Proposal of Visual Grammar for the Description and Compositional Analysis of Digital Videos in Sign Languages[1]

Proposta de uma Gramática Visual para Descrição e Análise Composicional de Vídeos Digitais em Línguas de Sinais

Luiz Alexandre da Silva ROSADO[2]
Cristiane Correia TAVEIRA[3]

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.

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.

1 Introduction: media ecologies and the search for a deaf-memory on video

> […] how do we actually know that these great writers (including the author of the Holy Scriptures) would not rather have spoken into a tape recorder or have been filmed? (Flusser, 2002 [1987], p. xxviii).

[1] http://dx.doi.org/10.1590/s1413-65382519000300001

[2] PhD in Education, Instituto Nacional de *Educação de Surdos* (INES) - National Institute for the Education of the Deaf, Department of Higher Education, Rio de Janeiro. alexandre.rosado@gmail.com. Rio de Janeiro/Rio de Janeiro/Brazil. ORCID: https://orcid.org/0000-0002-2702-5617

[3] PhD in Education, *Instituto Nacional de Educação de Surdos* (INES), Department of Higher Education, Rio de Janeiro. cristianecorreiataveira@gmail.com. Rio de Janeiro/Rio de Janeiro/Brazil. ORCID: https://orcid.org/0000-0001-9626-4921

Countless authors, from Marshall Mcluhan (2001 [1964])[4], to Pierre Lévy (1993), Derrick de Kerckhove (2009) and Lúcia Santaella (2010) have argued that new Information and Communication Technologies (ICT), when designed and then appropriate socially, generate new socio-cultural settings and even different modes of thinking and language (Santaella, 2004) as they expand and offer differentiated capacities to our bodies and broaden the perceptions gathered by our senses. It is a conception of technology as prosthesis and expansion of the capacities of the human body.

To these authors, our mind (cognition/perception) and our ambience (houses we inhabit, places we work in, streets in which we circulate) have been altered over the last millennia by successive technological waves. Such changes have accelerated in the last centuries with the insertion of new ICT, represented by their artifacts and their products. This way, we had the millennial technologies of alphabets and handwriting, the fixation of information through ink and papyrus, the invention, in the late Middle Ages, of the printing press by Gutenberg, civil and military experiments with the magic lantern , the ease of writing with the mechanical typewriter, remote communications with the telegraph and the AM and FM radio, the expansion, in the 20th century, of the cinema, television, and more recently the processing of telecommunications through computer and internet.

These technologies form inter-overlapping layers that, to date, are progressive and cumulative in their technical-scientific development, but also change in continuous interlinkages of mutual influence, generating not expected appropriations by those who created them. In the Flusserian view, according to the authors hitherto listed, every manufactured/created object transforms the user's relations with his/her environment in order to take advantage of it. Therefore, when producing an artifact, we also inform and give meaning to the world (Flusser, 2013).

Initial periodizations were proposed to distinguish this path, such as those described by Lévy (1993) and amplified by Santaella (2010), including basically three major phases of ICT throughout our known history: unregistered *tribal orality* (speech and fleeting gesture); *writing*, at first, manual and subsequently printed (accounting, economics, bureaucratic processes and laws), up to the present moment with the mass media; and the unifying *digital format* (*global village*, internet and computing).

These phases synthesize the anthropological and cultural approach of technical artifacts, which become defining pillars of communication periods in the societies that use them. At first, there is no mutual exclusion or absolute substitution between technical artifacts, but rather marked periods, such as the expansion of literate societies based on writing (Goody, 1987) and the printing press from the 15th century (Burke, 2003 [2000]) and its deepening with the expansion of mass culture and the necessary means for the technical reproducibility of the communicational contents in the 19th and 20th centuries (Benjamin, 1987). At the beginning of the 20th century, the growth of the industrial-communication complex caused astonishment and inquiries about the consequences on the human mind and everyday behaviors, like those reported by Simmel (1973 [1902]) when he observed the acceleration of

[4] We decided to put the original dates of the publications in square brackets in the citations and in the bibliography at the end of this paper, so that the reader can accurately locate the decade in which a certain author acted as a theoretical-conceptual timeline.

life and the development of certain systemic indifference among the thousands of inhabitants of the emerging European metropolises.

At the present time, the internet and computing (personal computers, cell phones and tablets) allow the large-scale adoption of digital video production and publication, leading to the emergence of a media ecology of moving image that, unlike the culture of mass media associated with radio, TV and cinema throughout the 20th century, is now fueled by a constellation of small-time new producers in their homes and private studios, youtubers and vloggers, for videos, and bloggers, for texts. In the early 2000s, the liberation of the hub issue (Lemos, 2009), at first, generated the fury of conservatives who attributed this phenomenon to a cultural (anti-intellectual) regression commanded by a legion of amateur producers, who would not have the same strength, reliability and status of media professionals (especially journalists) and academic university researchers (Keen, 2009 [2007]).

