Multimodal Cohesion in Panel Graphs:
A Pragmaticist Approach to the Gap Between
Comics Grammar and Aesthetics

Stephan Packard, University of Cologne

Abstract. The essay presents a Peircean or pragmaticist approach to the contradiction between two kinds of theories about comics interpretation: One set which assumes the existence of a fundamentally lingual grammar in accordance with concepts of 'visual languages', and another set which insists on a free artistic interpretation as the force grounding interrelations between comics panels in sequence and sense-making in comics. The approach presented supports both positions and attempts to reconcile them by explaining in which two different senses comics do and do not have grammar. It examines the perceived gap between the grammar and the aesthetics of comics in a quotidian example taken from the pages of *Amazing Spider-Man*. While admitting a possible hermeneutic corridor between regular and singular, 'heautonomous' interpretations in the terminology of Romantic aesthetics, this position then historicizes such accounts and looks for alternative treatments of the same cognitive processes. By moving from a grammatical through a hermeneutical to a pragmaticist semiotic account of the resolution of ambiguities and irritations in interpreting a short panel sequence, the argument pivots on the Peircean continua of continued semiosis, of the interplay of prescision and abstraction, and of the gradual differences between simple and creative abduction, to outline three distinct conclusions for a multimodal model of cohesion in panel sequences that straddles the seeming divide between grammar and aesthetics: First, that such a hermeneutic corridor can be elaborated as a specific kind of r e v i s i o n a l a t t i t u d e towards panel interpretation in sequence; second, that this allows the delegation of assumptions about conscious or reflective reading in favor of a comprehensive technical account of continuous interpretation; and third, that the historic contexts of previous interpretations have reasons to c o n f l a t e d i s t i n c t i o n s between pictorial grammar and aesthetics with distinctions between conscious and automatic reading as well as between script and pictures, but that such conflations may no longer have to hold today.

Keywords. Comics, semiotics, hermeneutics, multimodality, pictorial grammar

Zusammenfassung. Der Aufsatz präsentiert einen Peirceschen, pragmatizistischen Ansatz zur Auflösung der Spannung zwischen zwei Theorien der Comiclektüre: Einer Gruppe, die eine grundsätzlich sprachähnliche visuelle Sprache mit entsprechenden Grammatiken voraussetzt, und einer anderen Gruppe, die auf einer freien künstlerischen Interpretation besteht, die die Beziehungen in Panelsequenzen und die Sinnstiftung in der Comiclektüre begründe. Der vorgestellte Ansatz will beide Positionen bejahen und miteinander vereinbar machen, indem er erklärt, in welchen zwei verschiedenen Sinnen Comics eine Grammatik haben und nicht haben. Er untersucht die wahrgenommene Kluft zwischen der Grammatik und der Ästhetik von Comics an einem alltäglichen Beispiel aus *Amazing Spider-Man*. Zwar lässt sich von einem hermeneutischen Korridor zwischen regelgebundenen und heautonomen Interpretationen sprechen. Eine Historisierung dieser Position erlaubt es jedoch, andere Theoriebildungen über die so beschriebenen Kognitionen in Betracht zu ziehen. In einer Bewegung von einer grammatischen und einer hermeneutischen zu einer allgemeineren, pragmatizistisch semiotischen Konzeption der Semiosen, mit denen Ambiguitäten und Irritationen in der Comiclektüre durch Revision bearbeitet werden, geht das Argument zu einer Betrachtung dreier Kontinua im Sinne der Peirceschen Semiotik über: fortgesetzte Semiose, ‚Präszision‘ und Abstraktion, und der Übergang von einfachen zu kreativen Abduktionen. So ergeben sich drei Schlussfolgerungen für ein letztlich multimodales Modell der Kohäsionsbildung in Panelsequenzen, das die scheinbare Kluft zwischen Grammatik und Ästhetik übergreift: Erstens lässt sich die Spezifik einer r ü c k b l i c k e n d e n E i n s t e l l u n g in der Comiclektüre beschreiben, die sich zwar mit einem hermeneutischen Korridor vergleichen lässt, aber auch andere Deutungen zulässt. Zweitens lassen sich so Annahmen über die *b e w u s s t e   o d e r   r e f l e k t i e r t e* Qualität der Comiclektüre zugunsten einer umfassend technischen Beschreibung kontinuierlicher Semiose zurückstellen. Und drittens kann der historische Kontext anderer Deutungen die V e r m i s c h u n g konzeptuell zu trennender Distinktionen zwischen Grammatik und Ästhetik, zwischen bewusster und unreflektierter Lektüre sowie zwischen Schrift und Bild zwar motivieren, kann aber heute an Bindungskraft verlieren.

Schlüsselwörter. Comics, Semiotik, Hermeneutik, Multimodalität, Bildgrammatik

1.   In what sense do comics have grammar?

Studies on the regularities of visual language, especially by Cohn (2013 and beyond), have shown conclusively that comics are subject to cohesion, i.e. unifying principles that bind together their elements, beyond the mere logical coherence of their semantic reference: In other words, as we look for connections between several panels and their various parts, we no more have to start reasoning in a vacuum about the meaningful structures employed in the pictorial sequence than we need to consciously reinvent language through cumbersome code-breaking every time we decode a sentence. There are grammatical rules that go beyond, and indeed before, those implications of world knowledge, artistic appraisal, and sensemaking that are involved in the

active and deliberate inference of connections between separate depictions in the anatomy of the comics pages' assemblies. While the latter aspects of coherence all contribute to a general sense of cohesion, I will distinguish the two terms here and limit 'cohesion' (cf. Halliday and Hasan 1976) to relationships of elements within a media artifact that are formally and explicitly established. I will mostly focus not on (the visual equivalent of) lexical but on grammatical coherence, i.e. referential chains established in accordance with rules governing the syntactical arrangements of signs, belonging to the domain of legisigns rather than symbols in Peirce's terminology: These relations are mostly at issue when claims to a visual language in comics are defended or disputed. Crucially for the ensuing argument on historicizing hermeneutics, I will look at grammatical cohesion proper as encompassing qualities of artifacts rather than as cognitive interactions (cf. e.g. Menzel et al. 2017), setting apart some of the elements considered under the same name in recent treatments in multimodal linguistics (cf. e.g. Tseng and Bateman 2018; Stainbrook 2016), with which this account should however remain compatible.

However, many appraisals of comics' specific aesthetic devices have emphasized a distinction between rule-bound grammar and comics' cohesive structures. They describe the experience and cognitive performance of meaning construction between panels and panel elements as essentially different from the lingual comprehension of written and spoken words. While the latter is considered to be more direct, opaque, and determined, the former is proposed to be always already an intelligent and intense engagement with art and meaning (e.g. Groensteen 1999; Grünewald 2014; and most explicitly Grennan 2017, chapter 1.2.2). As Groensteen puts it: "La bande dessinée repose […] sur un dispositif qui ne connaît pas d'usage familier." (1999: 23) – 'Comics depend on a dispositive without familiar usage.' In this view, comics should be opposed to language. For lingual art, literary or poetic modes can be distinguished from everyday usage by their differing linguistic qualities. This opposes them to comics, which are, in the same view, always already artistic. They depend upon and elicit a strong interpretation, which not only goes beyond, but is liberated from conventional comprehension. It by no means eschews rules, but constructs rules beyond or even contrary to those taken from grammar. If this conception in which poetic modes could escape from grammar were true, then we might be tempted to accept that comics are understood in a manner fundamentally different to language; that this difference is due to the specificity of their pictorial code; and that this specificity is about a lack of grammatical regularity. In this case, a specific tension connected to the multimodal combination of script and pictures would be essential for understanding comics, and a theory that repeats that distinction would be indispensable for understanding comics comprehension. But how can this view be reconciled with the many empirical demonstrations of the grammar of visual language?

On the following pages, I want to outline the requirements for a semiotic conception of comics that can operationalize this seeming contradiction by clarifying the two senses in which we may say that comics have gram-

mar on the one hand, and that their poetics lack a conventional grammar on the other hand. Because I find the evidence for both claims to be overwhelming both in generalized arguments and perhaps even more so in many individual readings of specific examples of comics art, I believe we need to find a way to affirm both positions at the same time. Obviously, this can only be possible if we are talking about two different things in each case. The challenge is that these two aspects nevertheless seem closely connected and can hardly be satisfactorily described in separation. What I will explain as the h e a u t o n o m o u s poetics of comics cannot be conceived of as some merely additional quality that linguists may ignore, because it apparently engages and even interferes directly with the same structures of sensemaking that are involved in comics' grammatical regularities. Vice versa, any artistic appraisal of comics' devices that ignores their grammatical conditions will miss at least large parts of the picture. The interplay between the two positions, or so I will argue, reflects an actual tension between grammar and aesthetics in the comic books themselves.

