Comparatismi 4 2019
ISSN 2531-7547
http://dx.doi.org/10.14672/20191596

Grammatizing the Visible

Denitza Nedkova

Abstract • The Visual Narrative Grammar Theory counting the Parallel Architecture linguistic model—recently traced, the first one, by the cognitive and comics theorist Neil Cohn and, the second, by the linguist and cognitive scientist Ray Jackendoff—empirically evidence the same congenital but context modulated organizational cognitive structures founding the grammatical structure of verbal and visual languages. The identification of the visual morphology and syntax allows, then, the experimental application of the aforementioned grammar in every field of the image narration. Hence emerges the hypothesis of a dimensional counterpoint between the two communicative channels. The two-dimensional chronologically linear verbal expressions' phrase construction is compared with the three-dimensional visual representation construct based on irregular and diachronic combination of the images. Analyzing how the visual symbols' combination arises and to what degree it's similar to that of verbal languages, the present study suggests that the simultaneous perception of the visual grammar elements allows an immediate, but often partial, reading of the image narration.

Keywords • Visual Narrative Grammar; Image Morphology; Visual Syntax; Morphs; Iconic Text Storytelling

Ledizioni

Grammatizing the Visible

Denitza Nedkova

Notoriously the visual symbols, beyond the simple perception of the image, allow the sense construction. The latter occurs through the combination and ordering of the former. Consequently, it is possible to trace an analogy between the images and the verbal lexemes that form complete and meaningful messages (phrases) constructed in right combinations. Still the linear words ordering of the spoken language sequences contrast the simultaneous and parallel spatial and temporal structure of the visual ones. The hypothesis of a dimensional counterpoint between the two communicative systems could be ventured. On one side merge the two-dimensional string or verbal expressions' phrase construction, i.e. chronological arrangements of concepts representing morphemes and, on the other, the three-dimensional construct of the visual representation with the interferential interactions between the plural possible meanings generated by irregular and diachronic combination of the images.

This dichotomy causes the main queries of the present paper: How does the visual symbols' concurrence occur and to what degree is it similar to that of verbal languages? Is it possible to formulate a "theory of sequential image comprehension"?

1. Visual Language Theory

Inspired by Theory of Visual Language—born by the empirical and experimental research of the cognitivist Neil Cohn in seeking to understand the comics' communication mechanism—here is underlined the parallel between the cognitive structures that govern verbal and the visual languages. In confirm the threefold structure of the relationship between language and thought—signifying (word), meaning (mental imprinting of the referent), referent (an extra-linguistic reader body)—is stabile even with the switch of the first position from word to image. Therefore, every language, verbal or visual, is a system of conventional signifiers (phonemes, graphemes etc.) grammatically structured to form a rational message and depending on the cultural context that provides particularities and differences to distinguish it. The theory of common language grammatical rules (Cohn, *The visual narrative reader* 180) based on the same organizational cognitive structures - congenital but permanently modulated by the context—that govern all the communicative expressions (musical one included, Jackendoff and Audrin 470), opens the experimentation of the aforementioned grammar in every image narrative field. All possible because, contrary to the Chomsky's generative grammar, meaning is not grammar derived, but is a parallel and separate structure, whose organization rules are memory stored as complete archetype cognitive schemes.

Graphic abilities, as the discursive ones, are structured according to function, form and development. They involve, like the talking, motor skills, interaction with the graphic practice of a given culture, socialization and motivation, emotional intelligence, etc. The image, equally to the word, gives the psychosomatic capacity to express both the live (view-based depiction / description) and memory (object-based depiction / description) perception. Accepting the stability of Marr's visual perception tripartite division of mental

image processing dated 1978·*early vision* sketches the basic image elements (color, shape, movement) (Chatterjee 1573) activating the V1 zone of the primary visual cortex; *intermediate vision* adds spatial image elements (position, depth, orientation) activating the secondary extra striate visual cortex (V2); *late vision* provides three dimensional representation integrating new and existing knowledge through the motor-sensitive mirror neurons of the tertiary visual cortex (V3, V4,V5, V6)—the second level agrees with the view-based representation as recalls the visa emulation, while the third with the object-based one conceptualizing the perceived data. These two levels seem to found the neural organization structures that hold up every communicative construction facilitating the evidence of resemblance and contrast through visual and verbal cognitive stamps.