Going back a few decades in the 20th century, such changes and transitions derived from ICTs were already perceived and theorized by Vilém Flusser (2002 [1987]), which brings us to the conception of a manual and printed handwriting that, at first, collaborated for the decentralization of non-historical primary orality (in which time is conceived in circles), with the construction of a new collection of memory for humanity through libraries, public archives, museums, leading us to the conception of a linear memory-history (time seen as something that runs along a straight line). Writing, man living in the city-tribe, factory-manufacture, and the notion of a linear and progressive time are closely linked, as well as its materiality present in clocks and in the timing of calendars (Elias 1998 [1984]), which helped to ground our perception of modern-historical-man. Flusser, a radical thinker of the media, then predicts that, from now on, new analogue and digital media will have the power to superimpose, or even replace, various uses hitherto associated with writing, such as the production of scientific knowledge, political discussions, the poetic and even the conceptual philosophical construction.

Would new media have such power? In the deaf community, at least, they seem to play an increasingly central and structuring role, after a long period of predominance of writing. In relation to writing and structuring of memory in Western societies, we realize that such a path is being followed by sign languages[5], which build on the internet video collections aimed at the deaf community and hearing people who use these languages, which we may call a deaf-memory, little dependent on the alphabetic characters of the oral languages in their expression-communication and inscribed on hard disks of the diffuse digital mesh of internet servers. If in the first case the process of maturing of the Gutenberg revolution and its typographic press lasted for at least four centuries (17th to 20th), in the present time we are living a visual revival of deaf visual and gestural communication culture, through an increasing video sharing, throughout a short period of time (2005 -...), from the expansion of what has

[5] To the unfamiliar reader, we inform that there are many sign languages, usually associated with each country: Brazilian Sign Language (LIBRAS), American Sign Language (ASL) and Portuguese Sign Language (LGU) are some examples. Each language has its own set of signs (lexical items) and grammatical principles, distinguishing themselves more or less from one another through their genealogy: LIBRAS, for example, has a strong influence of the French Sign Language (LSF). These languages have a large number of signs already cataloged and dictionalized (Capovilla & Raphael, 2001), but in case they do not have a specific sign for a certain concept, they use what it is called "dactylology", that is, the letter-by-letter spelling of a word of the oral language through signs, using the hands.

been called the Web 2.0 (O'Reilly, 2005), digital cameras in smartphones and tablets and web-digital archives-networks (YouTube, Facebook, Vimeo, among others).

Although this rich collection of videos can be analyzed from the perspective of its expressed content (discourse analysis) or even through the reactions and perceptions of its audience (audience and reception studies), we opted in this paper for the compositional analysis focus (Rose, 2007). Compositional analysis is an approach that takes more into account the elements present in the material itself than the context in which it was produced or the culture of the actors who watch and interpret their meanings. Therefore, it is in this extremely rich and original environment of deaf communities and their productions on the internet that we will try to propose a grammar, not one already known in written texts and widely explored and matured in its structure by linguists, but of productions recorded in the form of videos and that use a visual-gestural form of communication: sign languages.

2 Methodological path

This paper addresses one of the objectives of our research conducted during the triennium 2015-2018 (Taveira & Rosado, 2015), namely: to meet the urgent need to collect, store, analyze, classify and organize the products derived from pedagogical practice, with an emphasis on artifacts and methods combining visual experience, translation and interpretation (Brazilian Sign Language) and communication in virtual and physical environments.

We focused on the analysis of a set of digital video productions, in which Sign Language is central and the deaf are the main target audience. From this analysis, we initially describe and propose the constituent elements of a visual grammar of the videos in sign languages. Our approach was focused on the formats (the form of the binomial form-content) or, more accurately, on the compositions obtained from frames captured from the videos in sign languages. There was, therefore, the rescue and appreciation of the understanding of the constituent elements of visual language, cataloged since the 1950s, 1960s and 1970s by scholars in the fields of Psychology, Art and Design (Leborg, 2015 [2004]; Dondis, 2007 [1973]; Arnheim, 1997 [1954, 1974]).

Historically, the interest in these studies is a direct consequence of the increased presence of painting, drawing and photography throughout the 20th century (visual arts), with the progressive expansion of technical image reproduction in communication vehicles, the cultural industry (Benjamin, 1987), such as newspapers, books and magazines; as well as the intention to extrapolate the false idea that it is enough to have *good eyes* to understand and describe works of art, which would lack a learning and development of that look (visual literacy). In spite of the increasing presence of videos reproduced on TV and in Cinema, the *grammatical* interest of these authors on elements of visuality was much more focused on the static media and on the variations and relations of these elements, which can be adapted to the static frames captured from a video.

Our empirical analysis was based on a sampling of videos produced for didactic and academic-cultural purposes, which included artistic-literal, novelistic and journalistic videos. For this, a standard analysis form was constructed, divided in three parts. The first part identifies *basic production data* (material title, year of production, year of publication, collection title and

length of the video). The second part seeks to reproduce both the *presentation of the material* made by its producers (when existent) and the *synthetic-analytical presentation* written by the member of the research group that watched all the material.

Finally, the third and most important part contains 14 categories of analysis, namely: (1) content type of the material; (2) languages present; (3) provision of languages; (4) reading facilitators; (5) language; (6) genres; (7) type of final material; (8) ambience of application of the artifact; (9) distribution mode; (10) target audience; (11) the age range of the target audience; (12) responsible for the production; (13) accessibility devices; and (14) editing and post-production effects. Each category, in the record, should have the marking of its items justified by the researcher/analyst, composing the defense of the choice and aiming to contribute to the progressive clarification of the meaning of the category among the researchers.

