

PROPOSTA DE UMA GRAMÁTICA VISUAL PARA DESCRIÇÃO E ANÁLISE COMPOSICIONAL DE VÍDEOS DIGITAIS EM LÍNGUAS DE SINAIS¹

PROPOSAL OF VISUAL GRAMMAR FOR THE DESCRIPTION AND COMPOSITIONAL ANALYSIS OF DIGITAL VIDEOS IN SIGN LANGUAGES

Luiz Alexandre da Silva ROSADO²
Cristiane Correia TAVEIRA³

RESUMO: Com o advento da produção e ampla circulação de vídeos por meio da internet, os surdos ganharam um espaço novo e independente para expressão cultural, formando, nos últimos 20 anos, acervos digitais que configuram materialmente uma surdo-memória em línguas de sinais. Assim sendo, este artigo apresenta o mapeamento e a proposta de uma gramática visual para os vídeos didáticos e acadêmicos-culturais em línguas de sinais que circulam hoje em repositórios digitais, em especial o YouTube. O objetivo deste estudo é duplo: conceber uma representação icônica para os artefatos produzidos pela/para a comunidade surda no formato vídeo e apresentar um catálogo inicial de soluções visuais capturadas dessas produções, sistematizando-as para futuros autores e criadores de conteúdos. Metodologicamente, focamos na abordagem composicional utilizada na análise de 24 produções selecionadas e estudadas por nosso grupo de pesquisa ao longo do biênio 2017-2018. Como resultado do estudo, definimos sete elementos basilares constituintes e um conjunto de variações e relações possíveis entre eles, tomando por empréstimo e inspiração categorias inicialmente apresentadas por teóricos da imagem e das artes visuais, em especial Arnheim, Dondis e Leborg.

PALAVRAS-CHAVE: Gramática visual. Análise composicional. Vídeo digital. Línguas de sinais. Surdos.

ABSTRACT: With the emergence of production and wide circulation of videos through the internet, deaf persons have gained a new and independent space for cultural expression, forming, in the last 20 years, digital collections that materially constitute a deaf-memory in sign languages. Thus, this paper presents the mapping and proposal of a visual grammar for pedagogic and academic-cultural videos in sign languages that currently circulate in digital repositories, in particular, YouTube. This study has a twofold aim: to propose an iconic representation for video artifacts created by/for the deaf community and to present an initial catalog of visual solutions captured from these products, organizing them for future authors and content creators. In terms of methodology, we focus on the compositional approach used in the analysis of 24 productions selected and examined by our research group in 2017-2018. As an outcome of the study, we have defined seven constituent basic elements and a set of possible variations and relationships between them, borrowing, for inspiration, categories originally presented by image and visual art theorists, including Arnheim, Dondis, and Leborg.

KEYWORDS: Visual grammar. Compositional analysis. Digital video. Sign languages. Deaf.

1 INTRODUÇÃO: ECOLOGIAS MIDIÁTICAS E A BUSCA DE UMA SURDO-MEMÓRIA EM VÍDEO

[...] como sabemos, na verdade, que esses grandes escritores (inclusive, o das sagradas escrituras) não teriam preferido filmar ou gravar seus textos? (Flusser, 2010 [1987], p. 14).

Inúmeros autores, desde Marshall McLuhan (2001 [1964])⁴, passando por Pierre Lévy (1993), Derrick de Kerckhove (2009) e Lúcia Santaella (2010) vêm afirmando que novas Tecnologias de Informação e Comunicação (TICs), ao serem concebidas e depois apropriadas

¹ <http://dx.doi.org/10.1590/s1413-65382519000300001>

² Doutor em Educação, Instituto Nacional de Educação de Surdos (INES), Departamento de Ensino Superior, Rio de Janeiro. alexandre.rosado@gmail.com. Rio de Janeiro/Rio de Janeiro/Brasil. ORCID: <https://orcid.org/0000-0002-2702-5617>

³ Doutora em Educação, Instituto Nacional de Educação de Surdos (INES), Departamento de Ensino Superior, Rio de Janeiro. cristianecorreiataveira@gmail.com. Rio de Janeiro/Rio de Janeiro/Brasil, ORCID: <https://orcid.org/0000-0001-9626-4921>

⁴ Decidimos, nas citações e na bibliografia ao final deste artigo, pôr, entre colchetes, as datas originais das publicações para que o leitor possa situar, de maneira exata, a década em que determinado autor atuou, ao modo de uma linha do tempo teórico-conceitual.

socialmente, geram novas configurações socioculturais e, até mesmo, modos diferenciados de pensamento e linguagem (Santaella, 2004), ao expandirem e oferecerem capacidades diferenciadas aos nossos corpos e ampliar as percepções hauridas por nossos sentidos. É uma concepção de tecnologia como prótese e ampliação das capacidades do corpo humano.

Para esses autores, nossa mente (cognição/percepção) e nossa ambiência (casas que habitamos, locais em que trabalhamos, ruas em que circulamos) foram alteradas ao longo dos últimos milênios por ondas tecnológicas sucessivas. Tais mudanças aceleraram-se nos últimos séculos com a inserção de novas TICs, representadas por seus artefatos e seus produtos. Nesse caminhar, tivemos as tecnologias milenares dos alfabetos e da escrita manual, a fixação da informações por meio da tinta e do papiro, a invenção, no fim da Idade Média, da prensa tipográfica por Gutenberg, os experimentos civis e militares com a lanterna mágica, a facilidade de escrita com a máquina de escrever mecânica, as comunicações a distancia com o telégrafo e o rádio AM e FM, a expansão, no século XX, do cinema, da televisão e, mais recentemente, o processamento das telecomunicações por meio do computador e da internet.

Essas tecnologias formam camadas entre-sobre-superpostas que, até o momento, são progressivas e cumulativas em seu desenvolvimento técnico-científico, mas também se modificam em entrelaces contínuos de influência mútua, gerando apropriações não esperadas por aqueles que as criaram. Na visão flusseriana, consoante com os autores até agora elencados, todo objeto manufaturado/criado transforma as relações do usuário com o seu entorno de modo a tirar dele proveito. Portanto, ao fabricar um artefato, também informamos e damos sentido ao mundo (Flusser, 2013).

Periodizações iniciais foram propostas para distinguir esse percurso, tais como aquelas descritas por Lévy (1993) e ampliadas por Santaella (2010), incluindo basicamente três grandes fases das TICs ao longo de nossa História conhecida: a *oralidade tribal* sem registro (a fala e o gesto fugaz); a *escrita* inicialmente manual e posteriormente impressa (a contabilidade, a economia, os processos burocráticos e as leis), até o tempo presente com os meios de comunicação de massa; e o *formato digital* unificador (a *aldeia global*, a internet e a computação).

Essas fases sintetizam a abordagem antropológica e cultural dos artefatos técnicos, que se tornam pilares definidores de períodos comunicacionais nas sociedades que os utilizam. A princípio, não há exclusão mútua ou substituição absoluta entre os artefatos técnicos, mas períodos marcantes, como a expansão das sociedades letradas baseadas na escrita (Goody, 1987) e na prensa tipográfica a partir do século XV (Burke, 2003 [2000]) e seu aprofundamento com a expansão da cultura de massa e dos meios necessários para a reprodutibilidade técnica dos conteúdos comunicacionais nos séculos XIX e XX (Benjamin, 1987). No início do século XX, o crescimento do complexo industrial-comunicacional causou assombro e indagações sobre as consequências na mente humana e nos comportamentos cotidianos, a exemplo daqueles relatados por Simmel (1973 [1902]) quando observa a aceleração da vida e o desenvolvimento de certa indiferença sistêmica entre os milhares de habitantes das metrópoles europeias nascentes.