The meeting point between the two communication systems, as previously highlighted, is the basic comprehension (lexicon) system with variable and individual number of cognitive meaning making schemes (concepts) expressed by image (drawing) or by systematic sounds and symbols' organization (spoken language). The cultural pre-establishment of the stamps' blend for each communication system conserves and conventionalizes them over time by the collective memory of each culture. This concurred process (culturally determined) allows distinguishing Greek art from the Maya or Japanese one exactly as it happens with their respective verbal languages. Consequently, there are pre-existing representative schemes of the objects in the mind and not just a direct perception of it at the base of the drawing, as well as of the language. Therefore the cognitive image system presupposes the availability of a graphic lexicon constructed from the pre-determined mental resemblances of the discerned object and organized by size—first the simple graphic elements (lines, dots, etc. equal to the letters), after schematized parts of the image (equal to the syllabation) and finally simple representations of the image (equal to the words) - and managed by a graphic syntax.

Accordingly, the main diversity between verbal and graphic cognitive models stays in the iconic nature of the picture. It presupposes that the graphic images are superimposed on the relative phrases as they are formally similar. Contrary the language signs (graphemes) are causally linked to the meanings—not having a visual similarity with them—which determines their purely symbolic nature. Thus, the drawing is more comprehensively universal, the temporal linearity of the spoken language is missing or in any case not imposed in the comprehension of the picture language even if both verbal and image languages maintain a high cultural-imposed conventionality. If in the spoken language it is the syntax to organize the meaning in the drawn one it organizes only graphic elements similarly to the phonology function that combines sounds to call and therefore to mean something. These different forms of linguistic structure and the relationship between the linguistic levels (phonetic, morphological, and syntactic) are balanced in the broader linguistic framework of Parallel Architecture (Jackendoff and Audrin 483). Concretely, the visual language implies an interaction between the graphic structure (lines and shapes are the phonemes) the morphological structure (linking of the phonemes) and the semantic structure (realization of a meaning).

2. Image morphology

Visual morphology treats the shapes of figurative elements, their formation and modulation in a system. The division of the language lexical units (single word, part of word or sequence of words) in open/variable elements (nouns, verbs, adjectives, adverbs) and closed/invariable elements (prepositions, articles, conjunctions, pronouns) is hardly applicable (Cohn, *Combinatorial Morphology in Visual Languages* 183) to the image

lexicon. The definition of a minimum unit in the image is useless, as the visual information allows a simultaneous, and non-linear as in speech, perception of the morphological elements. Therefore, the autonomous visual units are not morphemes—as the parallel architecture would have established—but mono, micro or macro *morphs*, term crystallized by Neil Cohn following the semiologist Walter Koch's interpretation from 1971 as *logemes*. Formally the morphs are represented by complete forms such as figures, objects or parts of them. These are connected in a morphology system through non-autonomous units such as affixes formally expressed by motion directions, color definition, dimensions etc. of the forms and aggregated both to open and close units. Accordingly, the monomorphs are composed of micromorphs which in turn can separate from the former and represent their own meaning such as synecdoche in verbal language. The macromorphs then are formed by several monomorphs as to form the phrases in the sentence.

This connection of the parts is exemplified in Cohn's visual grammar theory by the commix frame where the shapes—objects or persons—are matched with the insertion of affixes—in the word these are suffixes, prefixes etc.—that could be upfix like the speech bubble above the head of the speaking character, a circumflix when it comes to lines of action around an object that moves etc. The affixes can substitute completely the morphs similarly to the irregular verbs (eat → ate). The affix of the dust cloud, for example, replaces the macromorph of the fighting people representing an external substitution, while the internal one—as the metaphonesi (fall → fell)—is graphically expressed from hearts instead of eyes to indicate falling in love.