I want to discuss this semiotic cleft from four angles – grammatically in terms of some of the categories proposed for comics' visual language (2), hermeneutically in terms of the Romantic history of the ideas involved (3), strictly semiotically in terms of Charles Sanders Peirce' generalized theory of signs (4), and as a specific kind of staggered and hybrid modalities in terms of recent concepts of multimodality (5). Through the course of this argument, the most general approach from semiotics should work as a means of transference between the divided conceptions of empirically founded grammar and interpretatively liberated hermeneutics, and a unified model of multimodal comics comprehension. Crucially, the Romantic question of conscious or sophisticated interpretation as opposed to conventional and determined comprehension can be dissolved from the point of view of a semiotic theory for which consciousness is epiphenomenal (cf. Baltzer 1994; Colapietro 2014) and, for any individual case of semiosis, a facultative addition to an otherwise independent explanation of sign processing. We need not know whether a comics reader knows what they are doing in order to describe what they do.

2.   Grammar: Panel elements and revision

It follows that the phenomena that I want to look at here are of a lower or middle range in terms of their aesthetic quality. They occur across comic book syntax, and relate to specific devices employed more so than unique ruptures in comprehension. This separates them from another sense of aesthetics or poetics that focuses only on rules broken or flaunted, phenomena that do occur in comics as they probably do in all communication but are separate from the debate of grammaticality examined here. I thus avoid any especially unusual or marked moments of aesthetic interruption and foreground minor but pertinent devices that constitute poignant shifts in understanding as a part

of the ongoing progressive comprehension of a typical panel sequence. Crucially, their effect might or might not be explicitly conscious for readers.

Fig. 1. An emergent pipe. Spider-Man battling the Green Goblin in ASM #122: 4.

Consider this page from *Amazing Spider-Man* #122 from July, 1973 (Fig. 1, Conway and Kane: 4). The two triangular panels in the lower half describe a quite sudden reversal in the depicted combat between Spider-Man and his nemesis, the Green Goblin – exactly the kind of moment that is expected to occur in such combat scenes in superhero comics with some frequency, and with some flourish, but without breaking the format. If it first appears as if Spider-Man is gaining the upper hand, pummeling the Goblin as he sits on his shoulders while they ride in the air on the Goblin's hovercraft glider, this changes in the second panel, in which Spider-Man collides with a large steel beam and is thrown off the shoulders of his antagonist. What interests me about this sequence is the manner in which this reversal redefines the semiotic quality of the green beam. For Spider-Man, it suddenly appears – one might describe the immediate connection as an exemplary i n d e x i a l experience, in which the steal beam draws Spider-Man's attention by immediately connecting with his forehead, painfully. But we are interested in the readers' semiosis, not the hero's. If we look back at the previous panel, the beam is undeniably already there. In fact, upon closer examination of the first triangular panel, we might even be uncertain about the exact placement of the beam, Goblin's head, Spider-Man's fist, and Spider-Man's head. It is the important function of the beam in the second panel that encourages such a renewed consideration of the first panel. Upon such reflection, we might find the left triangular panel to be somewhat ambiguous in this regard, drawing on a vagueness in Gil Kane's artistic style that is demonstrably felicitous, though again not necessarily consciously intentional (although it very well might be).
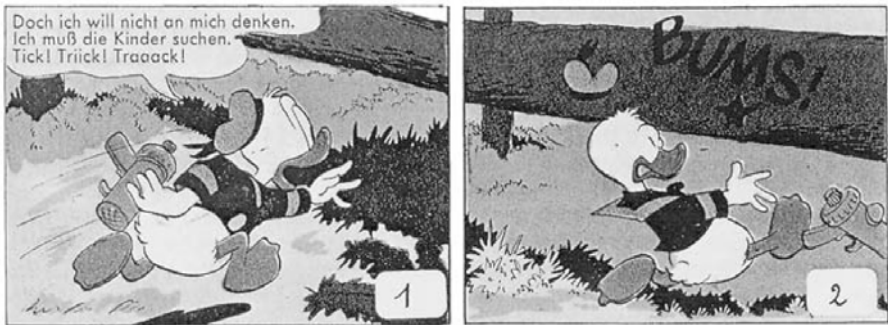

Fig. 2. An emergent branch. Donald running through a forest as cited in Krafft 1978: 48.

In one way, this example is certainly similar to a sequence that Ulrich Krafft, in his 1978 treatise on a textual linguistics of comic books, puts at the heart of his argument about different domains structuring consecutive panels. Here (Fig. 2, Krafft 1978: 48), Donald's head collides with a tree trunk in much the same fashion. Donald is surprised; but again, when we look back from the second to the first panel, we see that the trunk was clearly already there. In Krafft's model of panel domains governing referential movement,

this moves the trunk from the domain of spatial signs to object signs and even props, i.e. objects with which the characters directly interact. Carl Bark's art does not necessarily shift its depiction of the trunk from one panel to the next except for the fact that in the second panel, the contour of Donald's cartoonish character is directly connected to the circumference of the trunk; this is the main distinguishing stylistic marker for Krafft. Beyond it, the shift in domains in his theory is down to top-down ascriptions as readers make sense of the reappearance of elements across panels in continuous sequence. Functionally, the trunk in the second panel could not be erased without changing (or perhaps even destroying) the meaning of the picture; whereas spatial signs, according to Krafft, are specifically characterized by being subtractable. One might express a reversal of this idea (cf. Packard 2006: 207) by pointing out that comics spaces are productive, and allow for the emergence of object depictions in repeated spatial depictions.

The attention due the subsumption of graphical affordances under different panel elements reflects a number of claims about their functions in comics' comprehension, as well as possibly in their aesthetic appreciation. One first function throughout all similar theories is the precise description of different constituents of comics panels and pages. This would cover all attempts at a comics markup language, through to the most elaborate schemes for elemental annotation that can ground further corpus analysis (cf. Bateman et al. 2016). Krafft's model is more closely linked with one specific theoretical superstructure: distinguishing persons and objects across foreground and background, he arrives at four fundamental functions which are then taken up by the domains of action signs and spatial as well as canvas signs ('Handlungszeichen', 'Instrumente', 'Raumzeichen', 'Folienzeichen'; Krafft 1978: 15–41). While action signs are directly involved in marking the beginning, continuation, and suspension or replacement of phoric chains, spatial signs are only indirectly governed by such necessity. Krafft connects this logically to the way in which the reading process keeps track of signs: action signs can be enumerated, they are regularly repeated as long as the sequence keeps referring to their referents, and they are clearly delineated. Meanwhile, spatial signs can be added or eliminated with much greater liberty, their contours are often interrupted by panel borders, their internal structure of core elements and affixes is more liberally treated, and they are numerically indistinct: we know that there is only one Donald, or that there are only three of his nephews, in a given setup; but we do not expect the number of trees, bushes, or blades of grass to remain consistent. Surprisingly and hence persuasively, this even extends to architectural elements such as doors and windows, dissociating these rules of grammar from those of logic or even general concepts of salience. The specific salience of elements for each story told is upheld, tantamount to the concepts of focus or theme in other accounts of phoric movements. Any attempt at classifying domains within panels should be at least compatible with a concurrent theory of sequential construction or comprehension; additional reasons for this will become clear below.

A third possible use of panel domains as analytical categories bridges the intention to describe sequences and the goal of categorizing stylistic qualities. If action signs are more clearly and fully delineated and the contours of spatial signs do not need to be realized, even a single panel can suggest if not determine possible uses for different graphic elements. Here, Krafft's efforts to describe the compositional personhood of typical action signs (1978: 35) dovetails nicely with McCloud's famous emphasis on the differing drawing style of cartoons, usually employed as action signs, and less cartoonized backgrounds, which would usually be subsumed in Krafft's spatial sign (1994: 33–40). Both McCloud and Krafft, for differing reasons, consider an interim domain for objects which are taken up by, touched by, or interacted with by more fully realized agents depicted by action signs. Other models for the enumeration and distinction of types of panel elements and their interaction have suggested other motivations. A psychosemiotic model developed from Krafft's distinctions but founded on Peircean semiotics and Lacanian psychoanalysis has been presented (Packard 2006) for which cartoonized action signs, props, spatial signs, and canvas signs are all differentiated both by their function for the panel sequence and by their different interactions with the human gaze. Cohn has presented models founded on rich empirical accounts that continue the distinctions of Ray Jackendoff's parallel architecture model (Jackendoff and Audring 2020) into a six-part architecture comprised of external compositional structure, narrative structure, and event structure for the qualities of the sequence, corresponding respectively to a 'graphology', a 'visual morphology', and a lower level conceptual structure for the classification of units (e.g. Cohn and Schilperoord 2022). Notably, all of these models include a similar distinction between action and spatial signs, or more generally, activated signs and others. To the extent that visual language has a grammar beyond semantic comprehension, this distinction may be related to that between persons and objects, but cannot be reduced to it.