No tempo presente, a internet e a computação (computadores pessoais, celulares e *tablets*) permitem a adoção, em larga escala, da produção e publicação de vídeos digitais, levando ao surgimento de uma ecologia midiática da imagem em movimento que, diferentemente da cultura de massa associada ao rádio, à TV e ao Cinema ao longo do século XX, é agora alimen-

tada por uma constelação de pequenos novos produtores em suas casas e estúdios particulares, os *youtubers* e *vloggers*, para os vídeos, e os *bloggers*, para os textos. No começo dos anos 2000, a liberação do *polo da emissão* (Lemos, 2009), em um primeiro momento, gerou a fúria de conservadores que atribuíram tal fenômeno a uma regressão cultural (anti-intelectual) comandada por uma legião de produtores amadores, que não teriam a mesma força, confiabilidade e *status* dos profissionais da mídia (especialmente os jornalistas) e os pesquisadores universitários acadêmicos (Keen, 2009 [2007]).

Recuando algumas décadas no século XX, tais mudanças e transições derivadas das TICs já eram percebidas e teorizadas por Vilém Flusser (2010 [1987]), que nos traz a concepção de uma escrita manual e impressa que, em um primeiro momento, colaborou para a descentração da oralidade primária a-histórica (em que o tempo é concebido em círculos), com a construção de um novo acervo de memória para a humanidade por meio de bibliotecas, arquivos públicos, museus, nos levando à concepção de uma memória-história linear (tempo visto como algo que percorre uma linha reta). A escrita, o homem radicado na tribo-cidade, a manufatura-fábrica e a noção de um tempo linear e progressivo estão intimamente ligados, assim como a sua materialidade presente nos relógios e na contagem do tempo com os calendários (Elias, 1998 [1984]), o que ajudou a fundar nossa percepção de homem-moderno-histórico. Flusser, um pensador radical das mídias, então prevê que, a partir de agora, as novas mídias analógicas e digitais terão a potência de sobrepor, ou mesmo substituir, diversos usos até então associados à escrita, como a produção de conhecimentos científicos, as discussões políticas, o fazer poético e até mesmo a construção conceitual filosófica.

Teriam as novas mídias tal poder? Na comunidade surda, ao menos, parecem ter um papel cada vez mais central e estruturante, após longo período de domínio da escrita. Em relação à escrita e a estruturação da memória nas sociedades ocidentais, percebemos que semelhante caminho está sendo percorrido com as línguas de sinais⁵, que constroem na internet acervos em vídeo voltados à comunidade de pessoas surdas e ouvintes usuárias dessas línguas, o que podemos chamar de uma surdo-memória, pouco dependente dos caracteres alfabéticos das línguas orais em sua expressão-comunicação e inscrita em *hard disks* da difusa malha digital de servidores de internet. Se no primeiro caso o processo de amadurecimento da revolução de Gutenberg e sua prensa tipográfica durou, ao menos, quatro séculos (XVII ao XX), no tempo presente estamos vivendo um renascimento visual da cultura e comunicação viso-gestual surda, por meio do compartilhamento crescente de vídeos, acumulados em curto período (2005-...), a partir da expansão do que convencionou-se chamar *Web 2.0* (O'Reilly, 2005), das câmeras digitais em *smartphones* e *tablets* e dos sites-acervos-redes sociais (*YouTube*, *Facebook*, *Vimeo*, entre outros).

Embora esse rico acervo de vídeos possa ser analisado na perspectiva dos seus conteúdos expressos (análise do discurso) ou mesmo por intermédio das reações e das percepções de seu público (estudos da audiência e da recepção), optamos, neste artigo, pelo enfoque da análise

⁵ Ao leitor não familiarizado, informamos que existem inúmeras línguas de sinais, geralmente associadas a cada país: Língua Brasileira de Sinais (LIBRAS), American Sign Language (ASL) e Língua Gestual Portuguesa (LGP) são alguns exemplos. Cada língua possui um conjunto próprio de sinais (itens lexicais) e princípios gramaticais, distinguindo-se mais ou menos uma das outras a partir de sua genealogia: a LIBRAS, por exemplo, tem forte influência da Língua de Sinais Francesa (LSF). Essas línguas possuem vasto número de sinais já catalogados e dicionarizados (Capovilla & Raphael, 2001), porém, caso não possuam um sinal específico para determinado conceito, utilizam o que é chamado *datilologia*, ou seja, a soletração letra-por-letra de um vocábulo da língua oral por meio de sinais usando as mãos.

composicional (Rose, 2007). A análise composicional é uma abordagem que leva mais em consideração os elementos presentes no material em si do que o contexto em que ele foi produzido ou a cultura própria dos atores que os assistem e interpretam seus significados. Portanto, é nessa ambiência extremamente rica e original das comunidades surdas e suas produções na internet que procuraremos propor uma gramática, não aquela já conhecida dos textos escritos e há muito explorada e amadurecida em sua estrutura pelos linguistas, mas de produções registradas na forma de vídeos e que usam uma forma de comunicação nativamente viso-gestual: as línguas de sinais.

2 PERCURSO METODOLÓGICO

Este artigo atende a um dos objetivos de nossa pesquisa realizada no triênio 2015-2018 (Taveira & Rosado, 2015), a saber:

Atender à necessidade premente de coletar, armazenar, analisar, classificar e organizar os produtos derivados da prática pedagógica, com ênfase nos artefatos e métodos que combinem a experiência visual, a tradução e interpretação (Libras/LP) e a comunicação em ambientes virtuais e presenciais. (Taveira & Rosado, 2015, p. 2).

Focamos na análise de um conjunto de produções em formato vídeo digital, em que a Língua de Sinais é central e os surdos são o principal público-alvo. A partir dessa análise, descrevemos e propomos, de modo inicial, os elementos constituintes de uma gramática visual dos vídeos em línguas de sinais. Nossa abordagem foi centrada nos formatos (a *forma* do binômio forma-conteúdo) ou, de modo mais preciso, nas composições obtidas de quadros (*frames*) capturados dos vídeos em línguas de sinais. Houve, portanto, o resgate e a valorização do entendimento sobre os elementos constituintes da linguagem visual, catalogados desde as décadas de 1950, 1960 e 1970 por estudiosos dos campos da Psicologia, da Arte e do Design (Leborg, 2015 [2004]; Dondis, 2007 [1973]; Arnheim, 1992 [1954, 1974]).

Historicamente, o interesse desses estudos é consequência direta do aumento da presença da pintura, do desenho e da fotografia ao longo do século XX (artes visuais), com a expansão progressiva da reprodução técnica da imagem em veículos de comunicação, a indústria cultural (Benjamin, 1987), tais como jornais, livros e revistas; assim como a intenção de extrapolar a falsa ideia de que basta ter *bons olhos* para se compreender e descrever obras de arte, o que prescindiria de um aprendizado e desenvolvimento desse olhar (letramento visual). Apesar da também crescente presença de vídeos reproduzidos na TV e no Cinema, o interesse *gramatical* desses autores sobre elementos da visualidade se ateve muito mais às mídias estáticas e às variações e às relações desses elementos, o que pode ser adaptado para os quadros estáticos (*frames*) capturados de um vídeo.