The analogy that occurs in the combinatorial strategies of the spoken and the picture languages is confirmed by the cerebral functioning almost identical in the comprehension of morphological variations both on a graphic and written level. The electroencephalographic measurement of the electrophysiological brain response to an external stimulus (ERP, event-related potential) in cases of semantic or structure violation—on morphological as well as on syntax level—registers the same neural behavior both for a visual (Cohn and Mahler 78) and verbal (Kuperberg 176) narration. A negative electrophysiological peak in the frontal lobe occurs—the cerebral wave of the event-related potential drops for 400ms (N400)—in case of semantic incongruity, and a positive electrophysiological peak in the posterior left hemisphere rises—the ERP rises for 600ms (P600)—in the case of the syntactic one. N400 is a context specific default brain response, which indicates the extraction of semantic information from the perceived stimulus, namely the overlap of information currently received with the pre-existing mental models. Therefore, the more the stimulus is incongruous to the milieu the more the ERP potential falls. Instead, P600 reacts to situational changes, i.e. violations of the positioning and grouping of the informative components, with increasing its activity of modifying and updating the existing sequential mental model.

Relating the visual morphology to a real (photograph) or artistic (picture) representation emerge other formal elements proper to the image lexicon. Observing a masterpiece of Italian Renaissance art (Fig. 1), alongside with the conjunction graphic affixes—such as the drapery that wraps the figure or the clods that sustain them—the lightening and the chromatics of the monomorphs transform the simple physical relation—so well achieved by the action lines in the comics—in a dramatic event. Color affixes, for example, link the monomorphs in macromorphs i.e. phrases—note the dominant flesh color of the nudes—or hyperbolize one monomorph among the others—as the intense blue color of the Virgin's vest. Moreover, as in the comic strip a synecdoche or metonymy increase the dynamics of the narration, in this masterpiece it is entrust to the allegory with the attribution of a hidden meaning to images, beyond their apparent literal representation. In specific, the clusters of

corps become the waves of the vortex—namely the action lines—activated by Christ's gesture.

Fig. 1 – Michelangelo Buonarroti, *Last Judgment* (1536-41); detail: Christ Judge; Sistine Chapel, Vatican City.

3. Visual Syntax

The evidenced visual morphology is, by its nature, organized more or less sequentially in a narrative grammar through the syntax structures. The creation of a narration through images is analogous to the verbal one because of the same memory encoding of the systemic relations between forms and meaning, as by aforementioned empirical results. In Cohn's theory a narrative arc structure of cinematic origins is applied to identify a series of hierarchical narrative constituents (Cohn et al. 67) of the comics' storytelling, starting with these that establish and initiate the action/event and finishing with those that climax a release it. These syntactic elements form three types of narrative schemes: the regular one, the amplified regular one with additions to each constituent and fragmentation of the visual image (called conjunction) and narrative scheme with modifiers that focus the attention on a particular element of the image by attention windows. When the constituent is fragmented into several panels showing several points of view of the same event, different elements of the same semantic representation or an iterative progression of the event's completion, it needs conjunctions to maintain the integrity of the sequential flow. In these cases, the understanding of the sequence requires an inferential (deductive) neural process of updating that refers the detail image to the larger context of the whole sequence. The reader perceives this common milieu and understand the visual narration using pre-existing combinatorial structures separated from the perceptual updating of the panels (Cohn, *The Visual Language of Comics: Introduction to the Structure and Cognition of Sequential Images* 66) and represented by the conjunctions use in maintaining the narrative coherence. Actually the cerebral reaction, both of structural and sense correction to the same element of conjunction—namely high left anterior negativity (N400) in the presence of conjunctions both in semantically congruous and incongruous sequences and high left posterior positivity (P600) in both correct or incorrect syntactically sequences with conjunctions—in graphic (Manfredi, Cohn and Kutas 33) as well as in linguistic (Friederici

1357) and sound (Koelsch and Siebel 578) fields, highlights the share of the same conjunction mechanism and its presence both in terms of sense and structure in all communication systems. This is therefore a stable element of a unique narrative understanding system that underlies all the communication structures.

