

Studies in Comics

Volume 5 Number 1

© 2014 Intellect Ltd Article. English language. doi: 10.1386/stic.5.1.57\_1

---

**NEIL COHN**

Center for Research in Language, University of California, San Diego

# Building a better ‘comic theory’: Shortcomings of theoretical research on comics and how to overcome them

## Keywords

comics theory  
language  
linguistic  
    structuralism  
grammar  
semiotics  
cognition  
research methods

## Abstract

*Research on the understanding of ‘how comics work’ has grown tremendously over the past twenty years, with more articles and books emerging each year. Much of this research has discussed comparisons between comics and language, and/or has speculated on comics’ cognition. However, much of this research faces limitations, which hamper the seriousness of the endeavour and reflect the youth of this emerging field. This article points out these deficiencies that pervade theories about comics. These include inadequate background research, overly general and unsupported claims, a lack of adequate evidence, and limitations for research methodologies. To address these concerns, I draw from over 50 years of research from linguistics and cognitive science to inform how the field of ‘comic theory’ can move forward. In particular, I outline two primary*

*ways of progressing with this line of research: (1) explicit manipulation of the component parts of the structure used in comics and (2) cataloguing actual comics for various theoretically relevant phenomena. This data-driven approach is offered as a guiding vision for future works on the understanding of 'how comics work'.*

## Introduction

Over the past twenty years, the increasing number of papers in comics studies reflects the significant strides that have been made in people's interest in understanding the mechanisms at work underlying comics. However, much of this work faces significant problems, which often leave the overall conclusions from this work not very useful for building an understanding of the structure and comprehension of sequential images. In this article, I detail criticisms of this budding field, and give advice for doing substantive research on comics studies.

I should state at the outset that this article targets the branch of research involved with understanding the underlying structure of comics and how people understand them. Most specifically I target those studies that have been concerned with comparisons between comics and language, and studies that have incorporated aspects of cognition and cognitive science. Essentially, these works are concerned with how the medium used in comics is built and how it is understood by people's minds. These concerns have often been subsumed under the name 'comic theory', and thus I will therefore use this term – and the word 'theory' more generally – specifically to target this set of research questions. This is certainly not the only way to do research on or theorize about comics, but it is where these criticisms and advice will be directed.

## Misunderstanding language

Let us first consider discussions of how sequential images are or are not similar to language. In recent writings, it has become popular to disparage claims that sequential images are structured akin to language, most pointedly critiquing the idea that sequential images might be guided by a 'grammar'. In her book *Narrative Structure in Comics*, Postema writes that, '... images communicate largely without rules ... the smallest elements of images have no set meanings, and the way these elements are combined or even repeated are not governed by rules like grammar' (2013: xvi). An 'anti-language' argument is made more thoroughly in Miodrag's (2013) book *Comics and Language*, which uses criteria for comparing the structure of comics against language. Such views are repeated across many examples of comic theory (e.g. Hick 2012; Horstkotte 2013; Kukkonen 2008; Walsh 2006).

These works all make the argument that the sequential images in comics are not structured or understood in comparable ways to language. The foremost problem with these critiques is that they do not adequately understand what language is or hold an outdated view on language. In most of

these cases, the indictment against comics being 'like language' argues specifically against a structuralist conception of 'language'. This view of language first emerged in Ferdinand de Saussure's posthumous writings in the early twentieth century (de Saussure 1972), and became the dominant paradigm of studying language in America (e.g. Bloomfield 1933). It then continued as the foundational framework for a semiotic understanding of language and semiotics more broadly, extending to become a significant paradigm within the social sciences and cultural studies. This approach of 'linguistic structuralism' used 'language' as a metaphor to discuss the 'structure' of many aspects of culture, society and the arts by breaking them down into minimal units shared as part a cultural code. As part of this larger research paradigm, many European comic theorists applied these methods to comics (e.g. Gubern 1972; Hünig 1974; Koch 1971; Nöth 1990).