From October 2017 to December 2018, we completed the analysis of 24 videos, all in Brazilian Sign language, except for one of them which was in the Sign Language of Uruguay (LSU). This analytical work generated the respective files filled by the members of the research group "Education, media and deaf community", individually, sometimes in pairs, and then debated collectively in the meetings with the other members of the group at the *Instituto Nacional de Educação de Surdos* (INES) - National Institute of Education of the Deaf, at which time additions are performed and the record is validated and filed. These meetings had, on average, five analyst members.

In this paper, we focus on category number 14, related to the effects of video editing and post-production. This post-studio work, with the popularization of software such as Adobe Première (video editing) and Adobe After Effects (post-production effects), among others, have become fundamental in the creation of multimodal videos in sign languages, in which the composition includes, in addition to the sign language interpreter, photographic images, graphs and diagrams, texts in various typographies and computer animations of all these elements. In a way, these softwares and their multi-layered video processing possibilities condition the visual compositions of most of the videos we have found on YouTube and other online video repositories nowadays, constituting a contemporary version of McLuhan's classic claim (2001 [1964]): "the [post-production] medium is the message".

The initial goal of Category 14 was the highly accurate description of these post-production elements increasingly used by video editors and graphic designers as they have access to those resources on relatively inexpensive computers, and therefore more and more accessible. However, bit by bit, the members of the research group felt a need to also describe the different dispositions, relationships and variations of the basic elements detected in the videos in sign languages, requalifying this category for the compositional analysis of the frames of the videos. After 24 analyzed videos, we perceived the stability reached in our descriptive/perceptive process, which can now be disseminated through this paper.

3 Defining the contours of a visual grammar with focus on the image

> We think differently when we have a language to describe something (Leborg, 2015 [2004], p. 5).

> It often happens that we see and feel certain qualities in a work of art but cannot express them in words. The reason for our failure is not that we use language, but that we have not yet succeeded in casting those perceived qualities into suitable categories. (Arnheim, 1997 [1954, 1974], p. 2). How many see? (Dondis, 2007 [1973], p. 5).

Visuality, and consequent visual experience, is today a fundamental theme in the field of education of the Deaf, especially through the use of sign languages as visual-spatial (non-oral) languages in which body and look use are predominant in communication. At the same time, there is a search for more understanding of the development of school learning through the use of visual resources (Luckner, Bowen & Carter, 2001; Lebedeff, 2010), especially in the video format (Lebedeff & Santos, 2014). Deaf gesturality contrasts with hearing education, largely based on communication through speech and materials centered on the written text. Visuality has a greater power to manifest and be materialized in categories/artifacts such as photographic images, paintings, graphics, infographics, tables, mind maps, drawings, models, sculptures and films in which linear written text, the verbal element based on oral language, is not the predominant center of information organization and content expression.

In spite of this watertight division presented here, on the one hand, we agree with the triadic conception of language matrices (sound-visual-verbal) proposed by Santaella (2005) in which verbal can also manifest itself in the categories of the first group, as well as the visual also sometimes is basilar in written texts, more evident in genres such as the poetic and the fictional novel, genres in which we form more or less detailed images (representations) in our minds as we go through the text. On the other hand, we are also aware that an image can contain an argument (Mateus, 2016), that is, it presents a set of premises and a central thesis, typical structures of the written text and, before it, of the orality, when the rhetoric flourished in ancient Greece. Such mutual contributions between categories once so markedly different may be the result of the exponential growth of the so-called multimodal "texts" (Kress & Van Leeuwen, 1996; Kress, 2010), that is, in a single composition aimed at a specific communicational purpose at the same time, written texts, images, videos and sounds, such as the multimedia of the software and the hypermedia of the pages that we access on the internet, are gathered, making the so-called "literacy" more complex (a *multiliteracy*?).

Our approach here, however, will be more focused. The analysis of artifacts, experiences and didactics aimed at the teaching of deaf students was previously explored by us (Taveira & Rosado, 2016), and this paper now provides a brief description of the meaning of what we call visual literacy and the understanding of fundamental elements of visuality compiled under the expression visual grammar. From Santaella's point of view (2012), visual literacy means learning to read images, developing the observation of its constitutive aspects and features, detecting what is produced within the image itself. That is, it means acquiring the corresponding knowledge and developing the necessary sensitivity to know how the images are presented, how they indicate what they want to indicate, what their context of reference is, what the images mean, how they think, what their modes are to represent reality.

To the author, and also to us, the visuality is learned, that is, it is not innate, it is not natural or even less spontaneous, and for this it is necessary to be systematized and conceived in a language that enables the observer to dismember an image in parts, in order to decode

it in its entirety. The author cites aspects and constitutive features of the images, which once understood make subjects skilled in reading images. It is in authors related to the Arts and Visual Communication (Design) such as Arnheim (1997), Dondis (2007) and Leborg (2015) that we find, objectively, these described features.

