com outra garota, mas agora ele tem uma namorada. O que a Carla deve fazer?

FIGURA 5 – Dando conselhos em inglês (CMI 3)

Fonte: Plataforma Solar

Apesar de termos várias imagens diferentes nessa composição, todas elas se relacionam com o texto escrito da mesma forma. A relação de status é complementar. Há uma combinação de texto e imagem para formar uma unidade de sentido maior, a própria narrativa. O texto escrito apresenta-se em fragmentos que vão se relacionando às imagens e, ao final, pode-se recompor a narrativa. À medida que o aprendiz clica nos locais indicados, imagens surgem e dão uma resposta sobre qual opção, das descritas textualmente, o personagem escolheu. Assim, imagem e texto vão construindo os sentidos da narrativa. Essa construção se dá mediante a coesão componencial, ou seja, através de mecanismo de substituição de parte do texto por uma imagem com a qual este estabelece relações semânticas de sinonímia.

O conteúdo da imagem 1 da Figura 5, por exemplo, estabelece relações de similaridade entre a imagem de John ajoelhado e Carla virando o rosto e os conteúdos expresso verbalmente pelos fragmentos de texto: “*Que tal implorar por seu amor?*”, “*John implorou pelo amor de Carla, mas ela nem ligou*”. Essas relações se repetem em todas as imagens e seus respectivos fragmentos de texto. O conteúdo da imagem 2 da Figura 5, em que John aparece oferecendo um anel à Carla e ela é retratada virando o rosto para a oferta do presente, guarda relações de sinonímia com os fragmentos de texto: “*Que tal dar a ela um anel?*”, “*John deu um anel a Carla, mas ela nem ligou*”. Na imagem 3 da Figura 5, os conteúdos expressos pela imagem se relacionam com os fragmentos de texto: “*Talvez fosse uma boa ideia dar a ela outro presente*”. “*John deu a Carla um outro presente e, de novo, ela nem ligou*”. O conteúdo da imagem 4 da Figura 5 é expresso verbalmente pelos fragmentos de texto: “*Talvez fosse uma boa ideia dar a ela um presente ainda maior*”. “*John foi à casa de Carla com uma caixa enorme na mão, ela nem ligou*”. O da imagem 5 da Figura 5 é expresso verbalmente pelos fragmentos de texto: “*Talvez devesse mostrar um vestido belíssimo dentro da caixa*”. “*John abriu a caixa imediatamente e deu a ela um casaco de pele, mas a Carla nem ligou*”. O conteúdo da imagem 6 da Figura 5 é expresso verbalmente pelos fragmentos de texto: “*Que tal visitá-la de novo?*”, “*John decidiu ir à casa dela de novo. Ele foi de carro. Era um carro bonito*”. O conteúdo da imagem 7 da Figura 5 é expresso verbalmente pelo fragmento de texto: “*Talvez devesse mostrar a sua nova namorada dentro do carro*”.

É importante destacar que, apesar de podermos observar relações de sinonímia entre os conteúdos expressos nos dois modos mediante a coesão composicional, não podemos esquecer que “mesmo quando os dois modos podem realizar o mesmo significado, cada um irá adicionar seu tom próprio e sua ressonância.” (LEEuwEN, 2011, p. 673)<sup>23</sup>. Dessa forma, se substituíssemos as imagens pela repetição dos fragmentos de texto, teríamos um resultado de sentido diferente do obtido com a relação texto-imagem. Destacamos ainda que a ordem de apresentação dos conteúdos pode também contribuir para o resultado de sentido. Na CMI originalmente criada pela professora, as imagens permanecem escondidas e só são reveladas após a leitura dos fragmentos de texto aos quais se relacionam, contribuindo para uma relação leitor-texto diferenciada da obtida em nossa descrição. Nessa

---

<sup>23</sup> “(...) even when two modes can realize the same meaning, each will add its own overtone and resonances.” (VAN LEEuwEN, 2011, p. 673).

descrição não temos como, pelo texto escrito, replicar o efeito de esconder a imagem até o momento oportuno para causar o efeito surpresa desejado pela professora.