It is thus possible to think of the example in Krafft as a straightforward use of a comics grammar. In his model, the elements of each panel slot into specific domains, and the referential movement here is established precisely by shifting one element from one domain to another. All of this is also true for Kane's art on the *Spider-Man* page. However, on that page, several additional elements prepare and echo the reappraisal of that steel beam. First, the dynamic cut of the triangular panels assembled to form a rectangle on the lower half of the page draws attention to the different angles at which the beam appears. In fact, the two depictions of the beam and the diagonal panel gutter combine to roughly trace the shape of the 'Z-path' described by the gaze in a regular panel sequence comprehension in this genre (cf. Cohn and Campbell 2014). We rediscover that path, previously merely a transparent rule for grammatical comprehension, as a now-topical element of interpretation. This is not true for the trunk that stops Donald, whose shape and orientation in panel 2 is repeated almost identically from panel 1. Moreover, the three panels in the upper half of the *Spider-Man*

page distinctively and probably deliberately eschew clear spatial informa-
tion. The third panel shows a close-up of the combatants' faces; the sec-
ond panel arranges an overstated trail of smoke and condensation to fill
the space of the panel with the path taken by the hovercraft, as opposed
to filling in other markers that would place that movement in a relationship
with the rest of the world. Crucially, the first panel takes up the previously
established architecture from which the combat erupts briefly and in a com-
pressed fashion, without giving any hint to the relative placement and later
appearance of the beam.

One function of this first panel is to move the scene away from the dom-
inance of the space used in the previous sequence(s). On several pages
in this and the previous issue, the cliffhanger that connects the episodes
revolves around Spider-Man kneeling on a platform, Gwen Stacy's dead
body on his lap. This issue takes up the plot after the now-famous, then-sur-
prising death of this central character in *Spider-Man*'s continuity. In a par-
allel break with narrative expectation, Spider-Man has vowed to kill the
Green Goblin in response to the murder of Stacy. The scene examined here
seems poised to end with that very killing. The movement away from the
open and clearly laid out platform to the cluttered architecture in which a
steel beam may suddenly appear thus parallels the movement of the plot,
which suspends and retards the following action to give Spider-Man some
time to struggle with his revenge. In the eventual resolution (Fig. 3, Conway
and Kane 1973: 20), a similar play with spatial and object signifiers has the
Goblin's glider surprise Spider-Man from behind and kill the Goblin just as
Spider-Man is still trying to decide whether to kill the Goblin, saving Spi-
der-Man from making that decision and resolving the plot of this revenge
tragedy without a final commitment from the protagonist. Here, the move-
ment of the hovercraft in the fourth panel encourages re-examination of the
second panel, where a similar graphic ambiguity as before is at play regard-
ing the distance between the hovercraft and Spider-Man's back, and the
angle at which it approaches the Goblin.

In both cases, Kane's pages offer elements that connect to a re-apprais-
al of the beam or the hovercraft as spatial or object markers respectively.
Where Bark's panels convey the suddenness with which Donald connects
to the trunk by first hiding it in the spatial domain and then taking it into the
object domain, Kane's pages are apt to be re-considered at length precise-
ly with this question in mind: Where did the beam come from? What makes
its appearance so sudden? And eventually, the answers to those questions
lead into the deeper meaning of the plot. These ultimate consequences for
the interpretation of the revenge tragedy are of course not within the scope
of a panel grammar. In contrast, the distinction of spatial and object mark-
ers very likely is grammatical. But the revision invited by the elements par-
alleling the domain shift for the steel beam on Kane's page would be diffi-
cult to explain with any degree of precision without these grammatical cat-
egories, and yet goes beyond a straightforward referential movement. The
panels have grammar, in the sense that the graphical element depicting the

beam shifts between grammatically defined domains, as does the depiction of the trunk. But their additional interrelations with surrounding elements contextualize this shift, and do so crucially on the same formal and elemental level, by operationalizing the vagueness involved in the assignment of pictorial elements to panel domains, as well as the subtractability or productivity of comic panel spaces. The panel sequence has a grammar in the sense that it regularly distinguishes spaces from objects; its poetics are heautonomous in that they recreate and reappraise that rule through different aspects of the 'iconic solidarity', as Groensteen (1999) describes it, that binds these images together.

Fig. 3. An emergent solution to a tragedy of revenge. Spider-Man failing to kill the Green Goblin in ASM #122: 20.

3.   A hermeneutic corridor in comics comprehension

If we want to connect this element of comics interpretation to established vocabulary in the humanities, we might consider it a h e r m e n e u t i c aspect that comes between the scopes of g r a m m a r on the one hand and either a e s t h e t i c s or r h e t o r i c a l n a r r a t i v e on the other. In his influential account, and in a context of interlacing concepts between Idealism and Romanticism, Schleiermacher (1809/10) has presented the practice of hermeneutics in relation to – but not completely convergent with – grammatical comprehension. Hermeneutics famously begins with the fact of not (yet) understanding, and hence consists in a conscious and reflective engagement with the material. At the same time, the grammaticality of the text interpreted is not just assumed, but becomes one of two structures – alongside the technical or psychological dimension following expectations shaped by genre or authorship – involved in that conscious reconstruction. Schleiermacher speaks of a g r a m m a t i c a l i n t e r p r e t a t i o n in as much as it r e - considers lexical and morphological elements not understood fully at once, in order to better understand them eventually. Thus hermeneutical interpretation is distinguished from direct grammatical comprehension, but not from grammar. In a similar sense, the specific devices pivotal to the appraisal of comics' art might be considered to be realized in a hermeneutic corridor between rule-bound grammar and 'heautonomous' aesthetics in the sense of Romantic and some Idealistic accounts of aesthetics that have been continued and evolved from 18th and 19th century notions. This notion of heautonomy denotes the autonomy of ordered structures that ground the rules of their order within the confines of their own architecture rather than merely instantiating external rules established elsewhere and holds for poetic language whose internal regularities go beyond and sometimes even against standard grammar (cf. Homann 1999). Crucially, neither half of the semiosis of comics can be fully explained without a reference to this corridor between them.

It might be useful to point out what this corridor, and the distinct parts of comics' semiosis between which it mediates, are not. Their distinction is not the same as the (nonetheless relevant) distinction between top-down and bottom-up subsumption of graphical elements to grammatical constituents. When Krafft traces the function of Donald's cartoon, the pertinent tree trunk, and the other flora depicted in the Barks sequence by separating subtractable from non-subtractable elements, he argues top-down; when he points out the continued contour of Donald's and the trunk's circumference as well as Donald's regular arrangement of bodily constituents, he argues bottom-up. Both arguments pertain to comics' grammar. Conversely, when we find the arrangement of elements in Kane's first triangular panel ambiguous, we argue bottom-up; when it is the arrangement of both triangles that draws our attention to this element, we argue top-down. Both movements are involved in the hermeneutic circle, which travels from the general to the particular and back as many times as needed, and this oscil-

lation further characterizes this part of the practice of reading comic books as properly hermeneutic.

Equally, while this interpretation agrees with accounts of comics' special aesthetics that foreground heautonomous semiotic modes rather than grammatical semiosis, the border between grammatical and heautonomous aspects does not coincide with that between pictures and script, or between mere script and a text-image-combination. This is in contrast to several accounts (e.g. Grünewald 2014), and especially some treatments in critical theory, that conflate the semiotic opposition between grammar and heautonomy with the distinction between text and images, and sometimes between semiotics and meaning altogether (e.g. Frahm 2010). Text and image are inarguably combined in all examples discussed here; but their combination is not intrinsic to the re-examination of pictorial elements proposed. Many comics employ no script whatsoever, but those 'pantomimic' comics would still be subject to the hermeneutic interplay between a visual grammar and aesthetics. Any model of multimodality involved in describing the use of comic books will have to separate the combination of text and image from the interplay of heautonomous and grammatical semiotic modes, finding both kinds of modes realized in the graphical elements of comics' pages (and likely also, independently, throughout the scriptoral elements and in their combinations).

