Constructing Domains in Visual Narratives: Structural Patterns of Incongruity Resolution

BIEN KLOMBERG, JOOST SCHILPEROORD, NEIL COHN

Abstract: Understanding visual narrative sequences, like those in comics, requires readers to track situational information like the continuity of characters, locations, or events across panels. Yet, some sequences intentionally present incongruities, which may elicit the inference that the narrative presents two separate domains: an established, expected storyworld (primary domain) and an additional context surrounding the incongruous events (auxiliary domain), like a character's imagination, dreams, or memories. This paper describes how these inferences are supported by visual constructional patterns, which emerge across a wide range of narratives, yielding further insight into the fundamental role of graphic and structural cues within visual storytelling.

Keywords: incongruity resolution, domains, parallel architecture, visual language, comics.

To understand visual storytelling, readers need to track situational information like characters, places, and events across panels, and integrate these coherently in a mental representation of the narrative (Cohn, "Your Brain on Comics" 355; Dijk and Kintsch 5; Huff et al. 943; Loschky et al. 312). Generally, information that is discontinuous with the current context is a cue for an event boundary (Cohn, "Your Brain on Comics" 360; Gernsbacher, "Coherence Cues" 3; Loschky et al. 322), but some sequences may include intentionally incongruous information that needs to be meaningfully incorporated in the current storyline. How do we deal with such issues of incongruity?

Consider Figure 1, which shows short comics from *JA!*, a series about the experiences of its two collaborative creators (who appear as characters in the final scenes of each strip). In all the strips in Figure 1, the final panel is incongruous with the prior series of events, juxtaposing distinct places and agents within a layout. This divergent information typically prompts readers to engage in "incongruity resolution", i.e., finding an interpretation that explains the incoherence (Schilperoord 2–4, 18–19; Forabosco 47). For Figure 1, readers can resolve the differences across characters/locations when inferring that the first series of events are not *actually* occurring or present at that moment/place, but function as metaphoric depictions of the "real" characters' actions and/or feelings. In Figure 1a, the mole's actions are analogous to the woman's: she bumps into things while "as blind as a mole", seeking the glasses worn by the dog. Similarly, Figure 1b's piranha (devouring meat quickly) and snail (eating a leaf slowly) seem to reflect the eating habits of the man and woman, while Figure 1c compares the woman catching a spider for the man to a superhero defeating a monster to save a civilian. Here, the animals, superhero, and monster would not be physically present in the focal characters' storyworld but can be inferred as additional contexts applied to those characters. The recognition of these incongruities and their resolutions are essential to the humor intended by these comics (Forabosco 47), and give a reason for why this incoherence was shown (Schilperoord 4): they reflect the characters' inner thoughts and feelings.

Journal of Comparative Literature and Aesthetics Vol. 47, No. 3, Autumn 2024 [37-55]

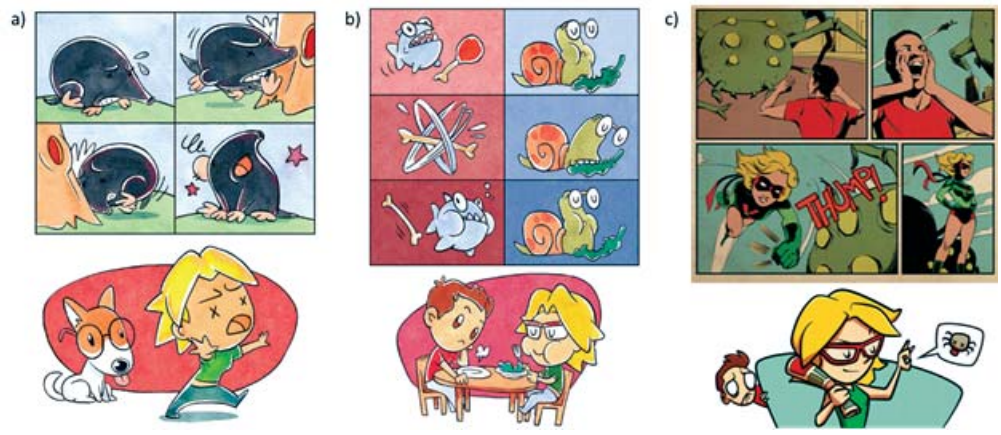

Figure 1. Examples of incongruous events that can be meaningfully incorporated.
All sequences are from the series *JA!* © Ángela Cuéllar and Jonás Aguilar (Creative Commons License).

Intentional incongruities like in Figure 1 have been widely recognized in visual narratives, united by the inference that some events do not actually occur at that place/moment. This gives rise to two contexts that should be integrated during incongruity resolution: a primary domain (the established storyworld, e.g. the man and woman) and an auxiliary domain (an additional context associated with events inferred to not occur, e.g. the analogy with animal habits/superheroes). Such inferencing may make sequences more complex but also engaging, rewarding, and fun, since readers are challenged but can still manage to construe meaning (Ramachandran and Hirstein 31; Giora et al. 138). In Figure 1, these domains (and their resolution) provide the strips with their aesthetic appeal. Such inferential challenges may also heighten narrative tension, e.g. for horror and mystery genres that may deliberately prolong cues that help to resolve incongruity. Since the resolution of these sequences relies on successfully integrating two domains, and such phenomena appear to rely on patterned structures, we refer to these as *domain constructions*, but similar phenomena have been subsumed by work on perspective taking or focalization (Maier and Bimpikou; Mikkonen 110; Horstkotte and Pedri 38–40).

Although prior approaches have attempted to describe the semantics related to resolving the incongruities within domain constructions (Abusch and Rooth 8; Bimpikou 246; Maier and Bimpikou), they have not addressed what visual forms might cue domain constructions, and how these cues correspond to meaning or the structural patterns that arise across such features of form. In Figure 1, the layout of the three sequences is such a form-related cue: all the panels showing the animals'/superhero's actions create a grouping, while the panel showing the human characters is its own row. These groupings are reinforced by panel borders, with the first four/six panels using explicit borders, while the final panel has none. Additionally, aspects of form may create connections between primary and auxiliary figures. In Figure 1a–b, the eyes of the mole/snail are drawn the same way as the woman's eyes, and in Figure 1b and 1c, the red of the meat/ civilian's shirt corresponds to the man's shirt, while the green of the leaf/superhero's top corresponds to the woman's shirt, and the superhero mask in 1c to her red glasses. Overall, formal features align with the semantic incongruity and may cue intended relations between domains and their respective figures.

This paper aims to account for such patterns and how they may signal incongruity to be resolved. Our examples include a variety of comics by different authors, highlighting how domain constructions appear widespread in storytelling for authors to convey characters' personal experiences, while aligning these meaning relations with conventionalized graphic features. The analyses are embed-

ded in a cognitive and linguistic approach to visual representations, and as such, aim to describe the patterned nature of these encoded constructions. We thus investigate the effects of graphic and structural cues on (inferential) meaning-making in storytelling and what constructional patterns persist across authors and their works. Additionally, such patterns could apply to other visual media, e.g., film, which maintain many of the same affordances as drawn visual narratives. Naturally, creators may also rely on similar schemas when working across media, often because film storytelling first occurs as drawn storyboards. To discuss domain constructions for comics, we first present how visual narratives are processed and how previous work has described issues of domain constructions similar to Figure 1. Next, we propose a model that describes patterned correspondences between form and meaning related to this construction.

Domain constructions in visual narratives

To comprehend visual sequences in general, readers integrate story components into one coherent, mental representation of the narrative (Cohn, "Your Brain on Comics" 355; Dijk and Kintsch 5; Huff et al. 943; Loschky et al. 312). This representation is known as a situation model, and includes situational aspects such as characters, their motivations, spatial locations, time, etc. (Dijk and Kintsch 46; Zwaan and Radvansky 167; Gernsbacher, "Coherence Cues" 5). With each incoming piece of information, the situation model gets updated, and sufficiently congruent elements can be mapped onto the existing model (Cohn, "Your Brain on Comics" 360; Gernsbacher, "Coherence Cues" 3; Loschky et al. 321). The more incongruent the incoming information, the greater the updating costs (Huff et al.; Hutson et al.; Magliano et al.), until significant discontinuity sponsors an event boundary (Gernsbacher, "Two Decades" 5; Loschky et al. 322).