Nossa análise empírica partiu de um recorte amostral de vídeos produzidos para fins didáticos e acadêmicos-culturais, o que incluiu vídeos artísticos-literários, novelísticos e jornalísticos. Para isso, foi construída uma ficha de análise padrão, dividida em três partes. A primeira parte identifica *dados básicos da produção* (título do material, ano de produção, ano de publicação, título da coleção e duração do vídeo). A segunda parte procura reproduzir tanto a *apresentação do material* feita pelos seus produtores (quando há) quanto a *apresentação sintético-analítica* redigida pelo membro do grupo de pesquisa que assistiu a todo o material.

Por fim, a terceira parte, mais importante, contém 14 categorias de análise, a saber: (1) tipo de conteúdo do material; (2) línguas presentes; (3) disposição das línguas; (4) facilitadores de leitura; (5) linguagem; (6) gêneros; (7) tipo de material final; (8) ambiência de aplicação do artefato; (9) modo de distribuição; (10) público-alvo; (11) faixa etária do público-alvo; (12) responsável pela produção; (13) dispositivos de acessibilidade; e (14) efeitos de edição e pós-produção. Cada categoria, na ficha, deve ter a marcação de seus itens justificada pelo pesquisador/analista, compondo a defesa da escolha e visando contribuir para o progressivo esclarecimento do significado da categoria entre os pesquisadores.

De outubro de 2017 a dezembro de 2018, completamos a análise de 24 vídeos, todos em Libras, com exceção de um deles em Língua de Sinais do Uruguai (LSU). Esse trabalho de análise gerou as respectivas fichas preenchidas pelos membros do grupo de pesquisa “Educação, mídias e comunidade surda”, ora individualmente, ora em duplas, sendo em seguida debatidas coletivamente nas reuniões com os demais integrantes do grupo no INES, momento que acréscimos são realizados e a ficha é validada e arquivada. Essas reuniões contaram, em média, com cinco integrantes-analistas.

Importa, neste trabalho, focar na categoria número 14, relativa aos efeitos de edição e pós-produção de vídeo. Esse trabalho posterior à gravação em estúdio, com a popularização de *softwares* como *Adobe Premiere* (edição de vídeo) e *Adobe After Effects* (efeitos de pós-produção), entre outros semelhantes, se tornaram fundamentais na criação de vídeos multimodais em línguas de sinais, em que na composição estão presentes, além do intérprete sinalizante, imagens fotográficas, gráficos e esquemas, textos em variadas tipografias e animações computadorizadas de todos esses elementos. De certa forma, esses *softwares* e suas possibilidades de tratamento de vídeo em múltiplas camadas sobrepostas condicionam hoje as composições visuais de grande parte dos vídeos que encontramos no *You Tube* e em outros repositórios *online* de vídeos, constituindo uma versão contemporânea da afirmação clássica de McLuhan (2001 [1964]): “o meio [de pós-produção] é a mensagem”.

O objetivo inicial da categoria 14 foi a descrição com alto grau de precisão desses elementos de pós-produção, cada vez mais utilizados pelos editores de vídeo e designers gráficos à medida que eles têm acesso a esses recursos em computadores relativamente baratos e, portanto, cada vez mais acessíveis. Contudo, pouco a pouco, os integrantes do grupo de pesquisa sentiram necessidade de descrever também as diferentes disposições, relações e variações dos elementos básicos detectados nos vídeos em línguas de sinais, requalificando essa categoria para a análise composicional dos *frames* dos vídeos. Depois de 24 vídeos analisados, percebemos a estabilidade alcançada em nosso processo descritivo/perceptivo, que agora pode ser apresentado publicamente neste artigo.

3 DEFININDO OS CONTORNOS DE UMA GRAMÁTICA VISUAL CENTRADA NA IMAGEM

Nós pensamos de maneira diferente quando temos uma linguagem para descrever o que pensamos (Leborg, 2015 [2004], p. 5).

Acontece com frequência vermos e sentirmos certas qualidades numa obra de arte sem poder expressá-las com palavras. A razão de nosso fracasso não está no fato de se usar uma linguagem, mas sim porque não se conseguiu ainda fundir essas qualidades percebidas em categorias adequadas (Arnheim, 1992 [1954, 1974], p. XIV).

Quantos de nós veem? (Dondis, 2007 [1973], p. 5).

A visualidade, e a conseqüente experiência visual, é tema hoje fundamental no campo de estudos da Educação de Surdos, sobretudo pelo uso das línguas de sinais como línguas viso-espaciais (não orais) em que o uso do corpo e do olhar são predominantes na comunicação. Ao mesmo tempo, existe uma busca por mais compreensão do desenvolvimento da aprendizagem escolar por meio do uso de recursos visuais (Luckner, Bowen, & Carter, 2001; Lebedeff, 2010), especialmente no formato vídeo (Lebedeff & Santos, 2014). A gestualidade surda contrasta com a educação ouvinte, baseada, em grande parte, na comunicação por meio da fala e de materiais centrados no texto escrito. A visualidade tem maior potência de manifestar-se e ser materializada em categorias/artefatos como imagens fotográficas, pinturas, gráficos, infográficos, tabelas, mapas mentais, desenhos, maquetes, esculturas e filmes em que o texto escrito linear, o elemento verbal baseado na língua oral, não é o centro predominante da organização das informações e expressão dos conteúdos.

Apesar dessa divisão estanque aqui apresentada, concordamos com a concepção triádica das *matrizes de linguagem* (sonora-visual-verbal) proposta por Santaella (2005) em que o *verbal* também pode se manifestar nas categorias do primeiro grupo, assim como o *visual* também por vezes é basilar em textos escritos, mais evidente em gêneros como o poético e o romance ficcional, gêneros em que formamos imagens mais ou menos detalhadas (representações) em nossas mentes à medida que percorremos o texto. Por outro lado, estamos cientes também de que uma imagem pode conter um argumento (Mateus, 2016), ou seja, apresentar um conjunto de premissas e uma tese central, estruturas típicas do texto linear escrito e, antes dele, da oralidade, quando a retórica floresceu na antiga Grécia. Tais aportes mútuos entre categorias outrora tão marcadamente distintas talvez seja resultado do crescimento exponencial dos “textos” agora chamados multimodais (Kress & Van Leeuwen, 1996; Kress, 2010), ou seja, que em uma só composição voltada a uma finalidade comunicacional específica reúnem-se simultaneamente textos escritos, imagens, vídeos e sons, como a multimídia dos *softwares* e a hipermídia das páginas que acessamos na internet, tornando mais complexo o que convencionou-se chamar letramento (um *multiletramento?*).

Nossa abordagem aqui, porém, será mais focada. A análise de artefatos, experiências e didáticas voltadas ao ensino de alunos surdos foi anteriormente explorada por nós (Taveira & Rosado, 2016), cabendo agora, neste artigo, o breve detalhamento sobre o significado do que chamamos de *letramento visual* (ou alfabetismo visual) e o entendimento sobre elementos fundamentais da visualidade compilados sob a expressão *gramática visual*. Sob o ponto de vista de Santaella:

A alfabetização visual significa aprender a ler imagens, desenvolver a observação de seus aspectos e traços constitutivos, detectar o que se produz no interior da própria imagem, [...]. Ou seja, significa adquirir os conhecimentos correspondentes e desenvolver a sensibilidade necessária para saber como as imagens se apresentam, como indicam o que querem indicar, qual é o seu contexto de referência, como as imagens significam, como elas pensam, quais são os seus modos específicos de representar a realidade. (Santaella, 2012, p. 13).