Quanto às relações lógico-semânticas entre texto e imagem da CMI 3, esta é de extensão. Isso significa que cada imagem adiciona, em dado momento, uma nova informação (o conselho escolhido pelo personagem) e essa informação está relacionada a ambos os modos.

Nota-se na CMI 3 uma relação direta entre a interação dos modos e o objetivo da professora – que os alunos fossem expostos às diferentes estruturas utilizadas pelos falantes da língua inglesa para dar conselhos a alguém.

#### **CMI 4 – Relação de desigualdade (ilustração)**

Contexto A professora esclarece que na Unidade 4 os alunos irão aprender a contar uma história sobre um colega de classe e que, para se prepararem para essa tarefa, eles devem fazer algumas atividades. No caso específico da CMI 4, os alunos devem ler a *história A* e depois observar a estrutura da *história B*<sup>24</sup>. (Figura 6)



A teacher is telling a story to her students. Read the story (A).

Then, read the narrative structure of stories (B).

Once upon a time there was a boy named David who wanted to be God. He wanted to make many changes in the world. One day he met God and he told God to make cashews as big as pumpkins. So, God denied his request. God said that all cashews would continue to be small and all pumpkins would continue to be big. David was very mad at God. He decided to take a nap under a cashew tree. Suddenly, he woke-up with three cashews falling over his head. He immediately thought that if God had accepted his request, he would be dead by now. So, David gave up being God. Perhaps David realized that God always knows best!

FIGURA 6 – Tempos Narrativos (CMI 4)

Fonte: Plataforma Solar

Nessa composição, a relação de status entre a imagem e o texto é de desigualdade, ou seja, a imagem relaciona-se apenas com uma parte do texto. A imagem é subordinada ao texto no sentido de que se relaciona apenas com

<sup>24</sup> “In this unit you will be asked to tell a story about a classmate. In order to get ready for the task, you will undergo some pre-task activities.”

o fragmento no qual a professora está contando uma história para seus alunos (*A teacher is telling a story to her students*), numa relação de sinonímia com uma única sentença. A imagem reapresenta o conteúdo expresso pelo fragmento de texto ao qual se relaciona. Não foi percebida nenhuma outra relação entre o texto e a imagem a não ser a relação que Barthes (1967) denominou de *ilustração*. A imagem duplica certas informações dadas no texto num fenômeno de redundância. Não foi detectado no uso dessa CMI nenhuma relação direta com o objetivo da professora de que os alunos percebam a estrutura das narrativas. O texto constitui-se, portanto, no principal elemento da CMI, a imagem apenas ilustra um determinado aspecto deste.

Vimos nesta seção algumas das formas através das quais as imagens podem interagir com textos verbais para construir sentidos diversos e ainda de que forma podem relacionar-se ou não com os objetivos didáticos para os quais são desenhadas. A seguinte, teceremos algumas considerações sobre a análise.

## 5 Conclusões e considerações finais

Após o levantamento dos modos de expressão de significado utilizados pela professora, detectamos que o ambiente on-line apresenta diversos modos interagindo entre si, às vezes em composições trimodais, como foi o caso da composição imagem/texto/vídeo, que apresentamos na *tabela de modos encontrados em interação com outros modos*. Foram detectadas majoritariamente composições de imagem estática e texto verbal – treze ocorrências. Dentre essas treze composições, encontramos quatro tipos diferentes de interações:

1. Interações cujo status entre a imagem e o texto era de igualdade. Os elementos, texto e imagem, eram complementares e havia entre eles relações lógico-semânticas sobreposta de expansão e projeção. Somente um exemplar de CMI com relações sobrepostas foi detectado no material – a CMI 1 que analisamos.
2. Interações cujo status entre texto e imagem é de desigualdade. A imagem é subordinada ao texto. Nesse tipo de interação as imagens se relacionavam com partes do texto especificando-o. Encontramos nessas composições relações de projeção de significado. Detectamos duas composições multimodais com esse tipo de relação entre os seus elementos. Como exemplo, analisamos a CMI 2.