Thirdly, this interpretation does not allow for a parallel reduction of grammar and aesthetics to two independent parts of reading practices – and hence, it also does not allow for a separation of analytical interest into a linguistic explanation of visual grammar on the one hand, and an aesthetic appraisal of further interpretability on the other. Instead, the elements constructed in the course of this hermeneutic semiosis make sense only in the interplay between reflexive grammatical and reflective heautonomous subsumption – the latter relating to the rhetorical problem of free hypotyposis as much as to issues of regular grammatical comprehension.

Finally, the gap described between comics grammar and aesthetics is not a distinction between rule-bound and irregular structures. While some irregular structures may have aesthetic effects, the central concept of aesthetics intended here, as in Groensteen's argument, is one of over-determination rather than indetermination, of an abundance of rules as opposed to an absence of rules. Accordingly, Jakobson has described this approach for the formalist and structuralist traditions when he referred to the 'poetical' as tantamount to a dominance of the 'grammatical' function (2010: 25). As the necessities of grammar demand relationships between elements of language uttered in linear sequence, and poetry creates additional, similar, but sometimes differently structured and always grammatically redundant similarities, the aesthetics of comics sequences and panel structures offer additional dependencies between their elements: in a structuralist sense, rendering the communication more grammatical than grammar demands. It is because of these additional interdependencies between elements that the hermeneutic look backwards from one panel towards a pre-

vious one is such a telling and typical movement for an aesthetic compre-
hension of comic book panels. It is after the second appearance of the trunk,
the steel beam, and the glider that their first appearances are subjected to
additional interpretative rules: Were they background objects or introduced
as relevant props, was their placement in line with the later collision, did
their presentation hint at, allow for, or exclude the way that the next panel
continues their depiction?

In parallel, heautonomy, here, goes beyond any mere autonomy of art
that would understand aesthetic objects as unregulated. It is intended not
only to codify the fact that art is not entirely bound by the rules of everyday
communication, but to describe the way in which individual pieces of art
introduce new rules, sometimes by repetition within their own confines, but
sometimes also by a constructed evidentiality in singular usage. Historical-
ly, the concept is equally connected to 18th and 19th centuries' Idealism's
engagement with seeming intractabilities of aesthetics (cf. Homann 1999),
in which it is the power of the genius, to which the sustained tension between
having and not having rules is eventually ascribed. This occurs in its histor-
ical context, in a Romantic interpretation of a democratization of aesthet-
ics for the bourgeoisie (cf. Rancière 2008: 27). Some of the emphatic state-
ments of irregularity and its interpretation in comics art seem to connect to
those evaluations of the worth of the art involved: If the comprehension of
comics is down to an ultimately inexplicably irregular implementation of per-
sonal insight or taste, then the practice of that engagement becomes valu-
able in the sense of art's value for the *Bildung* of an individual person-
ality, i.e. the free development of an individual mind and personality in accor-
dance with the same Romantic conceptions. This conforms with the con-
nection often drawn between this appraisal of comics and their use in didac-
tics and pedagogics (e.g. Grünewald 2014).

The historical and political context of the semiotic tension can illuminate
its structure and its historical evaluation, but does not deny its current poten-
cy. Perhaps we would view the relationship between grammar and aesthet-
ics differently, if it were not for this place in history in which comics current-
ly exist. But it is in this place that we find our work situated, and the analy-
sis of the tension between grammar and aesthetics challenges a multimod-
al theory of comics formally as well as historically. One way to face this chal-
lenge, then, is to affirm both the existence of one or more grammars for
comics and the importance of heautonomous aspects in individual comic
books, and to posit a hermeneutic corridor as part of the reading practice
in which the two are engaged in a circle of mutual revisions. In traditional
terms, the distinction between the hermeneutic and the merely grammati-
cal mode of interpretation would be understood as subjectivity or reflection,
qualities of the Romantic concept of the consciousness. Notoriously diffi-
cult to prove alongside individual feats of cognition in empirical studies, or
to describe alongside strict formal descriptions of grammatical elements in
multimodal linguistic analysis, that concept might be less well suited to con-
temporary interdisciplinarity. If it is true that its importance in this context

stems from attempts to connect such conscious cognition to class rather than to ubiquitous individual abilities, we might have additional reason to deny it further emphasis. But semiotics allows for an alternative.

4.   A semiotic approach: Reconsidering grammar, heautonomy, and abduction in a framework of multimodality

A complementary approach to face the same challenge moves beyond the historical language of its Romantic foundations by engaging with the continued work on a generalized theory of semiotics based in the ideas of Charles Sanders Peirce. As Bateman has recently argued (2018), Peircean semiotics are intrinsically multimodal. I believe that they excel through the precision with which they deal with multimodal differences and combinations. Not only is Peirce's semiotics generalized to the point where it can apply to pictures as easily as to script, but it allows for a description of each semiotic mode's differences and distinctions. Crucially, it allows for a definition of semiotic modes by which they are "generally far 'smaller', or more 'fine-grained', than those assumed in many other positions in multimodality studies, [while] [a]t the same time [...being] also 'broader' in the sense that they are not restricted to lie within single *perceptual* modes" (Bateman 2018: 19).

This aptly frames the distinctions proposed in the previous sections: Assuming one or more grammatically structured, as well as one or more heautonomous semiotic modes, we will find that these modes are more finely grained, on the one hand, than the monolithic dichotomy between script and pictures, with grammatical and heautonomous elements appearing both in script and in pictorial elements without governing any of them completely. At the same time, they are broader in a different sense, as they do not depend upon the distinction drawn by modes of perception: whether the immediate sensual affordance of a material artifact entices us to view a picture or to read writing, we will still find ourselves with grammatical as well as heautonomous elements in either mode.

The distinction nicely mirrors the debate on how to define codes as well as the various levels of modes within the description of any modality (cf. the summary in Wildfeuer et al. 2020, esp.: 135–143). It is founded on the now well-established distinction between material and semiotic properties, by which a comic book printed on paper is materially different from its scan appearing on a computer screen, but both may offer many of the same important material affordances for semiotic use. In the examples that interest us, each comic offers material for the same perceptual modes, i.e. the graphic and the literal (and perhaps more). In addition, we might distinguish different semiotic modalities, which are not limited to, but include different explicit or implicit codes and their conventional rules. Simply, if a comic includes script in English as well as in German, we might say that the material affordance and the perceptual modes as well as the semiotic modes

are identical, but the conventional code employed respectively differs. Less simply, the same logic might apply to the graphic elements of a comic book. If Cohn is correct in distinguishing a North American from a Japanese visual language in comics and manga (2010 and 2016), one might imagine a combination of several such modalities within the same sequence of panels. This might be the case, for instance, when an American super-hero book employs an 'exploded' style to depict a sequence of movements and poses in great detail, integrating a device popularized by manga into an otherwise differently styled artifact.

In this sense, the claim that comics have grammar would refer to the two (or perhaps even more) grammars in action. This clearly separates the specific grammar in use from the hermeneutic distinctions made in the examples above, as it is not a difference between two different visual languages that separates the depiction of a trunk or steel beam as a spatial or as an object marker, but rather different subsumptions under the categories involved in any one such language. The fact that spatial and object markers are involved, and mutually distinguishable at all, might be a candidate for a convergence between different comics grammars, or point towards some universals of visual language.

One important consequence of this concept of grammaticality is that it easily and quickly leads any analysis into the distinction of different elements within panels, as opposed to the description of panel sequentiality alone. The scandal for classical philosophies of the image is the ease with which this introduces the otherwise controversial idea of an intra-pictorial syntax. As Sachs-Hombach (2013: 125) has argued in dialogue with the more general semiotic vocabulary of Goodman, images might well be considered syntactically 'dense' in a sense that invites the ascription of several different functions to one image, but discourages the disassembly and precise delimitation of several different functional elements within a panel. It is worth noting that Sachs-Hombach immediately connects this presumed quality of pictures to the fact that they do not engage in specific syntactical combinations in context, reserving the option for more complex combinatorics and complexes of graphical elements beyond single images, and its assorted analogues of lingual grammaticality (2013: 140). But it is the existence of exactly such syntactical relations between separate images within the panel sequence as well as more far-reaching, and still more complex, inter-panel relationships, that drive the top-down disassembly of panels in Krafft's text-linguistic model of comics comprehension and lead him to identify some bottom-up qualities typical for elements with specific functions in a next step. It is precisely by pulling apart the seemingly dense interrelations between elements w i t h i n each panel that he grounds his explanation of the interaction b e t w e e n panels.