Para a autora, e também para nós, a visualidade é aprendida, ou seja, não é inata, não é natural e nem muito menos espontânea, devendo para isso ser sistematizada e concebida em uma linguagem que capacite o observador a desmembrar uma imagem em partes distintas, para decodificá-la em sua totalidade. A autora cita os aspectos e traços constitutivos das imagens,

que uma vez compreendidos, tornam os sujeitos hábeis na leitura de imagens. É em autores ligados às Artes e à Comunicação Visual (*Design*) como Arnheim (1992), Dondis (2007) e Leborg (2015) que encontramos, objetivamente, esses traços descritos.

Ser letrado visualmente, então, é aprender e exercitar na observação cotidiana o que Leborg chamou de *gramática visual*, um conjunto de elementos e de atividades/relações entre estes elementos que conseguiríamos distinguir ao analisar as composições visuais, *lendo-as* e, com isso, nos tornando visualmente alfabetizados. Para Leborg, o aprendizado desses elementos, no cotidiano, está relacionado mais à experiência física do sujeito, que, muitas vezes, não teve acesso a uma linguagem verbal e sua consequente sistematização, lacuna que ele procura preencher em seu livro.

Embora Leborg, e seus antecessores Dondis e Arnheim, sejam referências fundamentais deste nosso estudo, o ensino desses elementos está presente também em manuais voltados à prática do *Design* e da Arte Gráfica (análogica ou digital), como em Williams (1995), que reduz essa arte a quatro princípios basilares: (1) proximidade, (2) alinhamento, (3) repetição e (4) contraste. Leborg (2015), como vemos nas imagens da Figura 1 a seguir, extraídas do sumário de seu livro, expande o *vocabulário da visualidade* de modo que consigamos detalhar, de modo muito preciso, as composições visuais que experimentamos sensorialmente em nosso cotidiano.

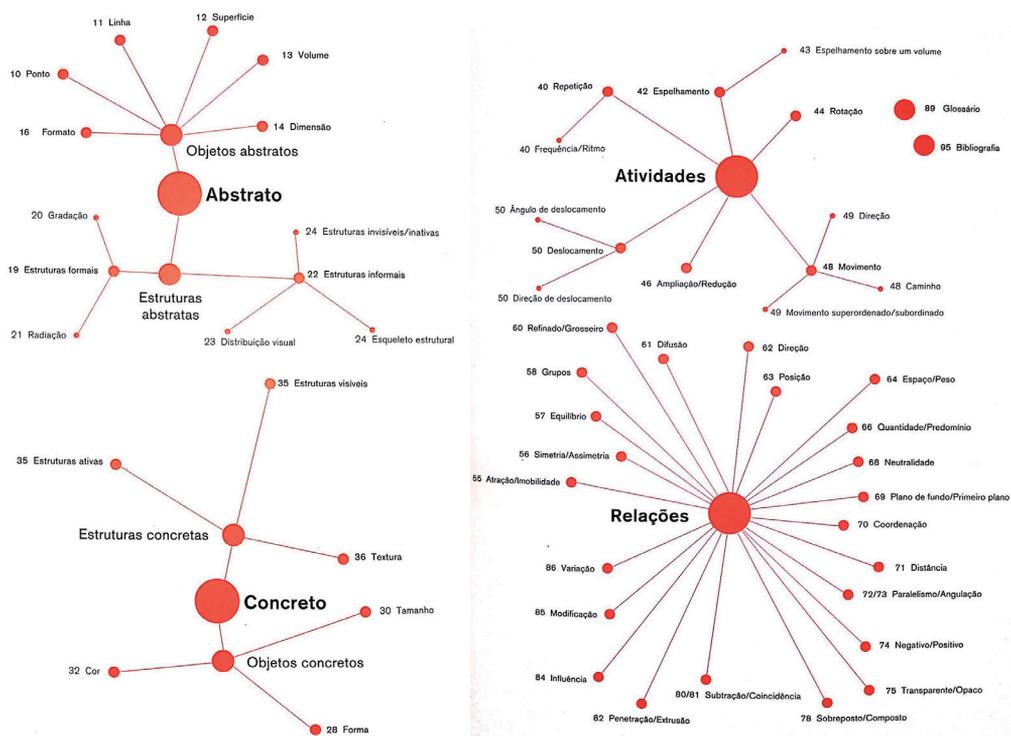


Figura 1. Síntese dos quatro grandes grupos propostos por Leborg em sua gramática visual. Fonte: Leborg (2015, p. 96-97).

Notamos como Leborg sintetizou categorias de análise de uma imagem inspiradas em categorizações prévias (Arnheim, 1992; Dondis, 2007), estando aquelas mais basilares relacionadas ao *abstrato* e ao *concreto* e as ações/relações dos/entre elementos basilares nos levariam às *atividades* e às *relações*. Não pretendemos aqui descrevê-las uma a uma, mas apresentar ao leitor o nosso ponto de partida. Foi com essa mesma concepção de separar os elementos básicos e relacioná-los em padrões combinatórios que nos voltamos à construção de uma gramática visual a ser utilizada na análise de vídeos em línguas de sinais.

4 PROPOSTA DE UMA GRAMÁTICA VISUAL PARA OS VÍDEOS DIGITAIS EM LÍNGUAS DE SINAIS

Procuramos, a partir da concepção de gramática visual desenvolvida por Donis A. Dondis (2007) para as imagens estáticas, resgatadas em uma refinada releitura sintética por meio de representações visuais por Christian Leborg (2015), os elementos constituintes basilares desses vídeos, que são basicamente imagens (*frames*) em sequência gerando a percepção temporal de movimento.

Propomos então sete elementos basilares para as composições de cada *frame* (Figura 2), a saber: (1) o ator/intérprete sinalizante, ou seja, a pessoa que utiliza a língua de sinais para se expressar; (2) o ator/intérprete usando língua oral, isto é, a pessoa que fala para se expressar; (3) a massa textual, na forma de títulos e textos descritivos em escrita alfabética; (4) a ilustração/imagem, gráfico ou fotografia; (5) a legenda em língua oral escrita alfabética; (6) o cenário natural ou fundo artificial inserido por intermédio de substituição pela técnica do *Chromakey*; e (7) o vídeo menor sobre o vídeo principal, o *Picture-in-Picture* ou PIP.