To be visually literate, then, is to learn and exercise in everyday observation what Leborg called visual grammar, a set of elements and activities/relationships between these elements that we could distinguish when analyzing the visual compositions, *reading them* and, with that, making us visually literate. To Leborg, the learning of these elements in everyday life is more related to the physical experience of the subject, who often had no access to a verbal language and its consequent systematization, a gap that he seeks to fill in his book.

Although Leborg, and his predecessors Dondis and Arnheim, are fundamental references of our study, the teaching of these elements is also present in manuals related to the practice of Design and Graphic Arts (analog or digital), as in Williams (1995), who reduces this art to four basic principles: (1) proximity, (2) alignment, (3) repetition, and (4) contrast. Leborg (2015), as we see in the images in Figure 1 below, extracted from the summary of his book, expands the vocabulary of visuality so that we can accurately detail the visual compositions that we experience sensorially in our daily life.

Figure 1. Synthesis of the four major groups proposed by Leborg in his visual grammar.
Source: Leborg (2015, p. 96-97).

We have noted how Leborg synthesized categories of image analysis inspired by previous categorizations (Arnheim, 1997; Dondis, 2007), with the more basic ones related to the *abstract* and the *concrete* and the actions/relations of/among the basilar elements would lead us to the *activities* and *relations*. We do not intend here to describe them one by one, but to present the reader with our starting point. It was with this same conception of separating the basic elements and relating them to combinatorial patterns that we focus on the construction of a visual grammar to be used in the analysis of videos in sign languages.

4 PROPOSAL OF A VISUAL GRAMMAR FOR DIGITAL VIDEOS IN SIGN LANGUAGES

We sought, from the conception of visual grammar developed by Donis A. Dondis (2007) for static images, rescued in a refined synthetic re-reading through visual representations by Christian Leborg (2015), the basic constituent elements of these videos, which are basically frames in sequence generating the temporal perception of movement.

We then proposed seven basic elements for the compositions of each frame (Figure 2), namely: (1) the sign language actor/interpreter, that is, the person who uses sign language to express him/herself; (2) the actor/interpreter using oral language, that is, the person speaking to express him/herself; (3) the textual mass, in the form of titles and descriptive texts in alphabetic writing; (4) illustration/picture, graphic or photograph; (5) the caption in the oral language written in alphabetical order; (6) the natural setting or artificial background inserted through substitution by the Chroma key technique; and (7) the smaller video over the main video, Picture-in-Picture or PIP.

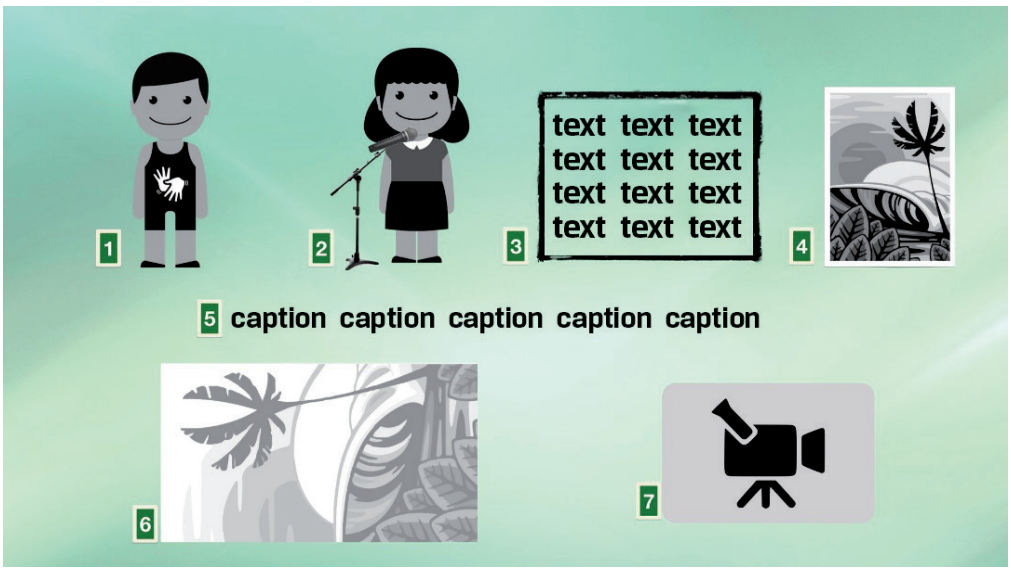


Figure 2. Proposal of seven basic elements for sign language videos.
Source: Elaborated by the authors.

These seven elements listed are a proposal to define basic units of composition of a video in sign language, being analyzed by means of the capture of a frame of the video under analysis. With these basilar elements, we can create an iconic representation of the video, abstracting and reducing the video to its basic elements with simple graphical representations positioned proportionally on the screen, which assists in the detection of their most typical relations and variations.

From these elements, we exemplify, below, some of these exercises of application in frames of videos that composed our experience up to that moment. It should be noted that on the right we find the screens captured with the identification, through numbers, of the elements present in the frame and, on the left, the iconic representation of the original frame (Figures 3 to 6).