While sequences with auxilairy domains are incongruous and have been shown to require more updating indeed, these were equally comprehensible as sequences with only a primary storyworld (Klomberg, Fadeeva, et al.). Moreover, domain constructions seem to be fairly common, with corpus work showing they are prevalent in around 65% of comics (van der Gouw et al. 24). The ubiquity and comprehensiblity of domain constructions give rise to the question of what mechanisms underlie this storytelling technique.

Prior work converges on a general semantic distinction between visual events that are perceptible to all characters versus those that are perceptible only to a single character (Abusch and Rooth 8; Bimpikou 247; Duncan 275; Horstkotte and Pedri 13, 147–48; Maier and Bimpikou). To mark this distinction, we use the notion of *domains* (Clausner and Croft 2; Langacker 147), which denote particular types of mental spaces (Fauconnier and Turner 40) within the context of narrative sequencing, and thus a certain context that visual events can be attributed to. The *primary domain* represents events that are inferred to transpire in the "actual" storyworld, i.e., the world physically accessible or perceivable to all characters. An *auxiliary domain* on the other hand, represents events inferred as "not physically occurring at that time". Auxiliary domains are typically resolved as depictions of a (private) mental experience of a certain character, such as their fantasy, hallucinations, dreams, or memories (Horstkotte and Pedri 40–48; Abusch and Rooth 12; Maier and Bimpikou; Bimpikou 255; Mikkonen 110). Work on focalization (Horstkotte and Pedri 1) describes domain constructions by analyzing how a narrative may present the experiences of a certain character, while approaches from semantics and discourse theory focus on the veridicity of events (Bimpikou 247; Maier and Bimpikou).

In addition to semantic distinctions between events, prior work foregrounds the role of perspective-taking for these sequences (Abusch and Rooth 8; Bimpikou 249; Maier and Bimpikou), with inferences of subjectivity as cue for auxiliary domain events. Consider the modified example in Figure 2a, exemplifying a two-panel construction that implies subjectivity. When a panel with a character looking off-panel (the man in red) is followed by an object incongruous with the narrative (the snail), the sense of perspective-taking evoked by this two-panel construction facilitates the

inference that the incongruity is a perception private to the character, i.e. their imagination, dream, or hallucination (Abusch and Rooth 8; Bimpikou 248–49; Maier and Bimpikou). This contrasts with "blended-perspective" (Bimpikou 246; Maier and Bimpikou), where elements of distinct domains are merged in a single panel. In Figure 2b, a man looks into the window of a shop but sees a reflection of a different person, with no apparent co-reference. Contextually, only the man can see the reflection of this dark-haired figure, who also haunts him in his dreams later on. Thus, this reflection seems to be imagined by the man, as if he and this figure were one in this moment (reinforced by the comic's title: *Bonding*). In this way, blended-perspectives merge the physical storyworld (the shopping street) with the private thoughts of a character (the reflection of the dark-haired figure).

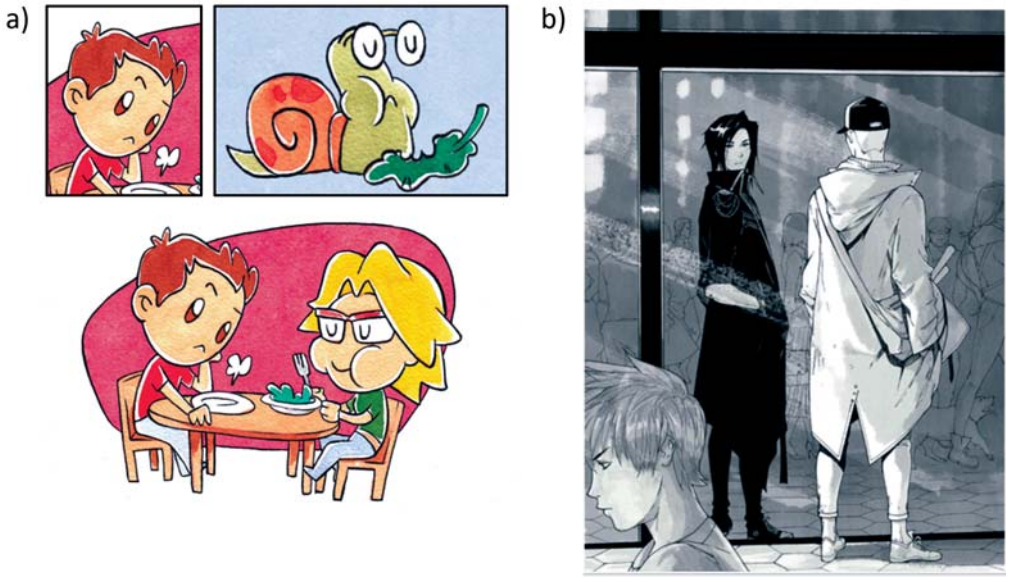

Figure 2. Examples illustrating a) a two-panel construction and b) a blended perspective. Figure a is a modified version of a sequence from the series *JA!* © Ángela Cuéllar and Jonás Aguilar and Figure b from *Bonding* © 2015 Lumarin and Marina Privalova.

While the availability of an identifiable experiencer may be helpful for the recognition of domain constructions, not all auxiliary domains can be readily interpreted as a particular character's experience. For Figure 1a-c, it is unclear whether these sequences are imagined by a storyworld character (does the woman imagine herself as a mole/snail/hero?) or a comparison made by a non-diegetic narrator. Despite this conceptual ambiguity of *why* the two event types are integrated together, the cues sponsoring those inferences (e.g. layout, visual patterns, colour correspondences, etc.) can be identified without issues. It thus seems that these cues can function independently from meaning.

These and other patterns have received little focus across research on visual narratives, despite some prior observations of cues for individual images. For instance, previous analyses acknowledge that divergent panel borders or changes in layout correspond to incongruous events (Duncan 275; Horstkotte and Pedri 40–46; Mikkonen 110–11), but they have not been generalized beyond the specific examples that are addressed. Moreover, there are patterns not yet accounted for, e.g. Figure 1's segmentation of panels and its ambiguous (or lacking) experiencer. Thus, while existing theories have yielded valuable insights into the meaning-making processes involved in domain resolution, these do not acccount for the formal patterns that signal these construals, which may give the false impression that such construals appear without patterned contexts. How then can we account for these various visual narrative constructions?

The Domain Constructions model

To account for the structured patterns of form and meaning that occur in domain constructions, we here propose the Domain Constructions model, or DC model. The model describes structural correspondences between form and meaning that occur for domain constructions, where incongruity across situational aspects and form-related features cue the search for a meaningful interpretation. This involves integrating an auxiliary domain (events inferred to not physically occur at that time/place) within the story's primary domain (the physical storyworld accessible to all characters). The proposed model is not restricted to cues signalling domains across panels but also pertains to elements within panels (Bimpikou 245–46; Horstkotte and Pedri 40–42; Maier and Bimpikou).

The DC model is embedded within the broader framework of Visual Language Theory (Cohn, *The Visual Language* 1–13) which posits that graphic communication maintains structures and cognition consistent with natural languages, as has been supported both theoretically and empirically (see Cohn, *Who Understands Comics?* (20–22)). Visual Language Theory itself is embedded within the linguistic model of the Parallel Architecture (Jackendoff 107–26), which characterizes language in terms of three parallel, but independent structural components: modality (phonology), grammar (syntax), and meaning (conceptual structure), as well as interfaces that emerge between these. The Parallel Architecture has been applied to characterize visual language in terms of the same components of modality (graphic structure), grammar (combinatorial structure), and meaning (conceptual structure) (Cohn, *Who Understands Comics?* 5–6). These three structures, shown in Figure 3 below, describe both single units, e.g. a single panel, and sequences of units, e.g. multiple panels on a page.