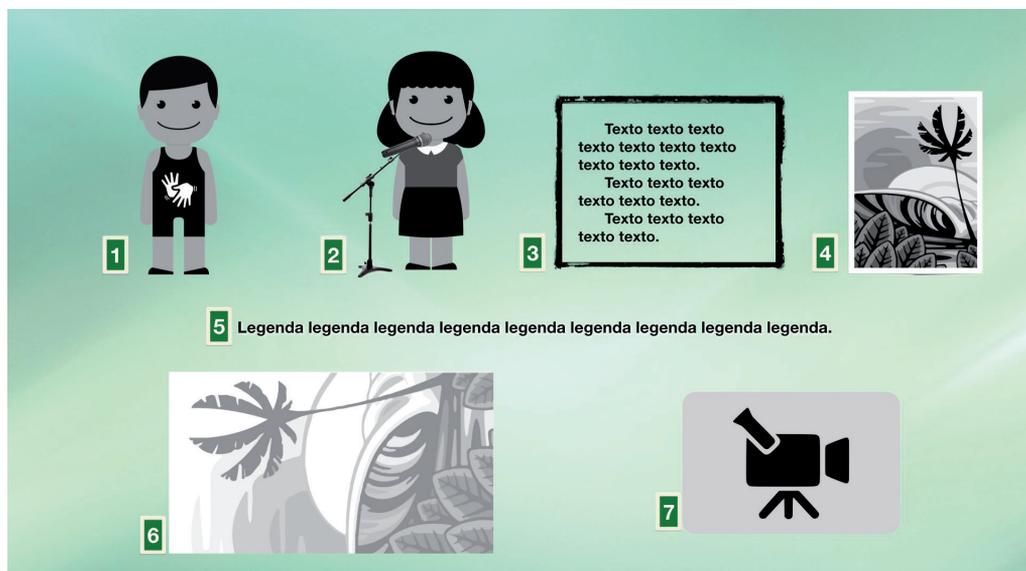


Figura 2. Proposta de sete elementos básicos para os vídeos em línguas de sinais.
Fonte: Elaboração nossa.

Esses sete elementos listados são uma proposta de definição de unidades básicas de composição de um vídeo em língua de sinais, sendo analisadas por meio da captura de um quadro (*frame*) do vídeo sob análise. Com esses elementos basilares, podemos criar uma representação icônica do vídeo, abstraindo e reduzindo o vídeo a seus elementos básicos com representações gráficas simples posicionadas proporcionalmente na tela, o que auxilia na detecção de suas relações e variações mais típicas.

A partir desses elementos, exemplificamos, a seguir, alguns desses exercícios de aplicação em *frames* de vídeos que compuseram nossa empiria até o momento. Cabe observar que à direita estão as telas capturadas com a identificação, por meio de números, dos elementos presentes no quadro e, à esquerda, está a representação icônica do quadro original (Figuras 3 a 6).



Figura 3. Exercício com representação a partir de dois elementos básicos.
Fonte: Elaboração nossa.



Figura 4. Exercício com representação a partir de três elementos básicos.
Fonte: Elaboração nossa.



Figura 5. Exercício com representação a partir de quatro elementos básicos.
Fonte: Elaboração nossa.



Figura 6. Exercício com representação a partir de cinco elementos básicos.

Fonte: Elaboração nossa.

À medida que fomos observando e analisando os vídeos componentes de nossa empiria (total de 24 até a presente data), detectamos e catalogamos alguns tipos de variações dos elementos básicos nas composições, chegando a cinco grandes tipologias: (1) tamanho (em algumas ocasiões os elementos estão maiores e, em outras, menores); (2) corte: um ou mais elementos estarem inteiros ou parcialmente exibidos; (3) posição na tela: se destacarem em diferentes áreas da composição (centro, bordas superiores e inferiores, laterais esquerda e direita); (4) grupos: quando se repetem sequencialmente e estão próximos uns dos outros, formando agrupamentos; (5) formato e espaçamento: quando variam em sua forma em *frames* diferentes ou na distância entre os elementos.

As variações dos elementos básicos só podem ser detectadas e definidas quando um conjunto de *frames*, sequenciais ou não em um mesmo vídeo ou de vídeos de origens variadas, podem ser cotejados conjuntamente, após seguidas observações e registros. Logo, é uma forma de descrição temporal do vídeo ou de um conjunto de vídeos.

De modo específico, as variações de tamanho e corte dos elementos de um vídeo se relacionam diretamente ao que é chamado de *planos de câmera* ou *enquadramento*, na linguagem cinematográfica e televisiva já exaustivamente detalhada em manuais (Thompson & Bowen, 2009): plano geral, médio, americano, primeiro plano, *close* e detalhe. Um plano *close*, por exemplo, teria um elemento do *frame* predominante com variação de tamanho e corte a ponto de ocupar completamente o quadro.

Pensamos que termos descritivos de planos de câmera são relacionados mais à descrição de pessoas e objetos presentes em um cenário – elementos típicos de vídeos narrativos novelísticos, documentais e ficcionais de cinema e TV –, mas não a elementos adicionados artificialmente após a filmagem obtida com o uso de uma câmera filmadora, tais como mensagens de texto, gráficos, fotografias e imagens, presentes na maior parte dos vídeos com fins didáticos que catalogamos.

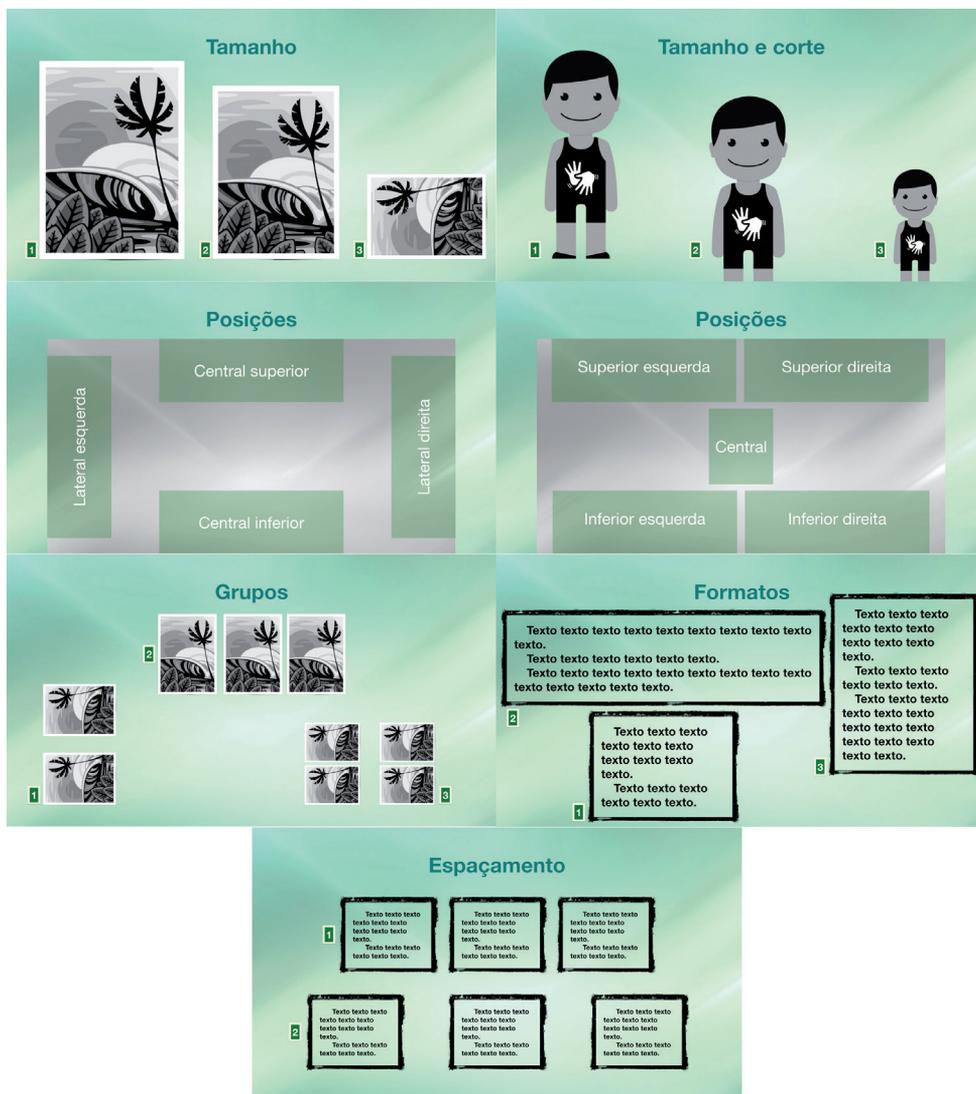


Figura 7. Catálogo de possíveis variações dos elementos básicos e sua exemplificação.
 Fonte: Elaboração nossa.