In addition, the left posterior positivity (P600) in the manga readers results lower as they more easily recognize the spatial connectors, opposed to their higher left anterior (N400) negativity left as they engage in more intense combinatorial experience-based processes. Contrary those who do not routinely read comics have a higher P600 due to a more sophisticated update of the mental models and a lower N400 as the combinatorial neural energy is in who is not used to such action manages a less extensive lexicon. These results explain the difference in the descriptive character—defined by Cohn expressive relativity (*The visual narrative reader* 323)—of the visual narratives of different cultures. Therefore, a Japanese, using a more descriptive spoken language, creates and reads without difficulty a similar visual language, unlike the American that, expressing himself synthetically, needs to find in a single or just few frames the entire narrative arc, as evidenced by its difficulty in combining multiple frames of the same action. This phenomenon is due not only to the difference in the verbal language typology but also to the different training of the visual perception as evidenced by the diverse use of the four main types of attention windows in the narrative strip (Cohn, *The visual narrative reader* 335). The *macro* attention window contains most or the entire information of the action, the *mono* window has only one element that, through the *micro* window is emphasized by enlarging certain parts of it, and finally the *amorphous* window containing elements not directly related to the action but rather to its contextualization. Westerners tend to focus on the leading character in the scene while the Orientals pay the same attention both to the protagonists and to the surrounding environment (Masuda et al. 1262). So, the Japanese manga simulates the whole process that the eye performs in the attention, posing consecutively on the characters (mono), on their details (micro) and on the context (amorphous). Western comics, instead, tend to offer immediately the integral scene (macro) as the readers easily capture the important traits, attracted only by the protagonists of the action.

Fig. 2 – Michelangelo Buonarroti, *Last Judgment* (1536-41); detail: Christ Judge; Sistine Chapel, Vatican City.

4. Universality of the visual language grammar

Despite the cultural specificities the unique visual narrative feature of the simultaneous perceiving and consequently reading of the information often leads to cultural interference and overlap. Returning to the Sistine Chapel, let's consider entire fresco narration of the Last Judgment (Fig. 2).

This iconic text is not apparently divided in separate frames—crucial for Cohn's Visual Grammar Theory—as is the vault of the same chapel frescoed by Michelangelo just few years earlier (Fig. 3) representing one of the most famous precursors of the comics strip.

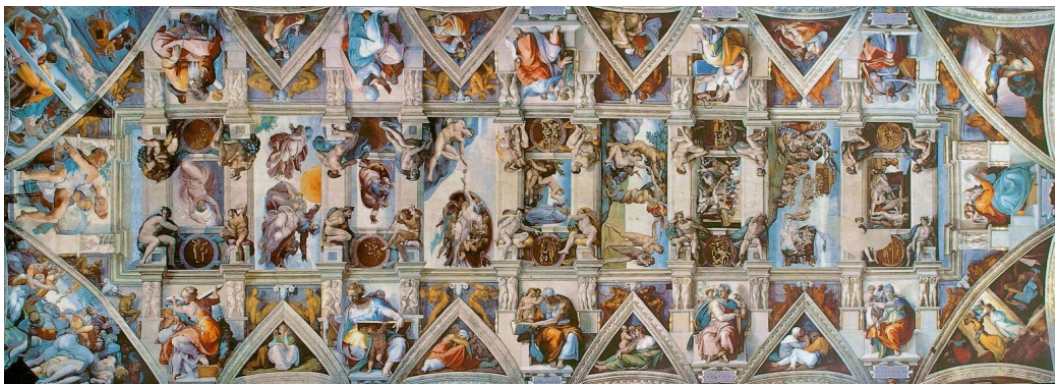

Fig. 3 – Michelangelo Buonarroti, Sistine Chapel vault (1508-12), Vatican City.