Most anti-language views on comics are critiquing the notion that comics are like this structuralist description of language. This also includes Groensteen (1999), who supports that comics are 'a language', but disagrees with the structuralist account. Thus, in context, critiques against a structuralist view of comics have been a response to developments within the field of comics scholarship more generally situated within the humanities and social sciences. By discussing linguistic structuralism as is, these researchers are upholding the traditions of their own fields (and the history of comics scholarship), which as broader disciplines may maintain threads of the structuralist paradigm.

The problem here is that the arguments made by these authors are not framed as critiquing 'language as metaphor' for larger research, or linguistic structuralism as a paradigm specifically. Rather, their arguments discuss comics being like or unlike 'language' as a phenomenon, holding up the structuralist account as accurately conveying what language is. This is made clear in the case of Miodrag and Postema by their discussion of specific aspects of language like phonemes, morphemes and grammar as evidence for their claims. It would be fully acceptable to argue against the use of linguistic structuralism specifically to describe comics, and perhaps this is implicit in these authors' writings for members of the audience who belong to their home disciplines. However, by not making this explicit, and by couching the discussion with regard to 'language' broadly, it perpetuates the belief for a broad readership that the structuralist paradigm accurately reflects how 'language' is structured, which is not the case.

Rather, no one who has seriously studied language for the past 50 years has held a structuralist viewpoint of language. Following the 'cognitive revolution' in the 1960s, structuralist ideas about the organization of language were rejected by nearly all researchers in linguistics in a massive paradigm shift (Harris 1993). This new understanding argued that 'language' is not a system of 'rules' and 'codes' floating in the cultural ether, but is instantiated in rules and principles that reside in the brains of speakers across the world. The goal then is not to understand minimal units or surface features of language's structure, but rather to understand the principles at work in the comprehension of language found in people's minds/brains. Because of this, it became easy to examine whether

something had grammar or did not: examples could be tested against the intuitions of actual speakers, and those intuitions could be used to build theories about how the cognitive architecture of language understanding worked. That cognitive architecture is the ‘grammar’ of language.

If claims about ‘language’ are to be made – rather than ‘linguistic structuralism’ as a paradigm – it is not unreasonable to address just what ‘language’ is by the standards of the field that studies it. Commendably, Miodrag’s (2013) book attempts to do this, offering a well-reasoned argument for similarities and differences between the structure of comics and language. She compares comics against a list of criteria for ‘what is language’ cited from Mitchell (1986), which includes the requirement for linguistic signs to be arbitrary, discrete and have minimal units. This list originally came from a classic article by the linguist Charles Hockett (1960) in which he laid out several ‘design features’ for what constitutes a language, with the intent of contrasting it from various animal communication systems (uncited by Miodrag). However, Hockett was steadfastly a structuralist (Harris 1993), and this list of criteria was never updated in light of the re-understanding of language from the cognitive revolution. This list also predated the insights from the subsequent study of sign language, which had to contend with – and later directly challenged – many of Hockett’s criteria, specifically discreteness and arbitrariness (Liddell 2003). Hockett’s list is still referenced in contemporary linguistics, though now discussion of it often focuses on what he got right and what he got wrong, given what we know now.

If theorists wish to argue that the sequential images in comics are or are not structured like ‘language’ (as opposed to arguing about linguistic structuralism specifically), then it seems reasonable that they should both know what language is given our current scientific understandings and cite relevant references to such ends. In doing so, scholars should engage with contemporary ideas from linguistics or psycholinguistics, rather than appealing to ‘everyday notions’ of language, to structuralism, or to semiologists like Barthes or Eco, which do not reflect the established thinking of the language sciences. Without doing this, it only perpetuates ideas about language that have long been considered outdated. Granted, this might force scholars to engage with a literature outside of their own traditions, but such is the demand of doing appropriate and informed background research. Besides, the study of comics has always lauded itself for its interdisciplinary scope (and rightly so), so such extensions may hopefully be more comfortable to scholars of comics in the first place.