Tomamos, então, a decisão de criar um conjunto de categorias mais abrangentes, que dessem conta da multimodalidade permitida pelos editores de vídeo e de pós-produção hoje presentes nos computadores pessoais e amplamente utilizados em ambientes profissionais, semi-profissionais e amadores. De certa forma, propomos aqui uma linguagem descritiva em que o centro analítico se moveu da câmera filmadora (processo de captura *natural*) para o editor de vídeo e efeitos (mixagem de elementos composicionais e finalização *artificial*).

Além das variações dos elementos basilares, também propomos algumas relações desenvolvidas entre os elementos básicos de uma composição, inspiradas em Leborg, que catalogou inúmeras delas, das quais destacamos algumas também encontradas nos vídeos analisados:

1. Repetição: quando um elemento se repete várias vezes na composição ou ao longo do vídeo e permite a detecção de continuidade de determinado padrão, a exemplo de textos que usam a mesma fonte e tamanho marcando títulos de capítulos, os cortes de transição usados sempre que determinada situação acontece, os cenários com cores e texturas que se repetem em momentos distintos, ou o uso de atores sinalizantes/falantes que vestem a mesma roupa/tipo/cor e estão dispostos na mesma área e com o mesmo tamanho em momentos distintos.
2. Simetria: quando há um equilíbrio da mancha de elementos nos dois lados da composição, esquerdo e direito, ou entre o topo e a base, gerando conforto no leitor de vídeos, que os percebe distribuídos de forma harmônica, a exemplo do ator sinalizante em um lado e imagens ilustrativas no outro.
3. Assimetria: é o desequilíbrio na composição, quando os elementos estão agrupados e “pesando” em um dos lados ou em uma das regiões do *frame*, o que leva o olhar do leitor para um dos polos intencionados pelo produtor do vídeo, um recurso para gerar ênfase em alguns elementos e secundarizar outros.
4. Ampliação/redução: quando um elemento de mesma natureza está representado em tamanhos diferentes ao mesmo tempo ou ao longo do vídeo, a exemplo do intérprete sinalizante que, liberto da janela que antes o restringia no canto da tela, pode agora ser ampliado em momentos de maior ênfase na língua de sinais, ou então uma imagem que pode ter um detalhe ampliado para gerar uma leitura mais precisa e enfática pelo espectador.
5. Atração/proximidade: são conjuntos de elementos agrupados que se atraem e indicam alguma relação entre eles, auxiliando o leitor de vídeo a relacionar e categorizar elementos dentro da composição, como no caso da legenda em língua de sinais que aparece sempre próxima ao ator oralizante no filme.
6. Peso: para qual lado da composição os elementos estão pesando mais, tal como uma força gravitacional sentida pelo espectador, também colaborando para a ênfase em determinado polo do *frame*, sendo uma experiência de contraste entre áreas mais *cheias* e áreas mais *vazias*, direcionando o olhar do leitor de vídeo, em geral, para a área *mais pesada*.
7. Quantidade/predomínio: a área da composição em que predominam os elementos, provocando a percepção de um campo bastante pigmentado com atributos, número maior de elementos e manchas maiores de leitura contendo textos, ilustrações, atores, entre outros.
8. Espaço: são as áreas vazias e as áreas densas que se destacam na composição, sendo as áreas vazias necessárias para o *respiro* do leitor visual, evitando a competição de elementos que podem levar a uma superocupação da tela e à perda do conteúdo/informação pelo superestímulo da percepção, algo que ocorre em vídeos em língua de sinais que abusam no uso de elementos de ênfase e destaque em um mesmo *frame*.

9. Sobreposição: quando um elemento está posicionado sobre outro elemento, a exemplo de trabalhos científico-acadêmicos em língua de sinais com citação e rodapé sobrepostos e em menor tamanho em relação ao ator sinalizante que está *pausado* e ocupando todo o *frame*.

As relações entre os elementos basilares só podem ser definidas e detectadas em um mesmo vídeo (unidade) e, em muitos casos, com a captura de um único *frame* deste vídeo.

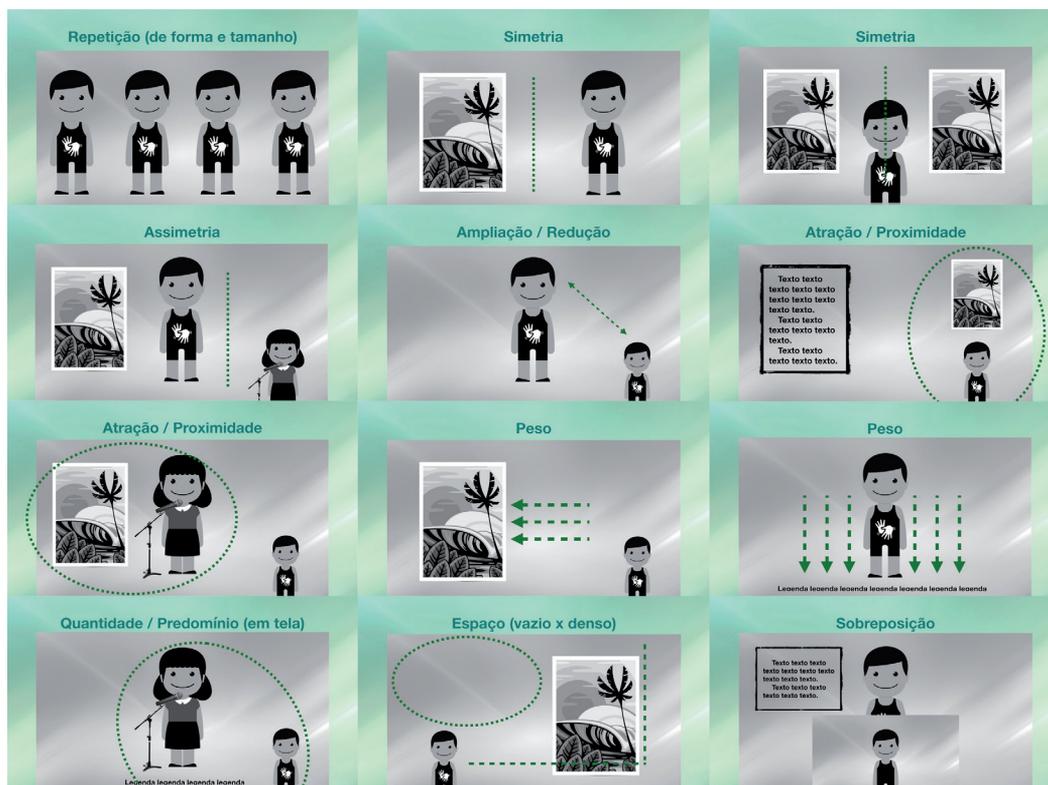


Figura 8. Catálogo de possíveis relações entre os elementos básicos.

Fonte: Elaboração nossa.