This might prove important, as empirical work to demonstrate the existence of a grammar in comics has hitherto been more successful for page or sequence structures between panels than for the internal structures of panels (cf. Cohn et al. 2012, Cohn and Campbell 2014; Bateman et al. 2018).

This suggests that more detailed work on grammaticality within panels and across panel borders is needed in addition to the existing insights into panel inter-relations, and that the hermeneutic gap which invites conscious reconsideration of previous panels' elements might be co-emergent with aesthetic, as well as grammatical codes that dissect panels into such functional elements in the first place.

Perhaps the most striking proof so far for the existence of grammaticality between whole panels, as separate from logical semantic inferences between their contents, lies in readers' ability to distinguish between re-arrangements of panels, e.g. from *Peanuts* strips (Cohn et al. 2012: 6), which are altogether scrambled, from those re-arrangements in which each panel has been taken from a different strip, thus destroying all semantic cohesion, but in which the relative placement of the panels is maintained. This in turn sustains a syntactical cohesion that need not make sense, but constitutes the panel equivalent of those 'colorless green ideas sleeping furiously' – a sentence that is famously cohesive grammatically while logically incoherent (famously employed in and since Chomsky 1956: 116). The example also happens to showcase one stark difference between lingual and visual languages: While readers show the passive ability to sanction violations of grammaticality in the strips they are shown, they are not necessarily able to produce the obvious correct alternative. *Colorlessly greenly ideas sleeps* is easily corrected to *Colorless green ideas sleep*, but the corrected version of a scrambled panel sequence is less easily determined in theory and much less easily drawn by most readers in practice. Similarly, once corrected, the sentence *Colorless green ideas sleep* might be translated into German or French with a high albeit not absolute expectation for convergence among translations, whereas an attempt at translating a structurally correctly formed *Peanuts* strip from North American into Japanese visual language would present a much less convergent result, if any. Certainly one may come up with parallel constructions to a Schultz strip while employing Japanese visual language, but there are far more such constructions and there is far less determination as to the one obvious solution than in comparably simple lingual translation tasks, owing only in part to the density of graphical realization beyond the grammatically necessitated aspects, and the strength of iconic semantic references.[1]

If this points to a relative lack of productive determinedness for the grammar of comics as opposed to typical grammars in spoken and written language, this might still give further weight to the importance of the hermeneutical corridor, in which irritation is interpreted rather than corrected. This would also conform with empirical results that show that backtracking is more prevalent in comics than in written narrative, suggesting that revision upon irritation – the core hermeneutic practice – plays an especially large role (cf. Foulsham et al. 2016; Kirtley et al. 2023). In the two examples from Kane's art above, the vagueness of the relative arrangement of Goblin, Spider-Man, and the steel beam, or of Spider-Man's back and the Goblin's hovercraft, in the respective earlier panels is turned into an operative ambigu-

ity to motivate the ensuing shift in panel domains, and from there, the pivotal moments for the plot. An alternative would of course have been, to criticize and perhaps correct an infelicitous drawing that fails to clearly convey the relative positions of objects in space in each depicted scene. But the sustained abstract grammaticality of the panel sequence, quite independently from its representative function, allows, and the hermeneutic interpretability motivates, an acceptance and appraisal for the artistic choice.

All of this allows for a translation of the general attitude of hermeneutics into a specific semiotic aspect: Foregrounding the need for revision by backtracking, reviewing, and comparing material backwards as well as forwards constitutes a circular movement that introduces and revises candidates for heautonomous rules as well as hypothetical subsumptions of elements under conventional rules of grammars. Backtracking, reviewing, and comparing are of course elements entertained in most grammatical accounts of comprehension, and some of the most pertinent examples for the aesthetic devices under scrutiny here might also be described as 'garden-pathing', as 'discourse pops' (cf. Wildfeuer 2013), or in similar fashions. What is at issue here, is whether an account of the specifics of this movement in comics' aesthetics allows for a purely grammatical explanation that nevertheless remains compatible with the ascription of aesthetic qualities to the same structures by some readers. This requires attention to at least three distinct continua as proposed by Peircean semiotics: As a framework, a Peircean view must insist on a continuum of ongoing reasoning as part of any semiosis, as ideas about connections and meanings are progressively 'filled in' on the path from an immediate through a dynamic to a final interpretant. Then, more specifically, a second continuum must be assumed between granular and summary treatments of parts of semiosis, in which each element may be broken down into several smaller elements or integrated into greater wholes as needed, a balance which Peirce has referred to as the 'prescision' and 'abstraction' of ideas. Finally, both of these continua also draw our attention to modes of conclusion other than fully determined deductions, especially by induction and the continuum between simple and creative abduction.

I will return to the three continua in a moment. A necessary preliminary logical step to realize such a fine-grained approach to rules across all three, is to redifferentiate what it means for rules to be conventional. Heautonomous rules by definition are not previously established conventions. But even rules of grammar may be conventional only in some of several conceivable ways. We should at the very least distinguish conventional syntax for the shape and arrangement of signifying elements (regularly formed and arranged 'representamens', i.e. 'legisigns') from the most commonly discussed conventional connection between signs and their semantics (regular object relations, i.e. 'symbols'), and each again from the perhaps most important conventionally rule-bound integration of signs into greater wholes (regular 'interpretants', i.e. 'arguments' in Peirce' diction). Crucially, as Bateman (2018) also emphasizes, the trichotomy of representamens and

the trichotomy of object relations are not entirely independent from one another – and hence, symbolicity and conventionality are n o t entirely orthogonal.[2] That is to say, while iconic signifying relations that operate through some similarity between a representamen and their content may be grounded on qualisigns, which convey a pure sensational quality, equally well as on legisigns, which are formed in accordance with a syntactical rule of the code employed, the reverse is not true: A symbolic sign, which refers to its object by some conventional rule, can only do so if it is a legisign, i.e. formed such that it may be recognized as the signifying element that satisfies the rule in the first place. In other words, while symbolicity is conventionality for sign-object relations, it is 'legisignality' that constitutes accordance with rules on a purely sensual or aesthetic level; and symbolicity always entails legisigns, even if legisigns need not be used as symbols.

A purely aesthetic approach to comics, if taken to an extreme, might focus on the same graphical elements merely as 'sinsigns', i.e. singular tokens in a usage that does not consider their regular repetition.[3] Similarly, the integration of an icon or a symbol into a fully determinate interpretable unit might include that icon or symbol in what Peirce refers to as a 'rhema', a 'dicent,' or an 'argument', depending on whether a statement is made as hypothetical across possible cases, as indicative of a given case, or as accordant to a general rule for all pertinent cases. Once again, any symbol and any legisign can be integrated into any of these semiotic triads, from rhema through dicent to argument. But if the interpretation is to arrive at generalized statements or arguments, the signs involved have to be legisigns and symbols, not mere qualisigns or icons: Only when a sign is recognizably formed in accordance with a certain rule and then connected to a certain signified object in accordance with its rule of interpretation, can it express a proposition about that interpretation that is itself rule-bound in a generalizable fashion.

This can mean forming a legisign in accordance with an exact model, such as is the case with even a singular replica of a letter from an established alphabet, used to express what that letter should look like. In the cases of legisignality in the examples discussed so far, it is not the exact shape of a picture of a trunk or beam that is predetermined, but it is its relational usage across arrangements of panel elements that accords to one or more visual languages. The shifting functions ascribed to tree trunk, steel beam, and hovercraft, as each moves from a spatial to an object marker, all assume recognizability for these elements (which is why we can name the graphical units by their semantic lingual equivalents in the first place). To say that tree, beam, or craft are in the situative background in one panel but interact with the plot-driving characters in the next panel, is to first be able to r e c o g n i z e these elements across panels by their r e g u l a r shape and arrangement, i.e. because they function as legisigns, which allows for but does not demand their iconic relationships to the objects denoted.