For those who wish to engage contemporary conceptions on the structure of language, there are many good resources. The textbook *Contemporary Linguistics* by O’Grady et al., includes introductions to most of the subfields of linguistics research (the sixth edition came out in 2008). A historical perspective on the downfall of structuralism, the rise of the cognitive revolution, and the internal conflicts within linguistics throughout the 1960s and 1970s are entertainingly detailed in *The Linguistics Wars* by Harris (1993). For an introduction to contemporary thinking on language as a whole, Pinker’s (1994) *Language Instinct* and Jackendoff’s (1994) *Patterns in the Mind*, are both written

for a popular audience. For a source more directly reflective of primary research, Jackendoff's (2002) *Foundations of Language* is among the best, coming from a linguist who has been central to the study of language and its connections to other aspects of cognition for over 40 years (full disclosure: he was also my mentor). He also starts by making a relevant statement to this discussion: one should know at least as much about the basic structures of language laid out in the introductory chapter if you are going to talk about it with any authority.

### **Background research**

The lack of providing up-to-date research does not rest solely with references to language. Studies of comics often invoke ideas from psychology and the cognitive sciences without citing recent work from the field. For example, Medley (2010) explores the differences between cartoony and realistic styles, posing in the abstract to describe 'the psychological mechanisms by which we understand images abstracted away from realism' and that the 'article explores some important faculties of the human visual system'. Yet, despite the interesting observations the article provides, it contains only a single reference to any substantive research done on perceptual psychology in the last 30 years (from 2002) with most citations pointing to works from the 1970s or *earlier*. Another instance occurs in Mallia's (2007) paper on the use of comics in education that purports to investigate 'comics and cognition'. The references for this paper include almost none of the prior research done on comics in education (e.g. those summarized by Nakazawa (2005) or many others), and includes few actual citations about cognition. Rather, most of the references are historical or literary treatments, and these appear to be mostly from 20 to 30 years old.

Research on cognition and psychology has made great strides over the past 30 years, often characterized by rapid changes made possible by the rise of computational modeling and neuroimaging. If papers on the structure of comics wish to make credible connections with ideas from psychology, cognitive science and cognitive neuroscience, it is worth directly engaging the recent work from these fields and not relying on decades-old sources or from works outside these disciplines. As with language, neglecting to engage with recent research risks perpetuating outdated notions of these fields, especially those that might have predated the contemporary insights of brain science and hold inaccurate ideas about how the brain works.

It is important to note that this statement is neither advocating nor disparaging the citing of ideas from cognitive science. Though cognition may have become a 'sexy' topic recently, the need to reference cognition should follow the aims of an author, and such citations may or may not be needed. For example, Kukkonen (2013a) frequently invokes the concepts of 'embodied cognition' in an analysis of Winsor McCay's 'Dreams of the Rarebit Fiend' and other comics. Here, it is claimed that we understand a character floating in the air by using 'image schemas' (e.g. Johnson 1987) of

generalized spatial knowledge of shapes and weights. However, no argument is made for why image schemas are necessary to understand this comic or how the analysis would change if image schemas *were not* invoked. The paper goes on to even discuss ‘mirror neurons’, which have been found to activate in the brains of monkeys (Rizzolatti et al. 1998) – but only controversially in humans (Hutto 2013; Turella et al. 2009) – when they both see and perform actions. It is unclear the relevance for invoking brain science given the aims of this paper, especially for a topic so widely controversial. Such a paper distinctly contrasts with useful and relevant applications of cognitive science concepts, like Kukkonen’s (2013b) own framing of mental models (e.g. Johnson-Laird 1983) to explain the comprehension of multiple fictional ‘universes’. Here, Kukkonen’s use of cognitive science concepts directly reveals insights of a phenomenon in comics (and, to her credit, in both cases Kukkonen engages directly with recent notions from cognitive science). Unless a paper directly discusses the cognition of comics – and cites the relevant studies on that topic – appealing to cognitive (neuro)science is unnecessary at best and hand-waving at worst (see, e.g., similar arguments by Trout 2008). Authors should therefore consider whether such concepts are needed in the first place, given their aims.