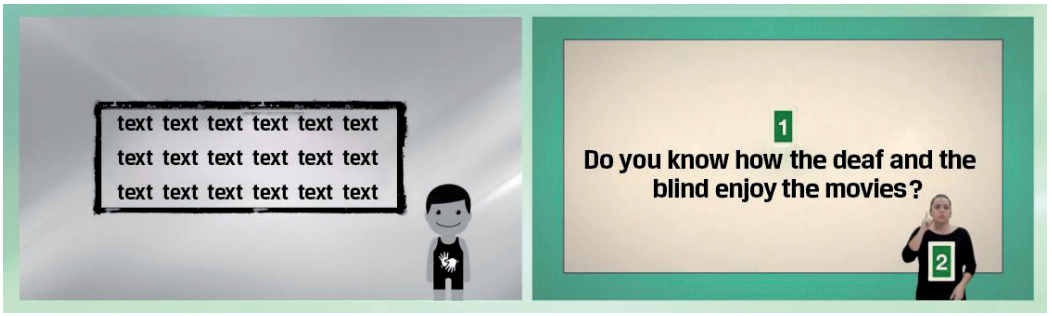


Figure 3. Exercise with representation through two basic elements.
Source: Elaborated by the authors.

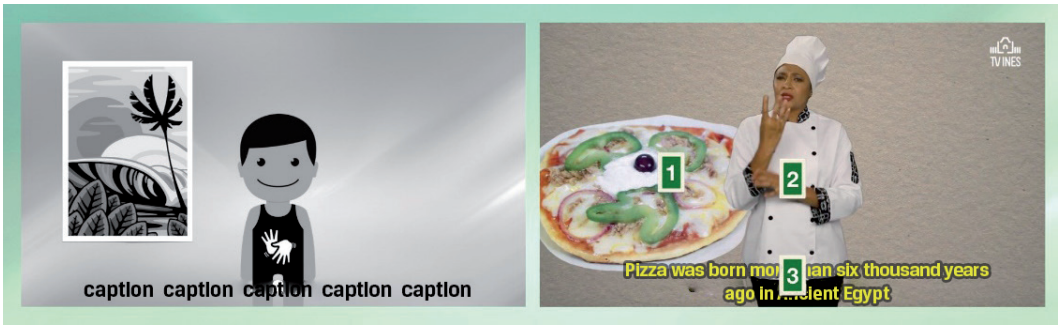


Figure 4. Exercise with representation through three basic elements.
Source: Elaborated by the authors.

Figure 5. Exercise with representation through four basic elements.
Source: Elaborated by the authors.

Figure 6. Exercise with representation through five basic elements.
Source: Elaborated by the authors.

As we were observing and analyzing the component videos of our collection (total of 24 to this date), we detected and cataloged some types of variations of the basic elements in the compositions, arriving at five great typologies: (1) size (on some occasions the elements are larger and, in others, smaller); (2) cut: one or more elements are either fully or partially displayed; (3) position on the screen: stand out in different areas of the composition (center, upper and lower edges, left and right sides); (4) groups: when they are repeated sequentially and are close to each other, forming clusters; (5) format and spacing: when they vary in their shape in different frames or in the distance between the elements.

Variations of the basic elements can only be detected and defined when a set of frames, sequential or not in the same video or of videos of varied origins, can be collated together, followed by observations and records. Therefore, it is a form of temporal description of the video or a set of videos.

Specifically, the variations of size and cut of the elements of a video relate directly to what is called camera shots or framing, in the cinematic and television language already comprehensively detailed in manuals (Thompson & Bowen, 2009): long shot, medium shot, medium-wide shot, medium close-up, close-up and extreme close-up. A close-up shot, for example, would have a predominant frame element with varying size and cut to the point of filling the frame completely.

We thought that descriptive terms of camera shots are more related to the description of people and objects present in a scenario - typical elements of novelistic, documentary and fictional film and TV narrative videos - but not to elements added artificially after the footage obtained with the use of a camera camcorder, such as text massages, graphics, photographs and images, present in most of the videos for educational purposes that we have cataloged.