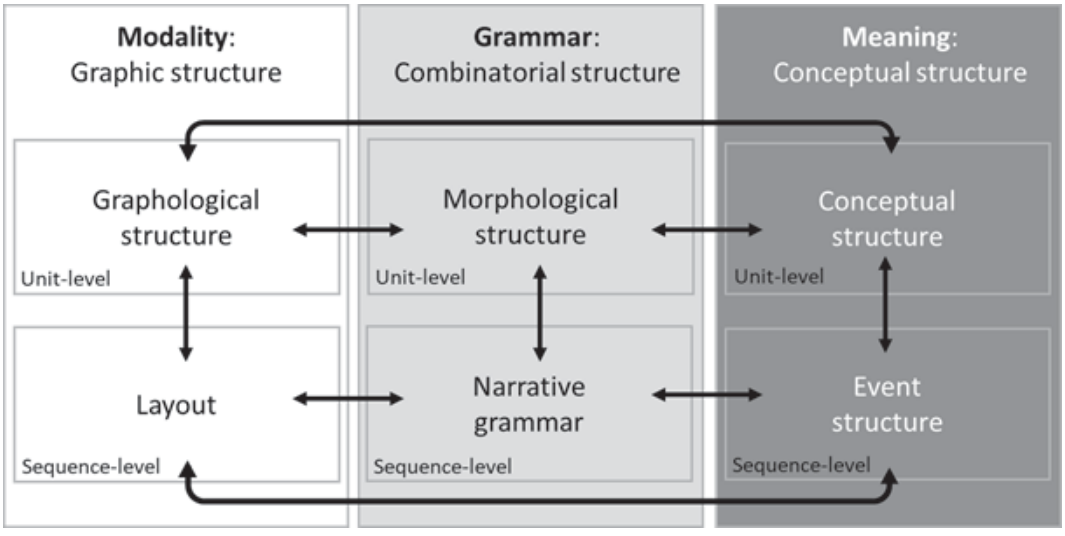

Figure 3. The Parallel Architecture framework.

In visual expressions, each component involves unit- and sequence-level structures, indicated by the vertical arrows in Figure 3. The modality component (i.e., graphic structure) related to single panels is termed graphological structure (Klomberg, Hacımusaoğlu, et al. 10), which characterizes the organization of the visual markings that make up the images, such as lines, dots, or colours, and the visual regions they compose. Graphic structure across panels is captured by layout, which accounts for how sequences of images are integrated spatially on a canvas. The grammar component characterizes the combinatorial structure of a single panel in terms of its morphological structure, and organizes sequences of panels according to principles of Narrative Grammar (Cohn, *Who Understands Comics?* 12–16). Finally, the meaning component describes how panels are understood

independently (conceptual structure) and in sequence (event structure). For domain constructions such comprehension relies on similar principles, as the same conceptual inferences apply within and across single panels.

While domain constructions would be resolved at the level of conceptual structure, the DC model accounts for the ways that graphic and combinatorial structures contribute significantly to their recognition and resolution; these interfaces across components are expressed by the horizontal arrrows in Figure 3. Hence, the following sections will detail how constructions encoded across these components signal auxiliary domain usage and/or support relations of co-reference across entities that initially appear non-co-referential. We first present structural correspondences between the graphic and conceptual structures, then between grammatical and conceptual structures, and finally, structural correspondences across all three components.

Graphic Structure

Graphological Structure. To describe how form may interact with meaning we must first characterize the graphic form of visual representations. Analogous to how phonological structures organize sound structures, a graphological structure (Klomberg, Hacımusaoğlu, et al. 10) organizes visual-graphic components. The basic units of those components are graphemes, i.e. dots, lines, curves, and spirals, which combine to create larger structures of "regions": enclosed visual groupings (either explicitly so or inferred) (11). Regions, in turn, can recursively combine to form even larger graphic structures that correspond with conceptual representations, i.e., they can be recognized as meaningful objects (Willats 94). Graphological structure is thus concerned with purely graphic information, such as the lines and proporties of regions. We here propose that differences across these graphic features are often used to signal domain constructions.

Figure 4. Examples illustrating graphological cues corresponding to meaning. Figure a is from *Socrate* © Francesco Barilli and Alessandro Ranghiasci, Figure b from *Jar of Fools* © Jason Lutes, and Figure c from *Raruurien* © Ann Maulina.

Figure 4 illustrates three types of graphicological cues that may signal different conceptual domains: colour, linework, and shape similarity (Horstkotte and Pedri 12; Mikkonen 116–20). Consider Figure 4a, where the first page shows the philosopher Socrates during his own trial, the second page presents Socrates amongst different people in a different location, and the third page shows different figures in an unknown environment. These differences in both characters and locations confront the comprehender with an incongruity. Crucially, each page here is marked with a distinct colour scheme. The incongruity across pages can be resolved by construing the second page's events as taking place at another moment (Socrates teaching his students), with the same agent in both scenes (Socrates). The third page can subsequently be construed as the contents of Socrates' lesson. These inferences are signalled by the pages' graphological differences, with the first page's trial (the primary domain) shown in colour, the second page's memories (an auxiliary domain) in black and white, and the third page's exposition (a second auxiliary domain) in black and orange hues, like a Greek vase (Andrew et al. 2). A similar alignment of colour schemes and domains occurs in Figure 4c, where the final incongrous panel can also be inferred as a past event featuring the same agent (the woman). This panel is coloured only in blue hues, as opposed to the range of colours used in previous panels.

Differences in the linework of drawing styles may also correspond to domain constructions. In Figure 4a, the third page's drawing style further mimics Greek vase art , with figures shown mostly from the side and without shading (Andrew et al. 5). This stylistic shift again aligns with the conceptual interpretations that arise for page 2 and 3. While both pages show colour differences that correspond to different domains, page 2 maintains the same drawing style as the one used for primary domain events on page 1, warranting the interpretation that the memory event bears a substantive conceptual relation to Socrate himself, i.e. his lived experience. In contrast, page 3's features elicit associations to the myths and legends typically shown on those Greek vases (Andrew et al. 12), which aligns with the idea that this event is fictitious even to Socrates himself. Seemingly, the use of differing drawing styles across pages 2 and 3 in Figure 4a functions to distinguish between the two different auxiliary domains.

Graphological differences signalling different domains also emerge from the linework of the panel borders. In Figures 4a–c, diffferent panel borders may support inferring those panels as corresponding to auxiliary domains. Figure 4a's third page has thick, decorated borders which contrast the previous two pages' borders, thus emphasizing that this last page may present a second auxiliary domain rather than a continuation of the memory on page 2. In Figure 4b, the use of different borders likewise supports a distinction between the events captured in those panels: the man dreams that he is dancing with the woman (shown in panels with rounded borders), and wakes up to find he is actually sitting in a park (shown in panels with square borders). In Figure 4c the final panel has distinctly frayed borders, compared to the previous rectangular borders, reinforcing that the final event is part of an auxiliary domain (here, a memory). Besides these differences in panel borders' linework, differences can also arise when borders are omitted, as in Figure 1. Overall, incongruity in lines across panels may signal that additional updating is needed.

Finally, similarities in shapes across panels can motivate different domains. In Figure 4b, the switch between domains occurs between panels 8 and 9, which both depict the main character. Noticeably, the character's composition is highly similar across panels, with his hands in a similar position relative to his face despite changes in his expression, hair, clothing, and background. The overall similarity alongside small differences seems to emphasize that this same moment occurred across two domains (him holding up his hands in a dream and in the physical storyworld). Experimental research (Klomberg, Fadeeva, et al.) suggests that familarity with comics may mediate whether such contour similarity affects processing, with more experienced readers slowing down for sequences lacking this similarity.