3. Interações imagem-texto cujo status é complementar e a CMI apresenta relação lógico-semântica de extensão entre seus elementos. Há uma combinação de texto e imagem para formar uma unidade de sentido maior e um modo estende o significado do outro, adicionando uma nova informação. Foram detectadas duas composições dessa natureza, das quais analisamos a CMI 3 como exemplo.

Ao longo das análises, podemos perceber que, embora a disciplina apresente treze composições multimodais instrucionais compostas de imagens estáticas e texto escrito, em oito delas as imagens relacionam-se com apenas um pequeno fragmento do texto, como exemplificamos com a CMI 4. Não se relaciona, portanto, com os objetivos das atividades de forma significativa, ou seja, não contribui para que o objetivo didático da atividade em que está inserida seja atingido. Em apenas três composições, a CMI 1, a CMI 3 e uma com relações similares às descritas na CMI 3, as imagens em interação com o texto verbal relacionavam-se com o objetivo da atividade. Nas três CMI, as relações de status entre texto e imagem eram complementares e/ou havia sobreposições de relações lógico-semânticas, ou um modo estendia os sentidos do outro. O estudo indica haver uma conexão entre os tipos de relação que são construídos entre a imagem e o texto e o potencial que a CMI tem de contribuir para que o objetivo da atividade seja atingido. Contudo, em um estudo de natureza qualitativa e tão restrito em termos de corpus, não se pode comprovar essa afirmação. Pode-se, no entanto, colocar esse fator em destaque para que em pesquisas mais abrangentes se possa testar a hipótese de que as relações que se estabelecem entre imagens e textos contribuem de forma diferenciada para que os objetivos didáticos das atividades nas quais as CMI estão inseridas sejam atingidos.

Nessa tentativa de percepção de como se dão as relações texto-imagem no material didático on-line chegamos a duas considerações. A primeira delas é que a teoria de análise dessas interações proposta por Martinec e Salway (2005) consegue dar conta, como metalinguagem, de explicar as relações que encontramos no material didático. Dessa forma, acreditamos que a teoria tenha muito a oferecer aos professores na tarefa de ajudar os alunos no esclarecimento de como as interações se constituem entre os modos analisados. Percebe-se também que o potencial de contribuição para o aprendizado oferecido por cada modo precisa ser mais bem compreendido pelos professores para que possamos fazer escolhas de elementos que



constituam CMI's facilitadoras do processo de ensino-aprendizagem. Em suma, se entendermos melhor como são construídas as composições multimodais que dão sentido aos conteúdos didáticos propostos, podemos construir CMI's que sirvam de suporte para que os objetivos de ensino sejam atingidos, ou seja, podemos contribuir de forma mais significativa para que o aprendizado ocorra.

Outro aspecto a ser considerado é que, apesar de alguns estudos já terem sido realizados nessa área (ROTH; POZZER-ARDENGHI; HAN, 2005; UNSWORTH, 2006; UNSWORTH; DALY, 2011), ela ainda carece de teorias que venham embasar análises como as que apresentamos neste estudo. Essa escassez se torna patente ao depararmos com mais de dois modos interagindo entre si em composições trimodais, como encontramos na disciplina analisada. Em nossa busca por teorias que embasem análise de construção de sentido de forma multimodal, não identificamos teorias que permitissem uma análise, por exemplo, de interações entre modos tão diversos, como exemplo, vídeo, imagem e texto. Talvez por ainda não termos uma gramática intermodal integrada que descreva características semióticas de forma mais ampla, abrangendo as particularidades de cada modo na construção dos significados contidos em uma composição multimodal.