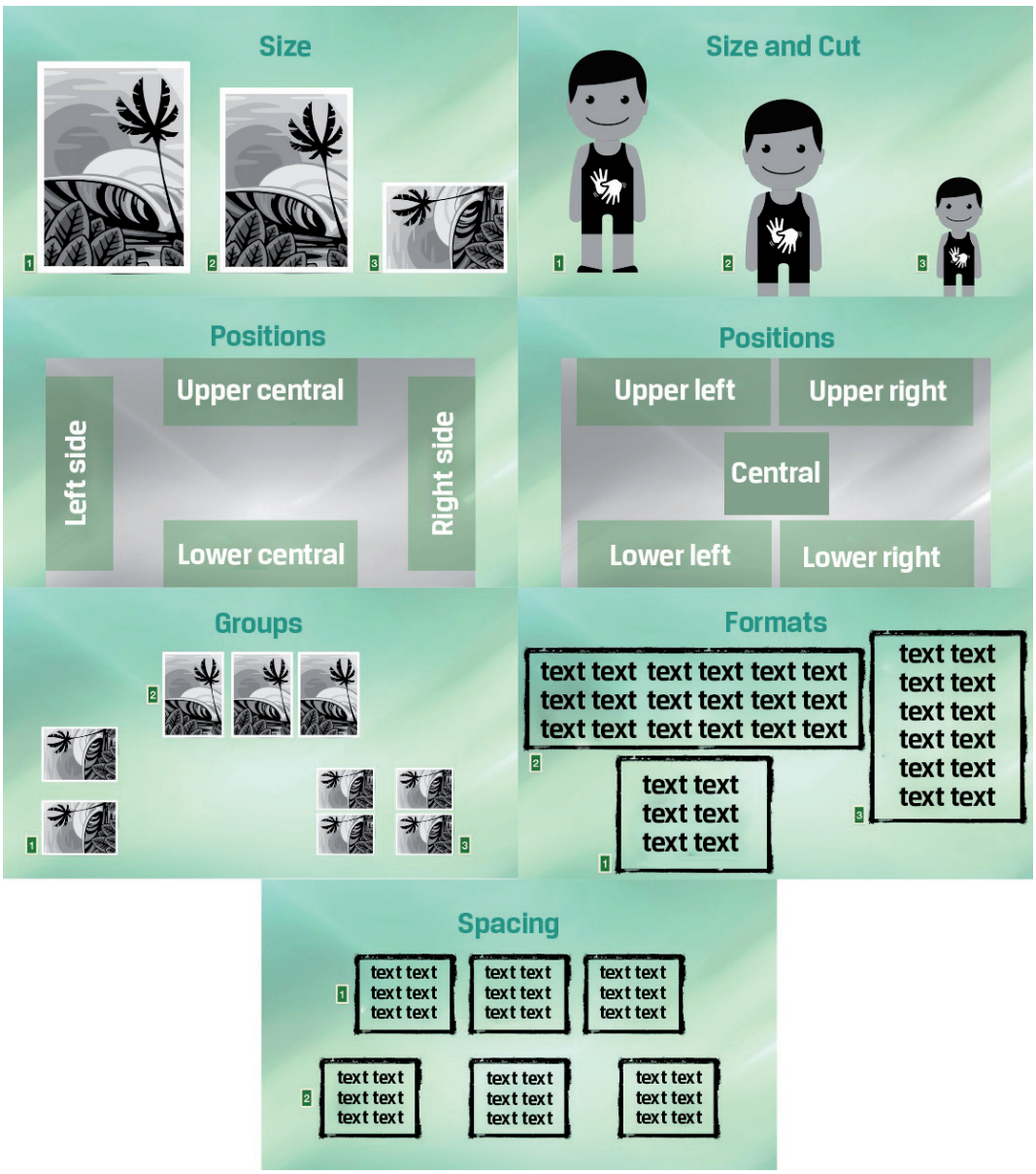


Figure 7. Catalog of possible variations of the basic elements and their exemplification.
Source: Elaborated by the authors.

We then decided to create a set of more comprehensive categories that would account for the multimodality allowed by the video and post-production editors present on personal computers today and widely used in professional, semi-professional and amateur environments. In a certain way, we propose here a descriptive language in which the analytical center moved from the camera (*raw* capture process) to the video editor and effects (compositional elements mixing and *artificial* finalization).

In addition to the variations of the basic elements, we also propose some relationships developed between the basic elements of a composition, inspired by Leborg, which cataloged a countless number of them, some of which we also highlight, as found in the analyzed videos:

1. Repetition: when an element is repeated several times in the composition or throughout the video and allows the detection of continuity of a given pattern, such as texts that use the same font and size by marking chapter titles, the transition cuts used whenever a specific situation takes place, the scenarios with colors and textures repeated at different times, or the use of sign language/voice actors wearing the same clothing/type/color and appear in the same area and of the same size at different times.

2. Symmetry: when there is a balance of element stain on both sides of the composition, left and right, or between the top and the base, generating comfort for the video viewer, who perceives them as distributed in a harmonic way, following the example of the sign language actor on one side and the illustrative images on the other.

3. Asymmetry: it is the unbalance in the composition, when the elements are grouped and "weigh" on one side or in one of the regions of the frame, which directs the reader's attention to one of the poles intended by the video producer, a feature to emphasize some elements and provide less importance to others.

4. Enlargement/reduction: when an element of the same nature is represented in different sizes at the same time or along the video, following the example of the sign language interpreter who, free of the window that previously restricted him/her in the corner of the screen, can now be enlarged in moments of greater emphasis on sign language, or an image that can have an enlarged detail to generate a more precise and emphatic reading by the viewer.

5. Attraction/proximity: they are groups of clustered elements that attract and indicate some relation between them, helping the video viewer to relate and categorize elements within the composition, as in the case of the sign language captions that always appear close to the oralizing actor in the film.

6. Weight: to which side of the composition the elements are weighing more, as a gravitational force felt by the viewer, also collaborating for the emphasis on a particular part of the frame, being a contrasting experience between *fuller* and *more empty* areas, directing the look of the video viewer, in general, to the *heavier* area.

7. Quantity/predominance: the area of the composition in which the elements predominate, provoking the perception of a pigmented field with features, greater number of elements and larger reading boxes containing texts, illustrations, actors, among others.