This progressive removal from the regular representational layouts—with the consequent linear reading of the content—towards structurally open and more random solutions is explained, century after, by the empirical analysis of the visual text. The phenomenon, in fact, is traced by Neil Cohn (Cohn and Campbell 197) examining the development in the last 70 years of the comic strip layouts and identifying a gradual shift from the Z-path conventional structures, namely from left to right and from top to bottom (the 40s and 50s of the 19th century) to most linguistically i.e. culturally free and productive ones—such to allow the creation of comics writers' personal style and the emergence of individual readers' storytelling. In fact, the authors initially feel obliged to follow the simple linguistic writing and reading structure (Z-path) in order not to confuse the readers in the narration following. Subsequently, however, the creative unconscious intuition that the navigation rules are systematic and implicit, namely the stability of the reading (verbal) order preference even in structurally unstable texts, allows to explore less regular decorative elements. The systematic application of the latter confirms the artists' treat of the page as a canvas regardless to compel the reader moving through the page in an intuitive way, obeying only to the embodied navigation principles. Once again, as in the Renaissance, the page is transformed from a simple base for a panels' flow (Molotiu 88) into a canvas to paint on (Groensteen 110) that can accommodate both the flow of narrative sequences—as the Sistine Chapel vault—and a unique widescreen—as the Last Judgment.

Indeed, the story fragmentation of the Michelangelo's Judgment in diverse scenes is accomplished by the vertical macromorphs' stratification straight from Environment - note the contrasting blue background that absorbs all the episodes—and Time—all actions refer to the same outcome - Conjunctions (Cohn, *The visual narrative reader* 326). The layers arrangement allows the definition of the constituents in a narrative arc that syntactically is far from the horizontal (linear verbal) narrative scheme—establisher/climax/release—which assumes the typical for the western languages (graphic and verbal) Z-path reading. Differently this fresco syntax develops an irregular narrative scheme structuring the constituents not sequentially but circularly around the central Peak constituent formally represented by the monomorph of Christ Judge. Even if each episode/phrase peripheral to the central scene can be used as an initial or establisher for the latter, one of the most probable reading directions, because of the viewers' most usual physical position, is the vertical one surprisingly conforms with the oriental verbal and graphic languages. If we

consider, for example, the narrative expression adopted by the father of the modern Japanese manga (*mangaka)* Osami Tezuka, we observe that the same narrative line direction is found (Fig. 3).

In fact, the contemporary manga graphic story, as set by its pioneer Tezuka, is usually made up of episodes that compose narrative arcs of variable length and direction within which the events are calibrated according to their importance by the syntax of the frames. In turn, the value of each panel is defined by the complexity of its morphological composition. *Mangaka* amalgamates the graphic essentiality of vertical oriental writing with the western narrative richness.

Fig. 4 – Osamu Tezuka, *Astro Boy, The birth*, vol.1, *Dark Horse* 27.

Observing one page of Tezuka's masterpiece *Astro Boy*, the assonance with the Michelangelo's narrative is discernible. The irregular narrative arc, created through three vertical panels sequence, increases the tension from the bottom to the top. The *consecutio temporum* of the entire page is determined by the main frame/phrase similarly to the one performed by the group of the Judge in Michelangelo's fresco. The western verbal and graphic space-time flow, namely a horizontally developing sequence of regular phrases/frames, is set aside by a syntactically irregular narrative scheme. The latter often abandons any reading direction inviting the user not to follow but to immerse into the image storytelling.