As shown by our analyses, these graphological cues do not indicate the presence of auxiliary/primary domains themselves, but rather do so in contrast to other panels (Mikkonen 119). For

example, a whole sequence can have rounded, wavy, or no borders or be shown in greyscale or blue hues, but this alone does not signify auxiliary events. When these features contrast other panels that have been established as normative to a primary domain, it may prompt readers to infer the events with incongruous features as part of an auxiliary domain. These graphological patterns thus include open slots, filled with whatever applies to a specific comic (e.g. Figure 4b has rounded borders as incongruent vs. square borders as expected, while Figure 4c contrasts wavy vs. square borders). Such cues often apply to several panels; e.g., Figure 4a–b's colours schemes and divergent borders persist for all panels depicting those auxiliary domain events. Additionally, graphological features can easily combine, whether border differences and shape similarity (Figure 4b), colour and border differences (Figure 4c), or colour, style, and border differences (Figure 4a).

Layout. Layout refers to the part of the graphic structure that concerns the spatial organization of panels on a larger canvas (e.g. a page) (Cohn, *Who Understands Comics?* 9–11). We already described one pattern of layout for Figure 1, where the auxiliary domain panels cluster together followed by a single panel showing the primary domain. This general construction is shown in Figure 5a where tree structures are used to specify the organization of these sequences.

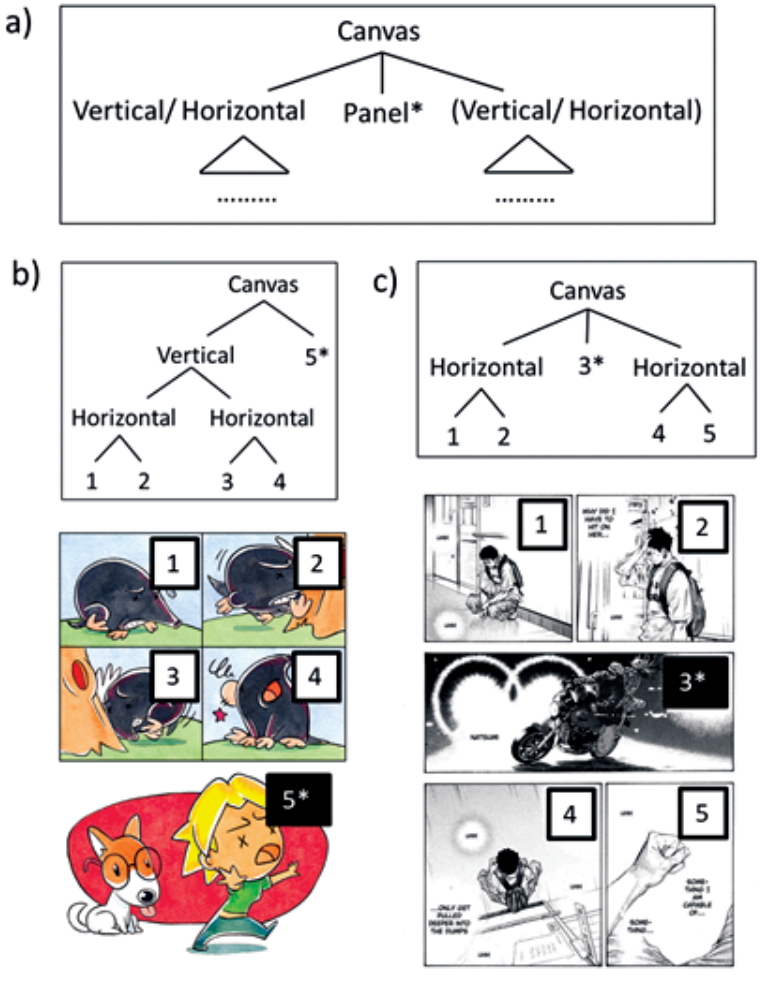

Figure 5. Examples illustrating layout trees with the * indicating the introduction of a second domain (whether primary (b) or auxiliary (c)). Figure b is slightly adapted from the series *JA!* © Ángela Cuéllar and Jonás Aguilar, and Figure c is slightly adapted from *REAL* © Inoue Takehiko.

Figure 5b shows how Figure 1's examples would fit this organization: the events in one domain (panels 1 to 4) are structured as multiple horizontal rows stacked on one another vertically. This column of panels then combines with another row underneath that corresponds to the second domain (panel 5), forming an ultimately vertical layout. This construction can have additional internal structure within its lower-level vertical grouping; e.g., Figures 1a and 1c feature two rows, while Figure 1b features three. Note that Figure 5b is ambiguous: it could also be viewed as three horizontal rows stacked upon each other, rather than a column upon a row. Nevertheless, either interpretation maintains the general construction illustrated in Figure 5a and demonstrates a recurrent pattern of segmentation across panels that aligns with conceptual inferences, with a sequence's second domain corresponding to a grouping on its own.

A second instantiation of the general pattern is shown in Figure 5c, where the horizontal row combines with other horizontal structures not only before but also after. Here, a boy in a hallway (panels 1–2) remembers himself and a girl on a motorbike (panel 3). This panel on its own (panel 3) corresponds to an auxiliary domain of a past event. In Figure 4a, different domains are distributed on different pages. Thus, it seems that layout segmentation may support conceptual interpretations of domains by presenting distinct panel groupings within page structures as well as across pages. Such segmentation echoes other observed alignments between coherence shifts and layout structure (Hacımusaoğlu et al.).

All the patterns discussed so far strictly concern the organization of panels and/or pages, and render available slots that can be filled by particular forms, rather than specific meaning. In Figure 5b, the single panel functioning as a row corresponds conceptually to a primary domain, while in 5c, the single panel corresponds to an inference of an auxiliary domain. This emphasizes that these constructions are aspects of layout which can interface with meaning in various ways. These patterns describe how the narrative switches domains in general, regardless of the direction of this switch, or the particular nature of the domains involved (Klomberg, Fadeeva, et al.).

Nevertheless, one layout structure appears to relate to auxiliary events more directly. Figure 6 shows two collage layouts, where multiple scenes graphically overlap without overt panel borders. Here, both pages show events that can be inferred to have taken place in the past and can thus be identified as auxiliary domain events. Collage layouts potentially emphasizes the sense that such events are "flashes" of memory. Corpus work indeed found that 83.8% of collage framing contained auxiliary domains (van der Gouw et al. 36), supporting a close relationship between them.

Figure 6. Examples illustrating collage sequences. Figures are from *Valley of White Birds* © Wolfsmoke studios and from *Pirozhki* © Sideburn004, respectively.

Combinatorial Structure

Morphological Structure. This component of the Parallel Architecture is concerned with the patterns of abstracted forms of visual units (Cohn, *The Visual Language* 23–24; Cohn, "Combinatorial Morphology"; Cohn, *Who Understands Comics?* 12). Similar to morphology in verbal language, visual morphology contains units of visual language that can stand on their own (e.g. a single character or object) next to bound morphemes that need to combine with units to be meaningful. These components use various strategies to combine, such as via suppletion, e.g. replacing a character's eyes with hearts to show love, or blending, e.g. depicting a character's normal body with an animal's head (Cohn, *The Visual Language* 44–47; Cohn, "Combinatorial Morphology").

These morphological patterns may signal domain switches and semantic (co-referential) connections across characters from different domains. Figure 7a illustrates a gradual replacement (partial suppletion), where the man's increasingly disturbing facial features from panel 4 to 7 signal that the woman slowly realizes his malicious intentions. Figure 7b's morphology signals the start of a dream: a man was watching a movie with a female protagonist (panels 1–3), but evidently fell asleep as the movie becomes personal and nightmarish, and he wakes up on a later page. He apparently falls asleep around panel 4, when the actress' appearance blends with his own. This "morphological blend" shows parts of one meaningful entity (the woman's hair, clothes, and earrings) merged with parts from another entity (the man's own face and body hair), which supports the inference that this is a dream rather than a movie. Finally, Figure 7c shows visual components repeated across different figures, with the woman and animals having the same eyes and eyebrows. These we term "echoic morphemes": visual components that mimic one another morphologically. Rather than signalling a domain switch, repeating visual components may signal readers to infer some semantic relation between characters, and as such, may support subsequent inferences of co-reference across these units (Klomberg, Hacımusaoğlu, et al. 21).