The salience of this point might become even more clear when we consider the alternative. An alternatively constructed semiosis might connect the two instances of tree trunks, steel beams, or hovercrafts across two panels not as two tokens of the same type of legisigns, but merely as the icons which they doubtlessly also constitute. This would read the second occurrence of each element as an iconic sign for the first and/or an iconic sign of the same referent, i.e. associate the two depictions of the same thing in two panels by their similarity without reducing that similarity to an implicit rule. In other words, we know that the trunk in panel 2 is the trunk in panel 1 because they each look like a trunk and like each other. And, of course, that is exactly the case. But because each picture of a tree trunk is similar to the other depiction of the same trunk, and also to an actual tree trunk, this runs the risk of conflating the repetition of signs with their iconic motivation, i.e. confusing syntax with semantics. We a l s o recognize that the trunk repeats syntactically in that place and in a pertinently shifted function. Iconicity is a possible, but not the only involved relationship: In each specific case, an element such as a tree trunk will be partially recognizable across panels because each instance is iconically similar to the appearance of an actual tree trunk, and yet the certainty with which the element is recognized as the s a m e tree trunk will draw on similarities between each depiction, which are not merely iconic relationships, but qualities of legisigns by which the first instance imposes limitations to the variability of further instances.

The semiotic principle involved here echoes the distinction between arguing that some lexical elements of language are iconically motivated, and that lexicalization does not govern them: *Cuckoo* and *Kuckuck* sound the way they do for a reason, but each language has recognizably different conventionalizations for their sound. Each sound might be interpreted as iconic by a person unacquainted with English or German, but to recognize the German or English word is an additional feat, establishing that a certain legisign is used, which in turn may carry a conventional, i.e. symbolic meaning. In parallel, each depiction of a trunk is motivated by the similarity in its appearance, but comics' syntax identifies the subsequent elements in subsequent panels through additional conventions.

What is at stake here for semiotic theory becomes clear when we compare this approach to the far more limited view of semiotics that considers only sign-object relations, and hence only recognizes iconic references from depiction to referenced object in panel sequences, denying the possibility of a visual grammar built on legisigns as opposed to mere icons. Strong versions of this latter view appear in critiques of semiotics in several phenomenological accounts. The most prominent example in comics studies might be in Lambert Wiesing's repeated (e.g. 2008) emphasis on the aesthetics of the speech bubble as the seminal moment in the history of sequential graphic art. In this account, comics become their own art form when Outcault interrupts the densely iconic depiction of scenes with 'rifts' in the canvas that contain written words. Having chosen the limited framework for

pictorial elements' interrelations, Wiesing has no choice but to once more return to the combination of text and images as comics' defining aesthetic feature: his denial of internal semiotic structures and intra-pictorial syntax has taken regular relationships between pictorial elements out of the argument. On the contrary, a visual morphology that allows for but is not limited to the iconic mode for graphical elements will assume "open-class lexical items" which "easily allow for new patterns to be created," and which can explain why in "visual form, these are typically," but n o t necessarily,

> iconic representations: it is easy to create a novel schematic pattern for iconic elements, based on the way they look (Cohn 2018b: 3).

The first continuum involved in such a semiosis is then built upon, and can be analyzed as, a series of conclusions, of deductions as well as more fragile induction and abductions. For our first example, one possible d e d u c - t i o n would be to recognize the unfinished contour of the steel beam as typical of what Krafft considers a spatial sign, and in a second step come to the conclusion that this is a spatial sign by a b d u c t i o n : If the steel beam is a spatial element, its contours need not be finished; its contours are unfinished; that might be the result of it being a case of a spatial element. Note how the parallel operation could be described as a consecutive deduction if the rules were certain. Assuming, then, that unfinished contours are only allowed for spatial signs, and recognizing that the contours of the steel beam are unfinished, we would h a v e to conclude that the steel beam is a spatial sign. But the actual uncertainty of the assignment negotiates grammatical relations considered across panel domains. This marks the previously described difference between the vagueness of comics' rules as opposed to language: *Colourless green ideas sleeps* contains a definite incongruity between the plural *ideas* and the predicate *sleeps*. Comic books' panels propose more probable assignments alongside definitive ones.

This fits Cohn's and others' employment – not of a strict generative grammar in a Chomskian sense – but of a probabilistic construction grammar whose empirical foundation does not expect that each competent comics reader will arrive at completely identical conclusions in every case, and in which lexicon and grammar are not entirely separated in their functions (e.g. Cohn 2018a). Crucially again, comics readers not only may arrive at different hypotheses, but they know that, and in which regards, they might do so, and can thus allow for ambiguity as a fact about a given communication, rather than a problem to be dissolved. As the visual linguist engages in an induction across as many possible readings of as many possible instances as they can empirically analyze, this account assumes that in hermeneutic reconsideration, each individual reader will engage in a similar inductive movement, considering real or imagined cases of parallel arrangements of panels and panel elements, to double-check their preferred interpretation against each individual case in the mode of the grammatical interpretation in Schleiermacher's hermeneutic circle.

A Romantic interpretation would make this consideration the hallmark of an educated interaction with the material. However, the second continuum taken from Peirce's descriptions offers an account that need not be explicit about assumptions of conscious or unconscious processes in this feat of sensemaking. As described in his seminal paper *On a New List of Categories* (1868), detailing the continuum of semiotic inference allows and even necessitates the distinction of two movements which summarize all that the hermeneutic circle accomplishes, in our instance without having to introduce the idea of consciousness as their specifying quality. Rather than assuming any one proper segmentation of ideas into propositional units, from which only reflective reasoning can hypothetically deviate by reconsidering whether what has been understood subliminally has been understood properly, it is on the contrary the very "function of conceptions to reduce the manifold of sensuous impressions to unity" (Peirce 1868: 278), and that function can be fulfilled in different parallel ways and is a l w a y s open to reconsideration with or without conscious reflection.

The general reason for this is that the granularity of the reduction is shifting and reversible. To adapt Peirce's example, I might consider a black stove a unit and add to its idea its placement at the far end of the kitchen; or consider the unit a stove, to which its color is added by predication; or consider the black stove at the far end a unit, for which I may then predicate its owner, etc. The far reaches of this movement are congruent to the general philosophical accounts of substance and being in Peirce's view, but conceiving of their opposition as continuous allows for a clear understanding of the logical movements between them: Movements which Peirce terms 'prescision' when they occur in one direction, and 'abstraction' in the other. If we follow his terminology,

> the terms 'prescision' and 'abstraction,' which were formerly applied to every kind of separation, are now limited, not merely to mental separation, but to that which arises from *attention to* one element and *neglect of* the other (Peirce 1868: 289).

The attention, or lack thereof, is the echo of the mind at work in more essentialist accounts of cognition, and becomes one dimension of secondness or volition in the semiotic account. But that volition is at once removed from the limitations to genuine human consciousness or 'mental' operations of separation. Rather, it allows us to think of prescision and abstraction as movements in any progressive semiotic process. Whether the steel beam, trunk, or hovercraft is to be summarized among the elements of the spatial sign, or grouped as a circumscribed unity of its own among the objects interacting with agents, is then due to just such a mutual movement of possible prescision and abstraction. It is a process open to conscious reflection, but not dependent upon it; the intricately staged emergence of the steel beam from one triangular panel to the next requires backtracking and is open to its conscious performance, but does not require consciousness. Readers will, as a general rule, appreciate the beam's sud-

den importance; they might, but need not, appreciate how that transformation has been signaled.

The third continuum of conclusions – deduction, induction, and abduction – has already been touched upon to describe the gradual process of semiosis. But in addition to the continuous movement from one conclusion to another in logical sequence, they also entail a continuum of more or less certitude. Not only are induction and abduction uncertain syllogisms as opposed to the determinedness of full argument of deduction (cf. Peirce 1878), but their conjunction allows for a yet more detailed typology. If deduction combines an established rule with an established case of that rule to determine the necessary result, abduction proceeds from a result and an established and possibly pertinent rule to a hypothetical case; and induction from a case and a result, or – as is more usually assumed – many such combinations of cases and their results, to a possible rule. But later versions of this theory distinguish a creative from a simple abduction such that a creative abduction introduces a rule that is less firmly established, and might be hypothetical itself: As Umberto Eco argues, it is in these cases, that

> the law must be invented *ex novo*. To invent a law is not so difficult, provided our mind is 'creative' enough (Eco 1988: 207).

Eco, here, is primarily considering the creativity in scientific reasoning that introduces new categories, focusing on the example of those populating a Linnean taxonomy in biology. However, the combination of all three continua allows us to conceptualize some important aspects of the creativity involved in heautonomous comics comprehension: that revision may often be gradual and subject to revision as we move from panel to panel; it might be open to prescision and abstraction in regrouping, summarizing, and redifferentiating elements and their subsumptions under panel domains: and it might either be ascribed to an author or reader, or none of them, as the concept of this creativity emerges from rather than determines the interrelations between signs.