Two images belonging to two completely diverse epochs, cultural systems and artistic styles but governed by the same grammatical representational organization confirm the universality of the Visual Language Grammar located in a system of pre-established mental patterns inherent in everyone and only subsequently culturally modulated that are crystallized by Gombrich's pupil Michael Baxandall in the concept of Period Eye (for a discussion of this topic, see Nedkova 199).

5. Works Cited

Chatterjee, Anjan. "The neuropsychology of visual artistic production." *Neuropsychologia*, vol. 42, no.11, 2004, pp. 1568-83.

Cohn, Neil. "Combinatorial Morphology in Visual Languages." *The Construction of Words*. Springer, Cham, 2018, pp.175-99.

---. *Early writings on visual language*. Emaki Productions, 2003.

---. *The Visual Language of Comics: Introduction to the Structure and Cognition of Sequential Images*. A&C Black, 2013.

---, editor. *The visual narrative reader*. Bloomsbury Publishing, 2016.

Cohn, Neil, et al. "The grammar of visual narrative: Neural evidence for constituent structure in sequential image comprehension." *Neuropsychologia*, vol. 64, 2014, pp. 63-70.

Cohn, Neil, and Hannah Campbell. "Navigating Comics II: Constraints on the Reading Order of Comic Page Layouts." *Applied Cognitive Psychology*, vol. 29, no. 2, 2015, pp. 193-99.

Cohn, Neil, and Marta Kutas. "What's your neural function, visual narrative conjunction? Grammar, meaning, and fluency in sequential image processing." *Cognitive research: principles and implications*, vol. 2, 2017.

Cohn, Neil, and Stephen Maher. "The notion of the motion: The neurocognition of motion lines in visual narratives." *Brain research*, no. 1601, 2015, pp. 73-84.

Friederici, Angela D. "The brain basis of language processing: from structure to function." *Physiological reviews*, vol. 91, no. 4, 2011, pp. 1357-92.

Groensteen, Thierry. *The system of comics*. UP of Mississippi, 2007.

Jackendoff, Ray, and Jenny Audring. "Morphological schemas." *The Mental Lexicon*, vol. 11, no. 3, 2016, pp. 467-93.

Koch, Walter A. *Varia semiotica*. Vol. 3, Georg Olms, 1971.

Koelsch, Stefan, and Walter A. Siebel. "Towards a neural basis of music perception." *Trends in cognitive sciences*, vol. 9, no. 12, 2005, pp. 578-84.

Kuperberg, Gina R. "The proactive comprehender: What event-related potentials tell us about the dynamics of reading comprehension." *Unraveling the Behavioral, Neurobiological, and Genetic Components of Reading Comprehension*, edited by Brett Miller, Laurie E. Cutting and Peggy McCardle, Paul Brookes Publishing, 2013, pp. 176-92.

Manfredi, Mirella, Neil Cohn, and Marta Kutas. "When a hit sounds like a kiss: An electrophysiological exploration of semantic processing in visual narrative." *Brain and language*, no. 169, 2017, pp. 28-38.

Masuda, Takahiko, et al. "Culture and aesthetic preference: Comparing the attention to context of East Asians and Americans." *Personality and Social Psychology Bulletin*, vol. 34, no. 9, 2008, pp. 1260-75.

Marr, David, and H. Keith Nishihara. "Visual information processing: Artificial intelligence and the sensorium of sight." *Technology review*, vol. 81, no.1, 1978, pp. 2-23.

Molotiu, Andrei. "Abstract Form: Sequential dynamism and iconostasis in abstract comics and in Steve Ditko's *Amazing Spider-Man*." *Critical Approaches to Comics: Theories and Methods*, edited by Matthew J. Smith and Randy Duncan, Routledge, 2012, pp. 84-100.

Nedkova, Denitza Roumenova. "The thinking eye: from Semir Zeki to John Onians." *Comparatismi*, no. 2, 2017, pp. 199-208.

Tezuka, Osamu. *Astro Boy* 1 & 2. Dark Horse Comics, 2008.