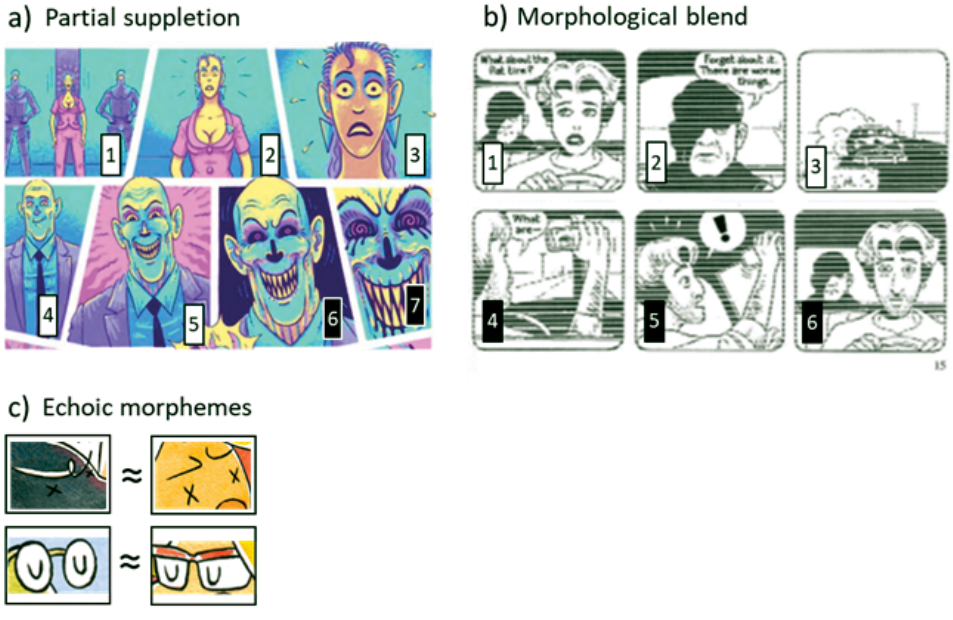

Figure 7. Examples illustrating the morphological strategies of a) partial suppletion, b) morphological blend, and c) echoic morphemes. Figure a is slightly adapted from *Vicki Lante* © Patrick Steptoe, Figure b is slightly adapted from *Jar of Fools* © Jason Lutes, and Figure c is slightly adapted from *JA!* © 2016 Ángela Cúellar and Jonas Aguilar.

Narrative Structure. Where morphology is concerned with the relations between visual units within panels, narrative structure characterizes the combinations of panel units, along with the function these panel units play within a narrative sequence. The combinatorial structure of sequences is captured by Visual Narrative Grammar (VNG) (Cohn, "Visual Narrative Structure" 420–38; Cohn, *Who Understands Comics?* 43–46), which identifies the narrative categories of panels and the hierarchic embedding of larger sequences of panels. We first explain the narrative categories panels can have, and then analyze how these categories may interface with domain switches.

Figure 8 shows the narrative structure of the comics in Figure 1 according to VNG (Cohn, "Visual Narrative Structure" 420–27; Cohn, *Who Understands Comics?* 43–46). Let's start with the first grouping in each of these sequences. In Figure 8a, panel 1 introduces the scene and characters (an Establisher), and then panel 2 sets up (an Initial) the climactic moment in panel 3 (the Peak), until this tension diminishes in panel 4 (a Release). These four panels follow the canonical narrative schematic order (E-I-P-R) within VNG. Figures 8b and 8c use an additional "conjunction schema" where identical categories follow one another. Figure 8b uses an action-conjunction (see subscript a), where panels of the same narrative categories repeat similar actions. Figure 8c shows an environmental-conjunction (subscript e), where the depicted characters or objects are inferred to belong to the same environment.

Narrative categories not only apply to individual panels, but also to constituents of panels. In Figure 8a, the first four panels constitute the climax of the sequence, thus forming a Peak sequence, while the final panel diminishes that tension and acts as a Release for all preceding panels. Figure 8b and 8c use this same pattern, with a final Release panel following a Peak segment.

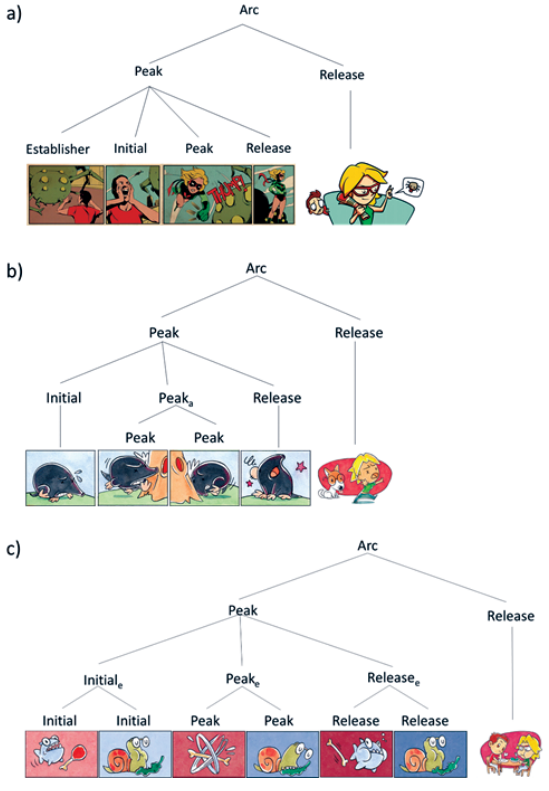

Figure 8. Examples illustrating grammatical tree structures with a a) canonical scheme, b) action conjunction, and c) environmental conjunction. Figures are slightly adapted sequences from *JA!*

For these sequences, narrative constituents correspond with the conceptual division between domains. Note that this division persists despite internal variation within the constituents, with the Peak-Release sequence on the highest level corresponding to a division between auxiliary and primary domain events. Other narrative constructions, e.g. conjunctions, could similarly support domain constructions across panels, and will be investigated more extensively in future works. As shown for other structures, these patterns can be expected to occur regardless of what conceptual domain fills what slot.

Meaning

Conceptual Structure. Conceptual structure specifies how representations are involved in meaning-making, which in the case of domain constructions often involves complex inferencing. Incongruity occurs across situational information, with time, space, or characters mismatching across panels, signalled also through differences in graphology, layout, morphology, and grammar. Such incongruity can then prompt the resolution that those depictions reflect events that do not actually occur in the storyworld (the primary domain), but instead correspond to some alternative context (an auxiliary domain). All examples shown so far involved such inferencing. Despite deviating from the primary domain, resolving these incongruities allows for coherent integration into situation models.

The way that domains are to be integrated to arrive at a sensible resolution will depend on the inferrred nature of the auxiliary domain, i.e., imagination, hallucinations, dreams, and memories (Abusch and Rooth 12; Bimpikou 255; Horstkotte and Pedri 40–48; Mikkonen 110–12). For instance, incongruent appearances of a character can be resolved as happening in a dream in which the character's appearance transforms, rather than this character actually (physically) changing or that a new character is being introduced. Particular graphic and/or structural patterns may be used for certain types of auxiliary domains; e.g., greyscale panels seem more indicative of memories than hallucinations, for which we may instead expect bright or atypical colours. So far, the DC model categorizes nine types of auxiliary domains: Imagination, Knowledge, Dream, Memory, Foresight, Hypothetical, Fiction, Fourth Wall Break, and Analogy, along with cases that may remain Ambiguous.