Até dezembro de 2018, catalogamos 187 combinações envolvendo os sete elementos básicos em uma composição (*frame*) de vídeo. São combinações que vão de um elemento presente no *frame* até sete elementos simultaneamente no *frame*. O grupo de pesquisa criou um arquivo PPT (*Microsoft PowerPoint*) com os sete elementos básicos, o que permite, a partir de cada vídeo assistido, a construção da representação icônica contendo as combinações entre os elementos básicos detectadas durante a análise, acrescentando essa representação na ficha de análise. Essa construção só é necessária quando as 187 combinações catalogadas são insuficientes para representar algum *frame* do vídeo analisado.

Importante ressaltar que, desse conjunto de combinações catalogadas a partir de nossa amostra de vídeos, podemos também inferir possíveis combinações ainda não encontradas,

porém perfeitamente viáveis de serem descobertas ou, até mesmo, criadas. O potencial da gramática visual dos vídeos em línguas de sinais não é somente descritivo/representativo do que já foi produzido, mas também prospectivo/criativo, permitindo a representação visual de possibilidades composicionais de futuras obras (virtualidades) para artistas, desenhistas, editores de vídeo e *designers* gráficos.

A partir dessa proposta inicial de uma gramática visual dos vídeos em línguas de sinais, é possível aproximar e diferenciar soluções visuais que estão sendo propostas, testadas e apresentadas em inúmeros vídeos atualmente em circulação na internet provenientes de instituições públicas, empresas privadas, organizações não governamentais e produtores independentes. Também é possível o catálogo de soluções originais, específicas, muitas vezes únicas, encontradas por determinado produtor de vídeo e, uma vez catalogadas formalmente e difundidas de maneira sistematizada, podem ser utilizadas em outras produções. Apresentaremos, brevemente, três exemplos que compõem o catálogo atualmente em fase de criação:

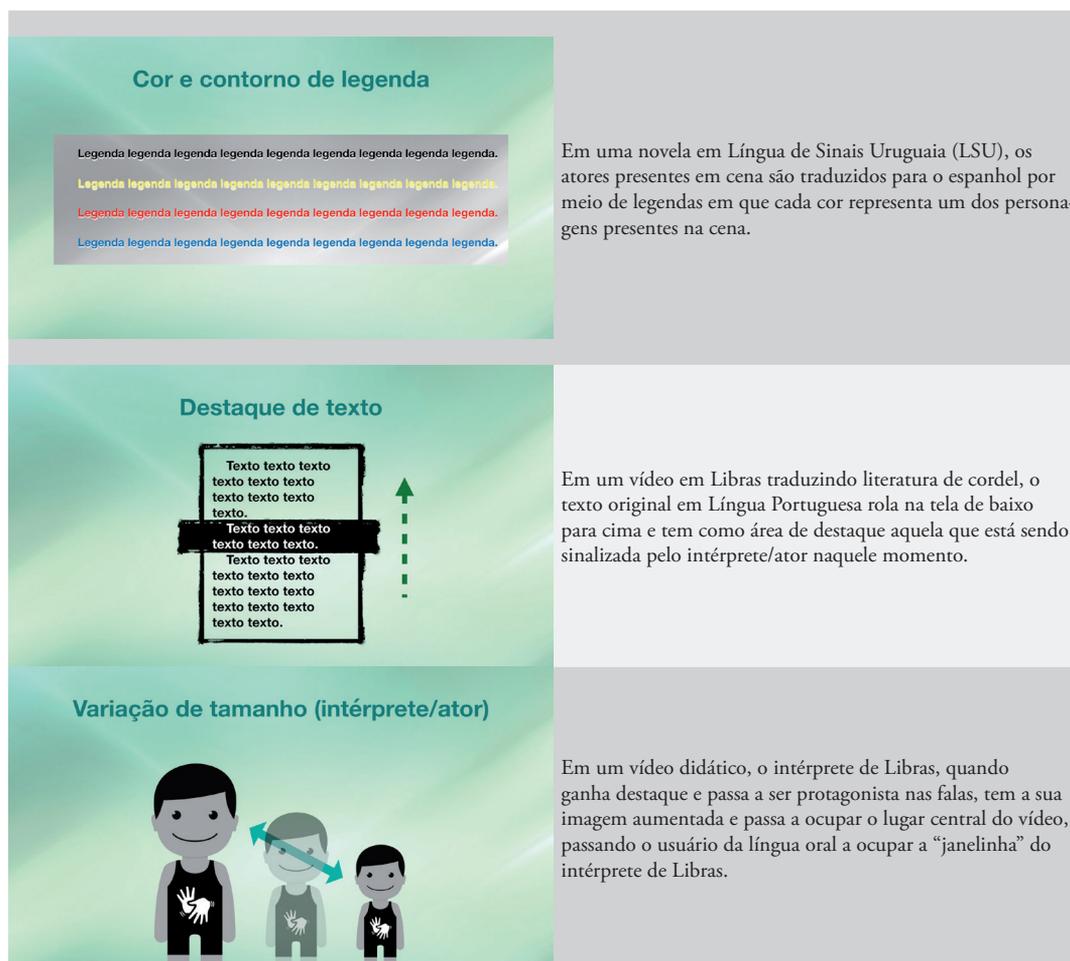


Figura 9. Três exemplos do catálogo de soluções visuais atualmente em construção.

Fonte: Elaboração nossa.

Como já mencionamos, são soluções aparentemente simples, porém pouco ou mesmo nunca utilizadas em outros vídeos em línguas de sinais. Uma vez reunidas, difundidas e conscientemente previstas na forma de catálogo de soluções visuais, os editores, os *designers*, os artistas e os produtores de vídeos poderão aplicá-las e ampliar o alcance de novos materiais e, conseqüentemente, o conforto visual para o público-alvo surdo.

5 CONSIDERAÇÕES FINAIS: UM VOCABULÁRIO DA VISUALIDADE EM CONSTRUÇÃO

Estamos em um momento singular de nossa história, em que tecnologias digitais conectadas em redes de alcance planetário possibilitam a criação e o compartilhamento de vídeos digitais em larga escala por meio de websites-repositórios. Esse cenário, com a *aldeia global* proposta por McLuhan (2001 [1964]), nos anos de 1960 e parcialmente concretizada nos anos 2000, tornou-se extremamente propício para a criação de acervos de vídeos em línguas de sinais, formando uma *surdo-memória* a partir dessas novas condições materiais sintetizadas na forma de computadores, câmeras e editores de vídeos digitais.

Neste artigo, apresentamos uma proposta de mapeamento e de codificação da composição desses vídeos, oferecendo as bases para a construção de uma gramática visual que dê suporte aos produtores de vídeos em línguas de sinais esquematizarem previamente suas concepções em representações icônicas. Apresentamos alguns exemplos de itens de um catálogo de melhores práticas encontradas neste emergente espaço de construção comunitária, histórica e identitária, para que esses mesmos produtores não partam do grau zero em suas criações, mas se inspirem em soluções já propostas, testadas e utilizadas antes deles.

Até o momento não encontramos, na literatura nacional e internacional, proposta semelhante a essa aqui apresentada⁶, com a análise composicional e a proposta de uma gramática visual totalmente voltada aos vídeos em línguas de sinais. Donis Dondis e Christian Leborg, assim como Gunther Kress, voltaram-se a análises universais da imagem e da fotografia, utilizadas por nós como ponto de partida, mas adaptadas ao conjunto de produções da/para comunidade surda.