This leaves us with a firm conception and a greater degree of freedom in explaining the phenomena from which we took our departure. In order to describe that the connection between Spider-Man and the steel beam might be interpreted as a categorical shift and then further connected to the interpretation of the whole episode as an interrupted and eventually avoided tragedy of revenge, we need to engage with the way in which pictorial elements are subsumed in panel domains; we need n o t make assumptions about whether or not the revision between two panels is conscious, or whether its creativity lies with the artist or the reader; but we need to admit room for the creativity of the abductions involved in either case. It is the interrelation of signs described as such, that allows for the possibility of these creative abductions[4], and in order to understand that, we need to consider the pictorial elements involved not merely as icons of denoted objects, but as legisigns governed by visual languages' grammars.

5.   A revisional attitude: Conclusions for multimodal panel segmentation

This leaves us with three major conclusions for any multimodal model of comics comprehension that can deal with formal cohesion as interrelated with, but not reducible to, logical narrative or aesthetic coherence.

First, the importance of r e v i s i o n as an active part in the comprehension of panel sequences has become obvious, and can be explained in specific detail. While revision is usually, perhaps always, possible in any kind of somewhat linear processes of comprehension along the sequence of consecutively arranged elements within a media artifact, the revision involved in the comprehension of comics' panel sequences allows for a more specific descriptions of its properties and functions, with five qualities that I want to emphasize: It is (1) expressible as abductions about the subsumptions of panel elements under panel domains. This assumes (2) that some regularity for the formation of panel domains and their appropriate tokens has to exist in the first place: Grammar has to ground these aesthetics. By the same measure, this introduces (3) a direction for the productivity as opposed to the receptivity of comics' visual language that is mirrored in such a backward movement: For comics' grammar, revision surpasses correction or production. At least in the phenomena discussed here, the ability to resolve irritation by backtracking takes prominence against abilities more readily prevalent in other grammars, especially the ability to correct malformed units or to produce likely continuations. Stylistically, (4) devices of ambiguity in previous panels allows for productive resolutions to irritation in subsequent panels, and that ambiguity plays out in an oscillation between prescision and abstraction as complementary movements associating and separating elements. In terms of (5) a formal aesthetics, this foregrounds the importance of the contour, realized or implied, as an orienting device in the segmentation of the intra-pictorial syntax of comic panels.

Secondly, in terms of a politics of aesthetics, this allows for, but does not necessitate a connection to the Romantic interpretation of c o n s c i o u s h e r m e n e u t i c s as a hallmark of taste and education. To the extent that comics and their accompanying meta-discourses make that connection, they remind us of a historical context in which this aspect was foregrounded beyond its abstract semiotic necessity. We might consider the gap between aesthetics and grammar a hermeneutical corridor, if we want to frame it in the concepts of Romantic theory, and doing so tells us something about comic books and the history of this art form. But we might as easily specify a revisional attitude integral to the semiotics of comics, which is less of a conundrum about the tension between grammar and aesthetics, and more a unified interplay of abstraction and prescision.

In other words, when Spider-Man's head connects with that steel beam, the rearrangement of elements makes sense not merely because the imagined fictional world allows for such accidents, but mostly because the ambiguity and invited revision of the pictorial representation motivates such an

interpretation despite its tension with the previous depiction. We might value such revision and interpret it as a deeper appraisal of aesthetic or narrative consequence (cf. Grünewald 2014). A discourse that pursues this angle will create metaphors of immediacy and reflection, or automation and liberal thought. It will assume that we are given each picture, but that the closure between them is not given, and constructing it promises great cognitive challenges and feats; or that we are given meanings of words, but have to create meanings for images. It is certainly possible to fill the gradual directedness of hermeneutic time with such an imagery of gradual gain. But at the same time, this gradual process relates to any other kind of continued semiosis, and the continuum of subsequent semioses is not specific to these cognitive actions. To foreground some of the cognition involved as conscious, reflective, or sophisticated to distinguish it from others, is to echo a likewise possible but unnecessary social interpretation distinguishing accomplished from other readers by the quality of their thought, as opposed to their trained literacy in the art form. In educational discourses especially, this focus seems to mirror and narrow down a more general tendency to elevate the objects of our study by emphasizing the hardship of their interpretation. But revisional attitudes are not hard; after all, they have their own grammars. They can be learned.

One of the things we learn are the limits of determined deduction as opposed to hypothetical, creative abduction – a better way of distinguishing what is given, to what remains open for discussion. Which is why thirdly, it might be that the historical context of Romantic interpretation and its role in distinguishing an educated social class from other media users is a better explanation than any abstract semiotic theory for the c o n f l a t i o n and even confusion that has been going on in parallel accounts of these devices. They far too readily identify the revisional reading attitude towards comics panels' sequential combination with other dichotomies: If one is the distinction of c o n s c i o u s   o r   r e f l e c t i v e   r e a d i n g from automatic comprehension, another is the combination of pictures and script within and across them. The latter is no more identical to the revisional attitude than the former. The process of revising assignments of graphical elements to panel domains by backtracking revision hardly seems congruent with that other semiotic interrelation typical of but not specific to comics which connects images to the written word. The combination of semiotic modes that happens between writing and pictures in most comic books is of an entirely different nature than the specific type of a semiotic mode that is characterized by the conditions for the revisional attitude. When Frahm, for instance, insists that the heterogeneity between script and pictures in comics underlines their self-referentiality and thus undermines a metaphysical belief in semiotic relations (2010: 146), he at once argues for a well-deserved attention to the parodistic elements of a fundamentally caricatural art form – but distracts from the very pictorial elements and interrelations that define it, to instead discuss their opposition to writing. Ironically, it is this hastily drawn connection to the distinction between pictures and language in comics that may obscure the grammatical features of the pictorial dimensions themselves.

This paper took its point of departure from the tension between two accounts of comics comprehension: One that emphasizes regularities of a grammatical nature, another that foregrounds aesthetic judgments. I hope to have shown that in contrast to many such accounts in recent debates, these views can be integrated. One reward of the effort to do so might lie in the specifications of some uses of backtracking enumerated above.

A second reward might be constituted by a better understanding of the historicity both of those two research positions, and the qualities of the art form that give rise to them. Here, different theories of how we understand comics point to different kinds of social imaginations, expressible by different assumptions about how a theory of comic comprehension should deal with individual differences in comics interpretation. What if you read those panels from *Spider-Man* differently to me?

Some of your interpretation might be separable from the specific functions of detailed panel elements. In my reading, Spider-Man's tragedy of revenge is resolved without saving or condemning the protagonist by the sudden elation of the Goblin's glider from a spatial marker to an interacting object or perhaps even an autonomous agent. A traditional view of the tension between pictures and script will look for a resolution to a possible ambiguity in the pictures in words: those spoken in the comic itself or those offered by readers. In the case at hand, the words spoken by the characters certainly mirror this interpretation of the plot. But that interrelation comes later than the immediate correspondence of lingual and pictorial markers in each panel, and much later than the interplay of the graphical elements, which exploit the productivity of the panel spaces involved, to motivate a sudden resolution that is congruent with the iconically represented world, but does not take its cohesion from that congruence. This might build up a parody of justice, even as the cartoonish representation of the actors implies a parody of iconicity. But it does not force us to identify an aesthetic tension between image and language with a political tension between parody and symbolic order, or of unconnected aesthetic images with grammatically well-ordered panel sequences and compositions. To do so repeats a social imagination of the artform grounded in the hierarchies of previous centuries and their continued quest for the genius interpreter that recognizes, reflects, and resolves or retains these tensions.

This leaves a broad scope for differences in resolving abductions even for the grammatical interrelations between purely pictorial elements on the page. As we know that creative abduction remains uncertain, we cannot presume that other readers will agree with ours. One aspect of the Romantic tradition, repeated in the insistence on a disruptive aesthetics for comics, is its translation of this fact into an unboundedness proper to aesthetics: We might only hope that others agree with our aesthetic judgments, a hope summarized by the *Ansinnen* in Kantian aesthetic theory, as despite evidence to the contrary, we turn to one another with an expectation or even demand, that they might agree with what we have read. But the Peircean concept of abduction does not force resolution; instead, it allows for accu-

mulation. We are equally free to take the openness of that interpretation as an invitation to turn to one another and ask: Here, in this panel, did you see that too?