Além disso, falta-nos a percepção de como essas interações entre os modos se alinham com os objetivos didáticos nos contextos em que composições multimodais instrucionais encontram-se inseridas. Percebemos essa escassez também devido ao fato de que, ao longo de nossas análises, sugeriram questionamentos acerca de como esses modos impactam o aprendizado, como a professora faz suas escolhas entre os modos e como utiliza as interações entre eles a seu favor para atingir seus objetivos de ensino. Durante nossa análise tivemos um pequeno vislumbre de que como isso pode acontecer nos modos imagem estática e texto escrito, mas precisamos de estudos maiores, mais profundos e mais abrangentes para poder estabelecer essas relações de forma mais consistente e derivada do uso que já fazemos de composições multimodais no ensino. Em suma, se entendermos como são construídas essas composições para dar sentido aos conteúdos didáticos propostos e para promover os objetivos didáticos que são estabelecidos nas aulas virtuais, podemos melhorar a eficácia dos materiais didáticos virtuais.

As questões colocadas neste estudo nos revelam que a pesquisa sobre as interações multimodais, embora já tenha caminhado, ainda está

em sua infância. A área oferece um enorme potencial para estudos sobre os conteúdos de ensino expressos em modos diversos com a intenção de atingir objetivos de ensino. Precisamos desenvolver estudos sobre as interações entre os leitores e os modos, estudos sobre o impacto de cada modo no aprendizado, estudos sobre as escolhas dessas interações feitas pelos autores das composições. Enfim, existe um mundo multimodal a ser pesquisado.

## Referências bibliográficas

ANSTEY, M.; BULL, G. *Defining multiliteracies: teaching and learning multiliteracies: changing times changing literacies*. Newark: International Reading Association, 2006.

BARTHES, R. *Image, music and text*. London: Fontana, 1977.

GODOY, A. S. Introdução à pesquisa qualitativa e suas possibilidades. *Revista de Administração de Empresas*, v. 35, n. 2, p. 57-63, 1995.

HALLIDAY, M. A. K. *Language as a social semiotic*. London: Edward Arnold, 1978.

HALLIDAY, M. A. K. *An introduction to functional grammar*. London: Edward Arnold, 1985.

HALLIDAY, M. A. K.; HASAN, R. *Cohesion in english*. London: Longman Group Limited, 1976.

HALLIDAY, M. A. K.; HASAN, R. *Language, context and text: aspects of language in a social semiotic perspective*. Oxford: Oxford University Press, 1989.

JEWITT, C. *The Routledge handbook of multimodal analysis*. London: Routledge, 2011.

KRESS, G. *Multimodality: a social semiotic approach*. New York: Routledge, 2010.

KRESS, G. What is mode? In: JEWITT, C. *The Routledge handbook of multimodal analysis*. London: Routledge, 2011. p. 54-67.

KRESS, G.; LEEUWEN, T. van. *Reading images: the grammar of visual design*. London: Routledge, 1996/2006.

KRESS, G.; VAN LEEUWEN, T. *Multimodal Discourse: The Modes and Media of Contemporary Communication*. London: Arnold, 2001.

LEEUWEN, T. van. *Introducing social semiotics*. London; New York: Routledge, 2005.

LEEUWEN, T. van. Multimodality. In: SIMPSON, J. (Ed.). *The routledge handbook of applied linguistics*. London; New York: Routledge, 2011. p. 668-682.

MARTINEC, R. *Cohesion in action*. *Semiotica*, v. 120, n. 1/2, p. 161-180, 1998.

MARTINEC, R.; SALWAY, A. A system for image-text relations in new (and old) media. *Visual Communication*, v. 4, n. 3, p. 339-374, 2005.

ROTH, W. M.; POZZER-ARDENGHI, L.; HAN, J.-Y. *Critical graphicacy: understanding visual representation practices in school science*. Dordrecht: Springer-Kluwer, 2005.

UNSWORTH, L. Image, text relations and intersemiosis: towards multimodal text description for multiliteracies education. In: 33<sup>RD</sup> INTERNATIONAL SYSTEMIC FUNCTIONAL CONGRESS PROCEEDING, 2006, São Paulo. *Proceedings...* São Paulo: PUC, 2006.

UNSWORTH, L.; DALY, A. Analysis and comprehension of multimodal texts. *Australian Journal of Language and Literacy*, v. 34, n. 1, p. 61-80, 2011.

Data de submissão: 16/01/2016. Data de aprovação: 15/09/2016.