Imagination and Knowledge domains are inferred as events that occur in the mind of a single experiencer during an active (awake) state. The page in Figure 9a (read from right to left) precedes the "blended-perspective" from Figure 2b, and implies the reflection is *imagined*: the man with the cap (the experiencer) sees an incongruent reflection in the shop window (panel 3), which only he appears (able) to see, as the reflection disappears when someone else is present (panel 6). Generally, Imagination domains show fantasized events originating from a single, awake character that other characters seem unable to perceive, e.g. active imagination, daydreams, or hallucinations. In Figure 9b, the swimming techniques can also be inferred to occur in the mind of the boy only, due to the juxtaposition of these techniques and the close up of the boy's face, and reinforced by the absence of physical materials he could be looking at in the rest of the sequence. Rather than a fantasized event, however, these seem a visual representation of the boy's *knowledge* relating to specific techniques. While both domains appear private to one experiencer, Imagination typically depicts made up information, while Knowledge reflects thoughts of a more substantive nature, such as spatial knowledge (e.g. a map of the environment), encyclopedic knowledge (e.g. how a certain object looks or functions), or instructions (e.g. how to perform certain actions).

While these events occur during characters' awake state, experiencers may also visualize events involuntarily, e.g. when dreaming. In Figure 9c, the third panel shows the experiencer jolting upright with a scream, suggesting he awakens, which is emphasized by the sixth panel depicting him on his bed. The incongruity between the first two panels and the rest can thus be resolved by inferrring the first two panels to represent a dream. Typically, Dream domains are preceded by a character lying down or closing their eyes and/or followed by a character wakening or opening their eyes, providing contextual cues that signal the incongruous events as part of a dream.

Figure 9. Examples showing an a) Imagination domain, b) Knowledge domain, and c) Dream domain. Figure a is from *Bonding* © 2015 Lumarin and Marina Privalova, Figure b from *Poisson* © Nie Jun, and Figure c from *Jar of Fools* © Jason Lutes.

These three domain types concern events assumed to originate completely from someone's mind, but two other types –Memory and Foresight– denote events that should be inferred to occur at some past or future point in the story's timeline. Figure 4a (Socrates's trial), Figure 4c (the woman remembering being angry), and Figure 6 (the collage examples) all show examples where incongruous events have to be construed as a protagonist's memories. Similarly, the collage panel in Figure 10a shows the protagonist overlayed with a series of events that have occurred before, some of which even have been shown in the story explicitly. In these instances of Memory domains, the primary assumption is that the depicted moment is something that the experiencer recalls from the past (which may or may not have been shown earlier in the story). Narratives can also show events that will occur in the future, e.g. via a prophecy or a character having a vision. Figure 10b shows an example of a witch watching her crystal ball with the vision illustrated on the right (a princess picking up a sword, coaxed by the snake on her shoulders). This exemplifies a Foresight auxiliary domain, representing (likely) future events.

Figure 10. Examples showing a a) Memory domain, b) Foresight domain, and c) Hypothetical domain. Figure a and b from *Valley of White Birds* © Wolfsmoke studios and Figure c © Sharitha van der Gouw.

In addition to past and future events, a depicted event may or may not come to pass, i.e., a Hypothetical auxiliary domain. Examples are thought experiments (an answer to the question "What if…?") or outcomes related to very specific circumstances. In Figure 9c the first panel shows an old man saying that "the white goddess" (i.e. the white creature in the second image) may be caught off guard only by the protagonist in panel 2 and 3, and so he "might just be the only one who could kill her". The panel in between these speech balloons shows such a scene where the protagonist has killed the white goddess. However, the reader knows this event has not (yet) happened, and so the panel shows an Hypothetical situation (cued predominantly by the modal keyword "might"), without indicating whether the event is likely to occur or not.

Thus far, we discussed auxiliary domains representing events revolving around the experiencer that visualizes, remembers, or otherwise inspires that event. However, auxiliary domains can also represent events that do not appear private to an experiencer. The first three panels of Figure 11a show different moments of a fight, while the final panel shows a person holding a comic book showing these exact three panels. Here, the first three panels depict fictitous events (a comic book narrative), and the final panel depicts the primary storyworld domain, where the story will continue. This auxiliary sequence is thus part of a Fiction domain, which can be described as a 'story within a story' and/or events that are fictitious to the characters themselves. Examples include depictions of folklore history, fairytales, or other accounts of (past) events that do not involve the narrator itself but other (fictitious) people.

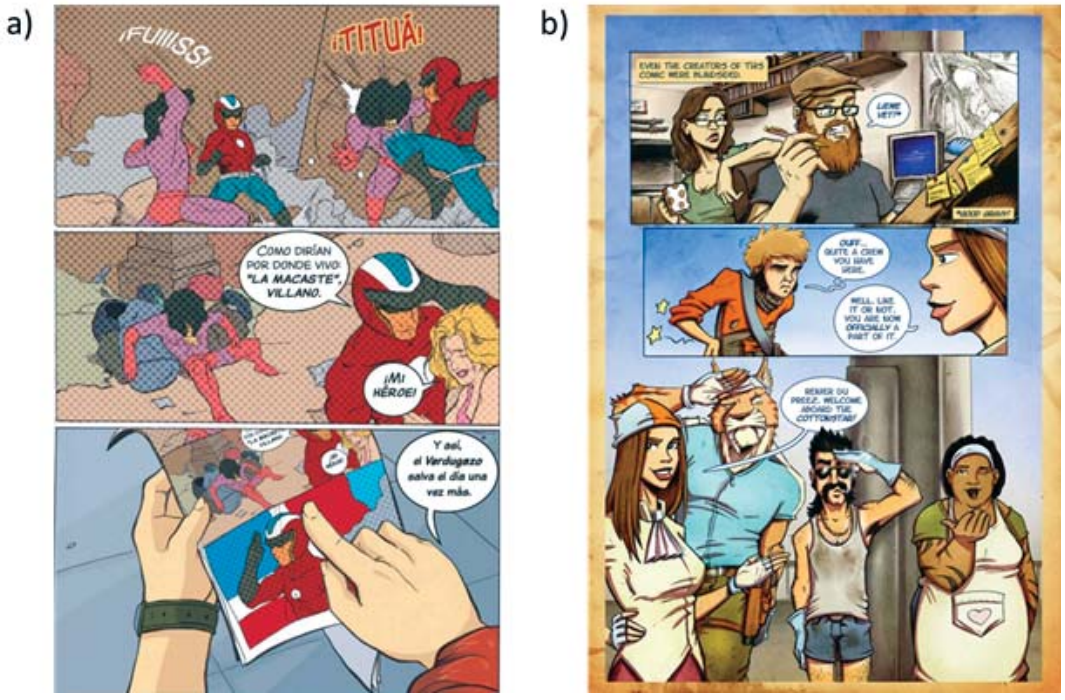

Figure 11. Examples showing a a) Fiction domain and b) Fourth Wall Break domain. Figure a is from *El Verdagazo* © Jorge E. Pérez and Francisco Alberto López and Figure b from *Cottonstar* © Ben Geldenhuys and Danelle Malan.

Additionally, narratives may also suspend the established storyworld altogether, as in Figure 11b where the first panel shows the comic creators reacting to a story event on the previous page, while the second panel continues in the primary storyworld domain. Thus, this first panel steps out of the story for a moment and visualizes figures with no relation to the narrative at all (except supposedly having created it). The incongruity across figures and places is resolved when the first panel is interpreted as an event from the creator's own world. We summarize these as instances of a Fourth Wall Break domain. It is worth pointing out that this also allows visual narratives to be self-referential, and to have creators comment on their own creation.

The last type of auxiliary domain is Analogy, which resolves incongruity via a (metaphoric) comparison between two domains. This was the case for Figure 1, where the animals and superhero figures metaphorically represented the actions and/or feelings of the actual characters shown in the final panel. For these domain constructions, it may be ambiguous whether an diegetic experiencer is behind the analogy (i.e., whether the woman imagines herself as a mole/snail/hero) or whether the comparison can be attributed to a non-diegetic narrator/author.