Notes

1   This fact seemed obvious to me, but the peer review has shown that others do not necessarily think it so. So we need more empirical evidence: Can we elicit convergent solutions from persons fluent both in North American and Japanese visual language for translating a sequence of panels from one to the other? I am not aware of previous work that answers this exact question.
2   I am indebted to Neil Cohn for clearly pinpointing this claim in the discussion of his presentation at the 2021 annual *ComFor* conference, even though I am here taking the opposite stance to his on this detail.
3   I am grateful to the peer review for this important point.
4   In general, the Peircean account of the possibility of a certain semiosis might leave questions about the choice of one of these possibilities, its interestingness, salience, or fit, unanswered. In other words, it explains why what happens is possible, but not why this possibility among many is realized. Cf. Packard (2006) for an attempt to combine these hypotheses with a concept of semiotic desire that motivates such choices while remaining equally non-committal about the ascription of these states to individual minds.

Bibliography

Baltzer, Ulrich (1994). Selbstbewußtsein ist ein Epiphänomen des Zeichenprozesses: Die Landkartenparabel von Ch. S. Peirce. *Zeitschrift für Semiotik* 16, 357–372.

Bateman, John A. (2018). Peircean semiotics and multimodality: Towards a new synthesis. *Multimodal Communication* 7, 1, 1–24.

Bateman, John A., Francisco O. D. Veloso, Janina Wildfeuer, Felix HiuLaam Cheung and Nancy Songdan Guo (2016). An open multilevel classification scheme for the visual layout of comics and graphic novels: motivation and design. *Digital Scholarship in the Humanities* 32, 2, 476–510.

Bateman, John A., Annika Beckmann and Rocío Inés Varela (2018). From empirical studies to visual narrative organization. Exploring page composition. In: Alexander Dunst, Jochen Laubrock and Janina Wildfeuer (eds.). *Empirical Comics Research*. New York: Routledge, 127–153.

Chomsky, Noam (1956). Three models for the description of language. *IRE Transactions on Information Theory* 2, 3, 113–124.

Cohn, Neil (2010). Japanese visual language: The structure of manga. In: Toni Johnson-Woods (ed.). *Manga – An Anthology of Global and Cultural Perspectives*. New York: Continuum, 187–204.

Cohn, Neil (2013). *The Visual Language of Comics*. New York: Bloomsbury.

Cohn, Neil (2016). The vocabulary of manga: Visual morphology in dialects of Japanese Visual Langue. *Journal of Pragmatics* 92, 17–29.

Cohn, Neil (2018a). In defense of a 'grammar' in the visual language of comics. *Journal of Pragmatics* 127, 1–19.

Cohn, Neil (2018b). Combinatorial morphology in visual languages. In: Geert Booij (ed.). *The Construction of Words*. London: Springer, 175–199.

Cohn, Neil, Martin Paczynski, Ray Jackendoff, Phillip J Holcomb and Gina R Kuperberg (2012). (Pea)nuts and bolts of visual language: Structure and meaning in sequential image composition. *Cognitive Psychology* 65, 1–38.

Cohn, Neil and Hannah Campbell (2014). Navigating comics II: Constraints on the reading order of comic page layouts. *Applied Cognitive Psychology* 29, 2, 193–199.

Cohn, Neil and Joost Schilperoord (2022). Remarks on multimodality: Grammatical interactions in the parallel architecture. *Frontiers in Artificial Intelligence* 4 (no page indication).

Colapietro, Vincent (2014). Notizen zu einer Skizze einer peirceanischen Theorie des Unbewussten. *Journal für Psychoanalyse* 55, 119–140.

Eco, Umberto (1988). Horns, Hooves, and Insteps. Some hypotheses on three types of abduction. In: Umberto Eco and Thomas A. Sebeok (eds.). *The sign of three – Dupin, Holmes and Peirce*. Bloomington: Indian University Press, 198–220.

Foulsham, Tom, Dean Wybrow and Neil Cohn (2016). Reading without words: Eye movements in the comprehension of comic strips. *Applied Cognitive Psychology* 30, 4, 566–579.

Frahm, Ole (2010). *Die Sprache des Comics*. Hamburg: Philo.

Grennan, Simon (2017). *A Theory of Narrative Drawing*. New York: Palgrave Macmillan.

Groensteen, Theirry (1999). *Système de la bande dessinée*. Paris: Presses Universitaires de France.

Grünewald, Dietrich (2014). Das Prinzip Bildgeschichte: Comics als eigenständige Kunstform und die Notwendigkeit, den sachgemäßen Umgang mit ihr zu erlernen. In: Barbara Lutz-Sterzenbach, Maria Peters and Frank Schulz (eds.). *Bild und Bildung. Praxis, Reflexion, Wissen im Kontext von Kunst und Medien*. München: kopaed, 453–80.

Halliday, Michael A. K. and Ruqaiya Hasan (1976). *Cohesion in English*. London: Longman.

Homann, Renate (1999). *Theorie der Lyrik. Heautonome Autopoiesis als Paradigma der Moderne*. Frankfurt am Main: Suhrkamp.

Jackendoff, Ray and Jenny Audring (2020). *The Texture of the Lexicon. Relational Morphology and the Parallel Architecture*. Oxford: University Press.

Jakobson, Roman (2010). Linguistics and poetics. In: *Selected Writings III: Poetry of Grammar and Grammar of Poetry*. Berlin: De Gruyter, 18–51.

Kirtley, Clare, Christopher Murray, Phillip B. Vaughan and Benjamin W. Tatler (2023). Navigating the narrative: An eye-tracking study of readers' strategies when reading comic page layouts. *Applied Cognitive Psychology* 37, 1, 52–70.

Krafft, Ullrich (1978). Comics lesen. Untersuchungen zur Textualität von Comics. Stuttgart: Klett-Cotta.

McCloud, Scott (1994). *Understanding Comics. The Invisible Art*. New York: Harper Collins.

Menzel, Katrin, Ekaterina Lapshinova-Koltunski and Kerstin Kunz (2017). Cohesion and coherence in multilingual contexts. In: Katrin Menzel, Ekaterina Lapshinova-Koltunski and Kerstin Kunz (eds.). *New Perspectives on Cohesion and Coherence*. Berlin: Language Science Press, 1–11.

Packard, Stephan (2006). *Anatomie des Comics. Psychosemiotische Medienanalyse*. Göttingen: Wallstein.

Peirce, Charles Sanders (1868). On a new list of categories. *Proceedings of the American Academy of Arts and Sciences* 7, 287–298.

Peirce, Charles Sanders (1878). Deduction, induction, and hypothesis. *Popular Science Monthly* 12, 470–482.

Rancière, Jacques (2008). *Ist Kunst widerständig?* Edited and translated by Frank Ruda and Jan Völker. Berlin: Merve.

Sachs-Hombach, Klaus (2013). *Das Bild als kommunikatives Medium: Elemente einer allgemeinen Bildwissenschaft*. 3rd ed. Köln: Herbert von Halem.

Schleiermacher, Friedrich (1809/10). Allgemeine Hermeneutik von 1809/10. Edited by Wolfgang Virmond. In: Kurt-Victor Selge (ed.). *Internationaler Schleiermacher-Kongreß 1984*. Berlin: De Gruyter, 1269–1310.

Stainbrook, Eric (2016). A little cohesion between friends; Or, we're just exploring our textuality: Reconciling cohesion in written language and visual language. In: Neil Cohn (ed.). *The Visual Narrative Reader*. London and New York: Bloomsbury, 129–157.

Tseng, Chiao-I and John A. Bateman (2018). Cohesion in comics and graphic novels: An empirical comparative approach to transmedia adaptation in *City of Glass. Adaptation* 11, 2, 122–143.

Wiesing, Lambert (2008). Die Sprechblase: Reale Schrift im Bild. In: Alexandra Kleihues (ed.). *Realitätseffekte. Ästhetische Repräsentation des Alltäglichen im 20. Jahrhundert*. München: Fink, 25–46.

Wildfeuer, Janina (2013). Formale Zugänge zur Diskursanalyse. *Zeitschrift für Semiotik* 35, 3–4, 393–417.

Wildfeuer, Janina, John A. Bateman and Tuomo Hiippala (2020). *Multimodalität. Grundlagen, Forschung und Analyse. Eine problemorientierte Einführung*. Berlin: De Gruyter.

Sources

Conway, Gerry (wri) and Gil Kane (art) (1973). *Amazing Spider-Man 122: The Green Goblin's Last Stand!* New York: Marvel.

Image Sources

Fig. 1. Conway (wri) and Kane (art) (1973: 4).
Fig. 2. Krafft (1978: 48).
Fig. 3. Conway (wri) and Kane (art) (1973: 20).

Prof. Dr. Stephan Packard
Universität zu Köln
Institut für Medienkultur und Theater
Meister-Ekkehart-Str. 11
50937 Köln
Germany
E-Mail: packard@uni-koeln.de