Na qualidade de pesquisadores no campo da surdez, estamos experimentando sensorialmente diversos vídeos produzidos pela/para/na comunidade surda, especialmente a brasileira, nos atribuindo a tarefa de decodificar e tornar legíveis as múltiplas camadas presentes nessas composições visuais. Somente a partir da formalização das bases de uma gramática visual para os vídeos em línguas de sinais que vimos aumentar nossa agilidade em detectar e representar esses elementos e suas mixagens, de modo prático e quase automático, com esquemas simples e, ao mesmo tempo, generalizantes.

A não-expressão da imagem por palavras, uma dificuldade sentida por Arnheim e que instigou posteriormente Dondis e Leborg a formalizarem um *vocabulário da visualidade*, deu-nos ensejo para realizar a passagem de uma linguagem visual sintética, global, porém dispersiva aos olhares, para a verbalização conceitual daquilo que experimentávamos sensorialmente.

⁶ A pesquisa bibliográfica nacional e internacional foi realizada a partir do acesso ao site da *Divisão de Bibliotecas e Documentação* da PUC-Rio, ao longo dos anos de 2017 e 2018, com as expressões “Visual Grammar”, “Deaf”, “Sign language”, “Deaf education” e suas derivações e combinações. O conjunto abrangente de bases de dados assinadas pela PUC-Rio está especificado no endereço http://testaremoto.dbd.puc-rio.br/sitenovo/base-de-dados.html#bases_assinadas.

Nesse exercício de pesquisa sensível, produzimos categorias descritivas e representações visuais necessárias, apresentadas aqui na forma de texto linear e suas respectivas imagens icônicas, uma forma de organização imperativa para a proposta de gramática visual ora aqui apresentada.

REFERÊNCIAS

- Arnheim, R. (1992 [1954, 1974]). *Arte & percepção visual: uma psicologia da visão criadora*. São Paulo, SP: Livraria Pioneira.
- Benjamin, W. (1987). A obra de arte na era de sua reprodutibilidade técnica. In W. Benjamin (Org.), *Magia e técnica, arte e política: ensaios sobre literatura e história da cultura* (3a ed., pp. 165-196). São Paulo, SP: Brasiliense.
- Burke, P. (2003 [2000]). *Uma história social do conhecimento: de Gutenberg a Diderot*. Rio de Janeiro: Jorge Zahar.
- Capovilla, F. C., & Raphael, W. D. (2001). *Dicionário Enciclopédico Ilustrado Trilíngue da Língua de Sinais Brasileira (Vols. I e II)*. São Paulo, SP: Edusp, Fapesp, Fundação Vitae, Feneis, Brasil Telecom.
- Dondis, D. A. (2007 [1973]). *Sintaxe da linguagem visual*. São Paulo: Martins Fontes.
- Elias, N. (1998 [1984]). *Sobre o tempo*. Rio de Janeiro: Jorge Zahar.
- Flusser, V. (2010 [1987]). *A escrita: há futuro para a escrita?* São Paulo: Annablume.
- Flusser, V. (2013). *O mundo codificado: por uma filosofia do design e da comunicação*. São Paulo: Cosac Naify.
- Goody, J. (1987). *A lógica da escrita e a organização da sociedade*. Lisboa: Edições 70.
- Keen, A. (2009 [2007]). *O culto de amador: como blogs, MySpace, YouTube e a pirataria digital estão destruindo nossa economia, cultura e valores*. Rio de Janeiro: Jorge Zahar.
- Kerckhove, D. (2009). *A pele da cultura*. São Paulo: Annablume.
- Kress, G. (2010). *Multimodality: a social semiotic approach to contemporary communication*. London, UK: Routledge.
- Kress, G., & Van Leeuwen, T. (1996). *Reading images: The grammar of visual design*. London, UK: Routledge.
- Lebedeff, T. B. (2010). Aprendendo “a ler” com outros olhos: relatos de oficinas de letramento visual com professores surdos. *Cadernos de Educação (UFPEL)*, 36, 175-196.
- Lebedeff, T. B., & Santos N. A. (2014). Objetos de aprendizagem para o ensino de línguas: vídeos de curta-metragem e o ensino de Libras. *Revista Brasileira de Linguística Aplicada*, 14(4), 1073-1094. Doi: <http://dx.doi.org/10.1590/S1984-63982014005000020>
- Leborg, C. (2015 [2004]). *Gramática visual*. São Paulo: Gustavo Gili.
- Lemos, A. (2009). Cibercultura como Território Recombinante. In E. Cazaloto, & E. Trivinho (Org.), *A cibercultura e seu espelho: Campo de conhecimento emergente e nova vivência humana na era da imersão interativa* (pp. 38-46). São Paulo: Itaú Cultural/Abciber.
- Lévy, P. (1993). *As tecnologias da inteligência: o futuro do pensamento na era da informática*. Rio de Janeiro: 34.
- Luckner, J., Bowen, S., & Carter, K. (2001). Visual teaching strategies for students who are deaf or hard of hearing. *Teaching Exceptional Children*, 33(3), 38-44.

- Mateus, S. (2016). Pode uma imagem ser um argumento? *Revista Famecos*, 23(2), 1-17. DOI: <http://dx.doi.org/10.15448/1980-3729.2016.2.21445>
- Mcluhan, M. (2001 [1964]). *Os meios de comunicação como extensões do homem*. São Paulo: Pensamento-Cultrix.
- O'Reilly, T. (2005). *What Is Web 2.0 Design Patterns and Business Models for the Next Generation of Software*. O'Reilly. Recuperado em 26 de agosto de 2018 de <http://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html>
- Rose, G. (2007). *Visual Methodologies: an introduction to the interpretation of Visual Materials*. Londres: SAGE Publications Ltd.
- Santaella, L. (2004). *Navegar no ciberespaço: o perfil cognitivo do leitor imersivo*. 4a ed. São Paulo: Paulus.
- Santaella, L. (2005). *Matrizes da linguagem e pensamento: sonora visual verbal: aplicações na hipermídia*. São Paulo: Iluminuras / FAPESP.
- Santaella, L. (2010). A aprendizagem ubíqua substitui a educação formal? *Revista de Computação e Tecnologia*, 2(1), 1-6.
- Santaella, L. (2012). *Leitura de imagens*. São Paulo: Melhoramentos.
- Simmel, G. (1973 [1902]). A metrópole e a vida mental. In G. Velho (Org.), *O fenômeno urbano* (2a ed., pp. 11-25). Rio de Janeiro: Jorge Zahar.
- Taveira, C. C., & Rosado, L. A. S. (2015). *Produção visual na comunidade surda: prática pedagógica, comunicação e linguagens*. Projeto de pesquisa. Recuperado em 22 de abril 2018 de <https://bit.ly/2D3S6A4>
- Taveira, C. C., & Rosado, L. A. S. (2016). O letramento visual como chave de leitura das práticas pedagógicas e da produção de artefatos no campo da surdez. *Revista Pedagógica* 18(39), 174-195. DOI: <http://dx.doi.org/10.22196/rp.v18i39.3691>
- Thompson, R., & Bowen, C. (2009). *Grammar of the edit*. Burlington, MA: Focal Press.
- Williams R. (1995). *Design para quem não é designer: noções básicas de planejamento visual*. São Paulo: Callis.

Submetido em 4/01/2019

Reformulado em 9/04/2019

Aceito em 10/04/2019