Finally, auxiliary domain events cannot always be easily categorized and may evoke multiple possible interpretations. A sequence may not imply an experiencer explicitly, e.g. Figure 1, or Figure 10c, where the Hypothetical situation of the protagonist killing the white goddess could be attributed to the mind's eye of either the protagonist or his mentor (or both). Even sequences with overt experiencers may be Ambiguous, when it is unclear what action the experiencer is involved in (imagining, remembering, etc.). In those cases, contextual cues then do not distinguish whether the depicted scene is private to the experiencer or not, or whether the depicted scene could have already occurred in the past. Alternatively, cues may overlap, such as when a characters's dream features a

memory or likely future event. This ambiguity is likely presented on purpose to leave readers in suspense or create more possible inferences (e.g. is it a dream or a prophetic vision?), requiring (or granting) more reader involvement and freedom to interpret sequences according to readers' own, strongest associations, which may contribute to the pleasure of reading.

Meaning across modalities

The discussion so far related to visual information; however, discontinuities also persist in multimodal interactions between text and images. Various examples illustrate such an interaction. In Figures 4c, 6a, and 9d, the text includes past tense and deictic indicators (4c: "that time", 6a: "back then", 10a: "my life like this…has been"), supporting the interpretation that the visual image depicts a past event. Figure 10c is even more overt, with the modal verb "might" reinforcing the inference of a Hypothetical domain. In Figure 11b too, the text in the first panel identifies who is depicted ("the creators of this comic"), which would have been difficult to establish otherwise. These references may even work in hindsight or advance, e.g. when auxiliary events are referenced in later panels (e.g., "I had a bad dream"), textully establishing an event as a particular domain type (here, Dream).

Interfaces across Parallel Architecture structures

So far, we have discussed how the DC model operates on correspondences between graphic structure and meaning, and between combinatorial structure and meaning, and how both support understanding of domain constructions in visual narratives. Correspondences across these structures may align in various ways for visual narratives, which gives rise to different ways that auxiliary domains are presented within a sequence.

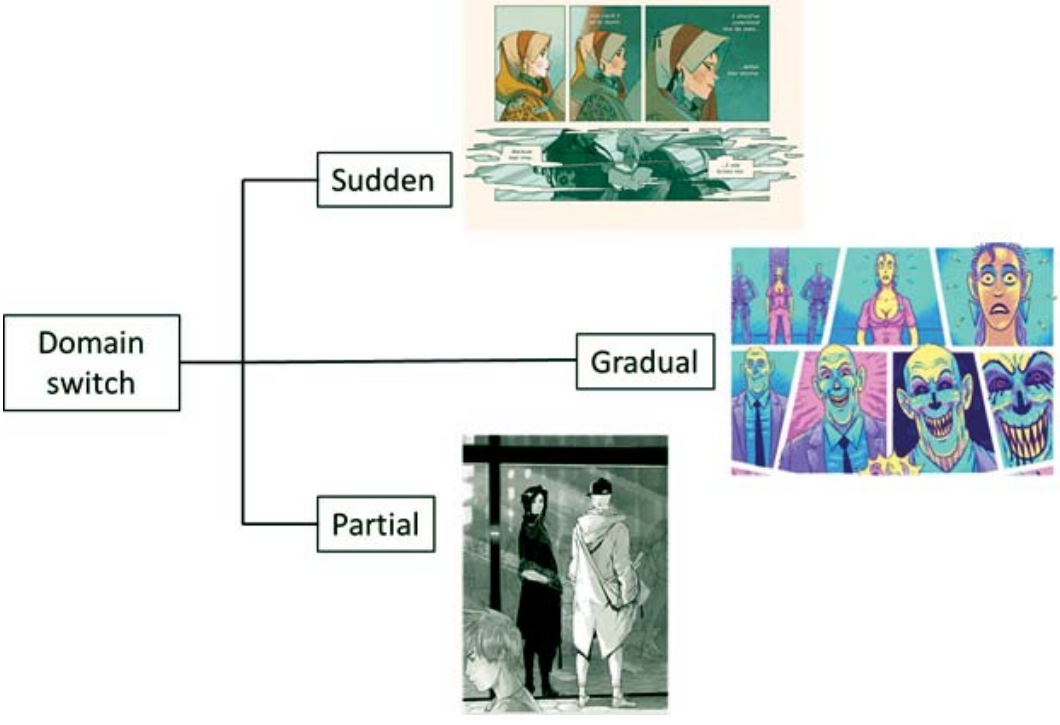

Figure 12. Illustration of the three domain switches: a Sudden transition, a Gradual transition, and a Partial transition.

Figure 12 summarizes how narratives may switch between their multiple domains. With a Sudden transition, layout constructions align with grammatical constructions, such that the narrative switches domains from one individual panel to another. In the example below, the narrative switches to a memory event across panels 3 and 4, with each domain reflected in its own panel. This type of transition would also apply to the two-panel constructions discussed by Abusch and Rooth (8). For Gradual transitions, the domain switch requires more than two panels. The example below illustrates how change across panels occurs progressively, with each subsequent panel amplifying the previous one until a certain end stage is reached (here, the woman realizing the man's malevolence). Finally, a Partial transition relies on graphological and morphological patterns rather than layout and grammar, since two domains are merged within a single panel rather than across a sequence, what previous studies term "blended-perspective" (Bimpikou 246; Maier and Bimpikou). Here, auxiliary elements are added to the primary domain, such that we assume the narrative remains in the established storyworld overall but overlays this scene with auxiliary aspects (the example here shows an Imagined reflection within an otherwise primary domain environment).

Discussion

Altogether, the Domain Constructions (DC) model describes how graphic, combinatorial, and conceptual structures contribute to construing incongruity in visual narrative sequences as depictions of auxiliary domains, such as character's (private) thoughts or other mental processes (e.g. dreams). Inferences of such specific predicates thus signal that the incongruity was likely intentional, meant to distinguish these events from (congruous) events belonging to the established, expected storyworld accessible to all characters, i.e. the primary domain.

The DC model expands upon prior (semantic) analyses of intentional incongruity by arguing that graphological, layout, morphological, and grammatical structures can all cue the conceptual interpretation of an auxiliary domain. For graphic structure, these include correspondences between meaning and colour schemes, (panel border) linework, or layout; within combinatorial structure, morphological strategies distinguish domain switches or connect seemingly distinct figures, and narrative patterns may align with domains. Interfaces across the three components then differentiate how domains are presented in a sequence, characterized in three types of domain transitions. In addition, these constructions appear together in visual sequences, inviting hypotheses related to the (mis)alignment of varying cues, e.g. that the alignment of multiple cues could facilitate the inferencing related to auxiliary domains.

As demonstrated in the range of examples throughout this paper, correspondences between form-related cues and meaning appear to be encoded patterns across visual narratives, emphasizing how incongruity forms a fundamental part of visual storytelling, in addition to congruity. Many visual naratives are analyzed for their contingency across discourse units (Maier and Bimpikou; Abusch), with discontinuity signalling event segmentation (Gernsbacher, "Two Decades" 5; Loschky et al. 322). However, the prevalence of incongruous information–semantic, graphic, and combinatorial–supports that such constructions are also central within visual storytelling. In addition, this also challenges the notion of a narrative as a temporal succession of events (Genette 25), as some depictions of experiences in domain constructions do not align with the established timeline (e.g. depictions of past/future events). While these can be construed as e.g. memories recalled in the current timeline, such inferences follow resolution based on (temporally) incongruous depictions. The question follows what effect such incongruity has on readers' narrative experience. As these sequences can still be comprehended well (Klomberg, Fadeeva, et al.) we speculate this complexity contributes to a sense of intruiging, artful storytelling.

Furthermore, if indeed visual narrative comprehension relies on such patterned structures, involving slots to be filled with appropriate content, it suggests authors and readers have learned these patterns, and that these might be modulated by visual fluency, as shown for other aspects of visual

languages (Cohn et al. 28; Cohn and Kutas 274; Cohn and Maher 79). Such learned constructions may have changed over time or differ across genres or cultures. Some constructions may even extend beyond comic narratives, depending on the affordances that these require. Colour differences, for instance, easily apply to film, while characteristics of panel borders remain a unique graphologial feature of drawn images. Overall, these questions provide relevant challenges for future corpus and/or experimental studies into the comprehension of domain constructions, within the studies of comics, discourse, focalization, perspective taking, and cognitive linguistics.