8. Space: related to the empty areas and the dense areas that stand out in the composition, being the empty areas necessary for the visual reader to *breathe*, avoiding the competition of elements that can lead to an overactivity of the screen and the loss of content/information by the overstimulation of perception, something that occurs in sign language videos that abuse the use of elements of emphasis and prominence within the same frame.
9. Overlapping: when an element is positioned over another element, such as scientific-academic works in sign language with citation and footer overlapping and in smaller size in relation to the sign language actor who is *paused* and occupying the entire frame.

Relationships between basilar elements can only be defined and detected within the same video (unit) and in many cases with the capture of a single frame of this video.

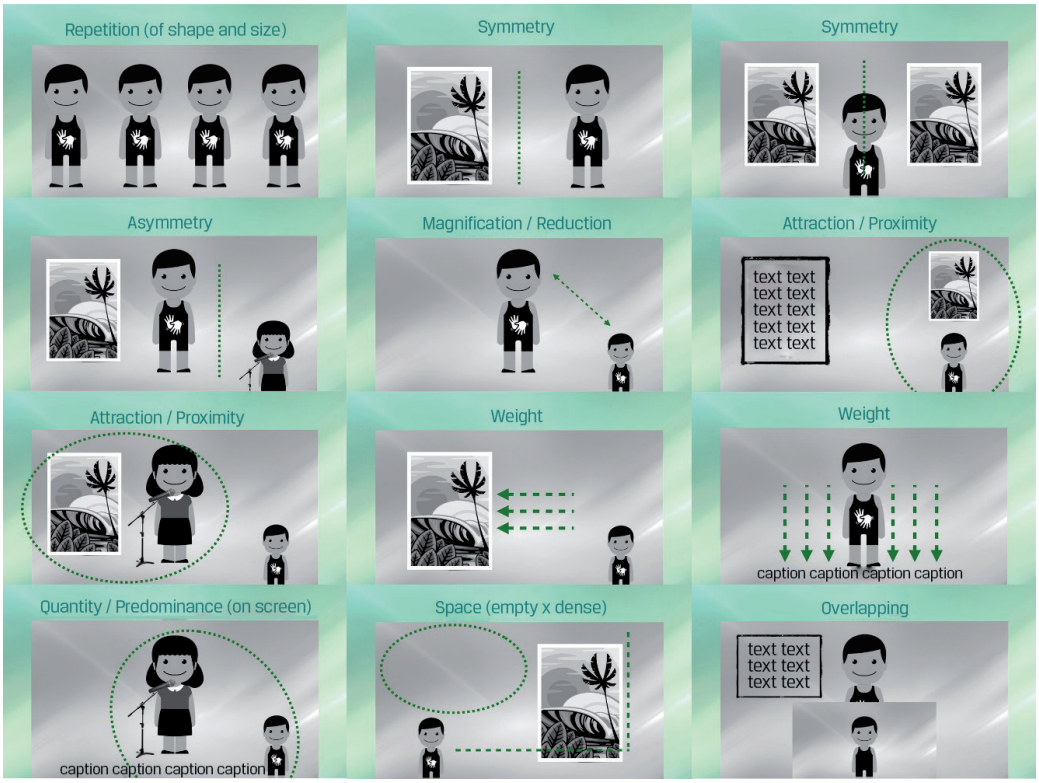


Figure 8. Catalog of possible relationships between the basic elements.
Source: Elaborated by the authors.

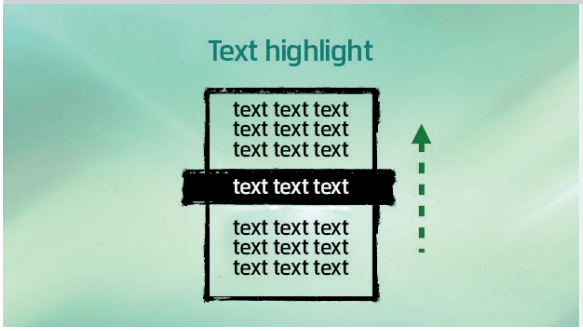
Until December 2018, we cataloged 187 combinations involving the seven basic elements in a video frame composition. These are combinations that range from one element in the frame to seven elements simultaneously in the frame. The research group created a PPT (Microsoft PowerPoint) file with the seven basic elements, which allows, from each watched video, the construction of the iconic representation containing the combinations between the

basic elements detected during the analysis, adding this representation to the analysis sheet. This construction is only necessary when the 187 cataloged combinations are insufficient to represent some frame of the analyzed video.

It is important to emphasize that from this set of combinations cataloged from our sample videos, we can also infer possible combinations not yet found, but perfectly feasible to be discovered or even created. The potential of the visual grammar of the videos in sign language is not only descriptive/representative of what has already been produced, but also prospective/creative, allowing the visual representation of compositional possibilities of future works (virtualities) for artists, designers, video editors and graphic designers.

From this initial proposal of a visual grammar of the videos in sign language, it is possible to approach and differentiate visual solutions that are being proposed, tested and presented in numerous videos currently circulating on the internet from public institutions, private companies, non-governmental organizations and independent producers. The catalog of original, specific, often unique solutions found by a given video producer is also possible and, once formally cataloged and disseminated in a systematized manner, can be used in other productions. We will briefly present three examples that make up the catalog currently under development:

In a novel in Sign Language of Uruguay (SLU), the actors present on the scene are translated into Spanish by means of captions in which each color represents one of the characters present in the scene.