All in all, the DC model provides a framework for constructions of domains in visual narratives, where graphic, combinatorial, and meaning structures combine as formal parallel components underlying a holistic interpretation of incongruity within sequences. These structures capture the complexity of visual narrative storytelling and its patterned nature, and may easily connect to how domain constructions could operate across modalities. The implications of such patterns raise relevant questions for future work into visual narrative comprehension.

Tilburg University, Netherlands

Acknowledgements

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 850975).

Works Cited

Abusch, Dorit. "Applying Discourse Semantics and Pragmatics to Co-Reference in Picture Sequences." *Proceedings of Sinn Und Bedeutung*, vol. 17, 2012.

Abusch, Dorit, and Mats Rooth. "Temporal and Intensional Pictorial Conation." *Proceedings of Sinn Und Bedeutung*, vol. 26, 2022.

Andrew, Clark, et al. *Understanding Greek Vases: A Guide to Terms, Styles, and Techniques*. Getty Publications, 2002.

Bimpikou, Sofia. "Perspective Blending in Graphic Media." *Proceedings of the ESSLLI 2018 Student Session*, ESSLLI, 2018, pp. 245–57.

Clausner, Timothy C., and William Croft. "Domains and Image Schemas." *Cognitive Linguistics*, vol. 10, no. 1, 1999, pp. 1–31, https://doi.org/0936±5907/99/0010±0001.

Cohn, Neil. "Combinatorial Morphology in Visual Languages." *The Construction of Words*, edited by Geert Booij, vol. 4, Springer International Publishing, 2018, pp. 175–99, https://doi.org/10.1007/978-3-319-74394-3_7.

—. "(Pea)Nuts and Bolts of Visual Narrative: Structure and Meaning in Sequential Image Comprehension." *Cognitive Psychology*, vol. 65, no. 1, Aug. 2012, pp. 1–38, https://doi.org/10.1016/j.cogpsych.2012.01.003.

—. *The Visual Language of Comics: Introduction to the Structure and Cognition of Sequential Images*. Bloomsbury Academic, 2013.

—. "Visual Narrative Structure." *Cognitive Science*, vol. 37, no. 3, 2013, pp. 413–52, https://doi.org/10.1111/cogs.12016.

—. *Who Understands Comics? Questioning the Universality of Visual Language Comprehension*. Bloomsbury Academic, 2020.

—. "Your Brain on Comics: A Cognitive Model of Visual Narrative Comprehension." *Topics in Cognitive Science*, vol. 12, no. 1, Jan. 2020, pp. 352–86, https://doi.org/10.1111/tops.12421.

Cohn, Neil, and Marta Kutas. "Getting a Cue before Getting a Clue: Event-Related Potentials to Inference in Visual Narrative Comprehension." *Neuropsychologia*, vol. 77, Oct. 2015, pp. 267–78, https://doi.org/10.1016/j.neuropsychologia.2015.08.026.

Cohn, Neil, and Stephen Maher. "The Notion of the Motion: The Neurocognition of Motion Lines in Visual Narratives." *Brain Research*, 2015, p. 12.

Dijk, Teun A. Van, and Walter Kintsch. *Strategies of Discourse Comprehension*. Academic Press, 1983.

Duncan, Randy. "Persuasive Comics." *What Political Science Can Learn from the Humanities: Blurring Genres*, edited by R.A.W. Rhodes and Susan Hodgett, 2021, pp. 259–86.

Fauconnier, Gilles, and Mark Turner. *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities*. Basic Books, 2002.

Forabosco, Giovannantonio. "Is the Concept of Incongruity Still a Useful Construct for the Advancement of Humor Research?" *Lodz Papers in Pragmatics*, vol. 4, no. 1, Jan. 2008, pp. 45–62, https://doi.org/10.2478/v10016-008-0003-5.

Genette, Gérard. *Narrative Discourse: An Essay in Method*. Translated by Jane E. Lewin, Cornell University Press, 1980.

Gernsbacher, Morton Ann. "Coherence Cues Mapping during Comprehension." *Processing Interclausal Relationships in the Production and Comprehension of Text*, edited by J. Costermans and M. Fayol, Erlbaum, 1997, pp. 3–21.

—. "Two Decades of Structure Building." *Discourse Processes*, vol. 23, no. 3, 1997, pp. 265–304, https://doi.org/10.1080/01638539709544994.

Giora, Rachel, et al. "Weapons of Mass Distraction: Optimal Innovation and Pleasure Ratings." *Metaphor and Symbol*, vol. 19, no. 2, May 2004, pp. 115–41, https://doi.org/10.1207/s15327868ms1902_2.

Hacımusaoğlu, Irmak, et al. "Navigating Meaning in the Spatial Layouts of Comics: A Cross-Cultural Corpus Analysis." *Visual Cognition*, vol. 31, no. 2, Feb. 2023, pp. 126–37, https://doi.org/10.1080/13506285.2023.2198271.

Horstkotte, Silke, and Nancy Pedri. *Experiencing Visual Storyworlds*. Ohio State University Press, 2022.

Huff, Markus, et al. "Cross-Codal Integration of Bridging-Event Information in Narrative Understanding." *Memory & Cognition*, vol. 48, no. 6, Aug. 2020, pp. 942–56, https://doi.org/10.3758/s13421-020-01039-z.

Hutson, John P., et al. "Understanding Moment-to-Moment Processing of Visual Narratives." *Cognitive Science*, vol. 42, no. 8, 2018, pp. 2999–3033, https://doi.org/10.1111/cogs.12699.

Jackendoff, Ray. *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford University Press, 2002.

Klomberg, Bien, Irmak Hacımusaoğlu, et al. "Continuity, Co-reference, and Inference in Visual Sequencing." *Glossa: A Journal of General Linguistics*, vol. 8, no. 1, 2023, pp. 1-43, https://doi.org/10.16995/glossa.998.

Klomberg, Bien, Klavdiia Fadeeva, et al. *The Costs and Contours of Incongruous Identities in Visual Narratives*. 2023. Department of Communication and Cognition, Tilburg University.

Langacker, Ronald. *Foundations of Cognitive Grammar. Volume I: Theoretical Prerequisites*. Stanford University Press, 1987.

Loschky, Lester C., et al. "The Scene Perception & Event Comprehension Theory (SPECT) Applied to Visual Narratives." *Topics in Cognitive Science*, vol. 12, no. 1, 2020, pp. 311–51, https://doi.org/10.1111/tops.12455.

Magliano, Joseph P., et al. "The Relative Roles of Visuospatial and Linguistic Working Memory Systems in Generating Inferences during Visual Narrative Comprehension." *Memory & Cognition*, vol. 44, no. 2, Feb. 2016, pp. 207–19, https://doi.org/10.3758/s13421-015-0558-7.

Maier, Emar, and Soa Bimpikou. "Shifting Perspectives in Pictorial Narratives." *Proceedings of Sinn Und Bedeutung*, vol. 2, 2019, pp. 91–105.

Mikkonen, Kai. *The Narratology of Comic Art*. Routledge, 2017.

Ramachandran, V. S., and William Hirstein. "The Science of Art: A Neurological Theory of Aesthetic Experience." *Journal of Consciousness Studies*, vol. 6, no. 6–7, 1999, pp. 15–51.

Schilperoord, Joost. "Ways with Pictures: Visual Incongruities and Metaphor." *Visual Metaphor: Structure and Process*, edited by Gerard Steen, John Benjamins Publishing Company, 2018, pp. 11–47.

van der Gouw, Sharitha, et al. *Visual Storytelling: A Corpus Analysis of Domain Usage in American and Japanese Comic Styles*. 2022. Tilburg University, Master's thesis.

Willats, John. *Art and Representation: New Principles in the Analysis of Pictures*. Princeton University Press, 1997.

Zwaan, Rolf A., and Gabriel A. Radvansky. "Situation Models in Language Comprehension and Memory." *Psychological Bulletin*, vol. 123, no. 2, 1998, pp. 162–85.