In a video in Libras (Brazilian Sign Language) translating *cordel* literature, the original text in Portuguese Language scrolls across the screen from bottom to top and has as a highlighted area, the one that is being signaled by the interpreter/actor at that moment.

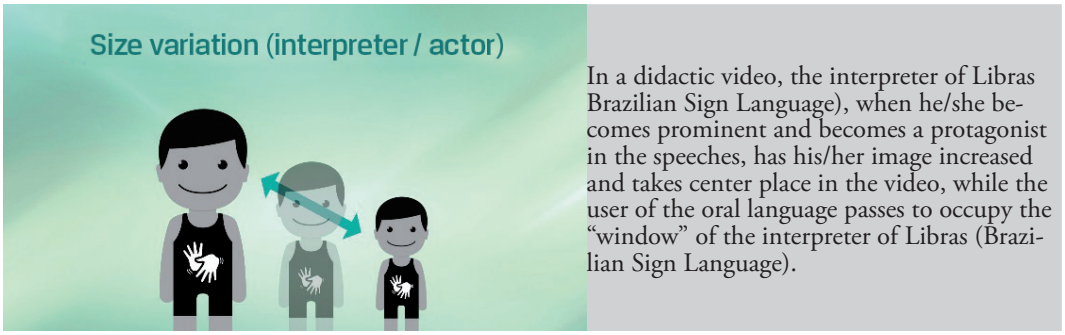


Figure 9. Three examples of the visual solutions catalog currently under construction.
Source: Elaborated by the authors.

As we have already mentioned, these are apparently simple solutions, but rarely or never used in other sign language videos. Once assembled, disseminated and consciously anticipated in the form of a catalog of visual solutions, publishers, designers, artists and video producers will be able to apply them, and expand the reach of new materials and, consequently, visual comfort for the deaf target audience.

5 Final considerations: a *vocabulary of visuality* under construction

We are at a unique time in our history, in which digital technologies connected in networks of global reach enable the creation and the sharing of digital videos on a large scale through websites-repositories. This scenario, with the *global village* proposed by McLuhan (2001 [1964]), in the 1960s and partially realized in the 2000s, has become extremely conducive to the creation of video collections in sign languages, forming a *deaf-memory* from these new material conditions synthesized in the form of computers, cameras and digital video editors.

In this paper, we presented a proposal of mapping and coding the composition of these videos, offering the bases for the construction of a visual grammar that supports the producers of videos in sign languages, to previously schematize their conceptions in iconic representations. We presented some examples of items from a catalog of better practices found in this emerging environment of community, historical and identity construction, so that these same producers do not start from scratch in their creations, but rather get inspired by solutions already proposed, tested and used at an earlier moment.

So far we have not found, in the national and international literature, a proposal similar to the one presented here[6], with the compositional analysis and the proposal of a visual grammar totally focused on the videos in sign languages. Donis Dondis and Christian Leborg, as well as Gunther Kress, focused on universal analyzes of image and photography, used by us as a starting point, but adapted to the set of productions by/for the deaf community.

[6] The national and international bibliographic research was carried out through the access to the site of the Division of Libraries and Documentation of the Pontifical Catholic University of Rio de Janeiro (PUC-Rio), throughout the years of 2017 and 2018, with the expressions "Visual Grammar", "Deaf", "Sign language" , "Deaf education" and its derivations and combinations. The comprehensive set of databases signed by PUC-Rio is specified at the following address http://testaremoto.dbd.puc-rio.br/sitenovo/base-de-dados.html#bases_assinadas.

As researchers in the field of deafness, we are sensorially experimenting with several videos produced by/for/in the deaf community, especially Brazilian, assigning us the task of decoding and turning the multiple layers present in these visual compositions into something readable. Only from the formalization of the bases of a visual grammar to the videos in sign languages that we have seen our agility in detecting and representing these elements and their mixes increased, in a practical and almost automatic way, with simple and at the same time generalizing schemes.

The non-expression of the image through words, a difficulty felt by Arnheim and which later instigated Dondis and Leborg to formalize a *vocabulary of visuality*, gave us an opportunity to realize the passage from a synthetic visual language, global but dispersive to the eyes, to the conceptual verbalization of what we experienced sensorially. In this exercise of sensitive research, we produce descriptive categories and necessary visual representations, presented here in the form of linear text and their respective iconic images, a form of an imperative organization for the proposal of visual grammar.

References

Arnheim, R. (1997 [1954, 1974]). *Art and visual perception: a Psychology of the creative eye.* Berkeley, Los Angeles, London: University of California Press.

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* (3rd 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. (2002 [1987]). *Writings*. Minneapolis, London: Universtity of Minesota Press.

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. Cazeloto, & 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. Retrieved on August 26, 2018 from 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*. 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* (2nd 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.* Research project. Retrieved on April 22, 2018 from 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.

Submitted on 4/01/2019

Reformulated on 9/04/2019

Accepted on 10/04/2019