A final point should be made on this issue of background research related to cognition. We have reached a point where psychological research has been carried out on various facets that go into understanding comics. If one has an interest or is concerned with ‘how people comprehend comics’, researchers can no longer assume that the field is ‘new’ enough that it has not been done. Rather, various studies on comics and sequential images over the past 30 years can be found scattered amongst the psychological literature using eye-movements, cognitive neuroscience, and other experimental methods. They are not copious (yet), but there are enough out there that if a paper wishes to talk about the ‘understanding’ of comics, it seems reasonable to ask that they seek out and cite such research.

## Providing evidence

In papers that argue against a ‘language’ comparison with comics, the most pointed criticism is that sequential images have no structure to their sequence, and thus have no ‘grammar’. For example, Postema states that, ‘The order of words in sentences and syllables in words is governed by various kinds of grammar, which allow for certain combinations but not others. Panels in sequences can be combined in infinite ways, because they are not governed by such grammar’ (2013: 57). Hick is even more sceptical, stating with preemptive defeat that ‘the notion of a *syntax* of comics is a difficult concept to even wrap one’s head around ... it is not at all clear how (if at all) systematized concatenation rules might even be described – and if there *are* such formalizable rules, we certainly don’t know them’ (2012: 140, original emphasis).

The problem here is that no adequate *evidence* is provided that sequential images communicate 'without rules' (in Postema's terms) or that narratives cannot feasibly be 'ungrammatical' (Walsh 2006). These authors merely state this claim as a presupposed fact with no proof or argumentation to support it. In Hick's case, he does not even seem to know how to *conceive* of providing such evidence, and thereby dismisses the whole endeavour outright without even looking to see if, in fact, such rules had been proposed before for sequential image comprehension – which they had (Branigan 1992; Carroll 1980; Cohn 2003). In order to substantiate a claim about the existence or non-existence of grammar in sequential images, it must be backed by evidence, not by opinion.

When papers do attempt to provide 'evidence' against a grammar, they often use pages from comics that they feel defy these (unstated) principles. This is the tactic taken by Horstkotte (2013), who at least attempts to support an 'anti-grammar' position using particular pages from published comics. However, any single example of an 'interesting' or 'rule-breaking' page taken directly from a comic is a bad form of evidence about the structure of the medium. The reason for this goes back to the ideas from the 'cognitive revolution' for defining 'what is' a grammar. A grammar is a set of rules and constraints within the mind of a speaker that governs the production and reception of an expression (such as a sentence, or, here, a sequence of images). Thus, any normal, comprehensible comic page should have been created by a grammar – present in the mind of its author – and thus most comic pages will not depart in ways that support an argument against systematicity.

If a page departs from the grammar it should appear as an *exceptional* example that noticeably departs from the normative feeling of most comics. It should feel 'wrong' compared to the usual reading experience, just like *Strong the hero smacked villain the* should sound 'wrong' as a sentence because it violates the rules of English grammar (despite the strained recoverability of its meaning, which is separate from its grammar). Many examples offered by theorists are indeed exceptional – or at least artistic – instances that challenge the conventions of standard comics. However, because these works are inherently 'poetic' in nature and do break the 'rules', they do not provide evidence against systematic principles. Rather, the observation that such examples are interesting betrays the systematic nature of more normative examples. Yet, basing theory on such exceptions cannot then extend to explaining the comprehension of more normative examples themselves.

Evidence to support or deny the existence of a grammar in sequential images can only be provided by manipulating a sequence (or a panel) and comparing the generated results. Because a 'grammar' is a collection of cognitive constraints in the mind of a creator (or reader), explicitly violating those rules provides the only way to discover how they work (or do not). This relates to a second type of critique of the 'grammatical' approach to sequential images (or all narratives): the assumption that there are no 'ungrammatical' narratives (Walsh 2006), and thus the belief that there are an infinite number of possible sequences. This belief is echoed in the idea that all relationships between panels – even non-sequitur images – are somehow comprehensible (McCloud 1993; Saraceni 2001).

However, the only way to verify such a critique is to manipulate sequences and see if there are indeed no unacceptable permutations.

Manipulating language is how linguists have studied the structure of language since casting aside structuralism. The structure of the system is altered in order to test the principles that govern it (i.e., examine the cognitive structures in the minds of speakers). This methodology can also work for the study of sequential images to show that a grammar *does* constrain their comprehension. This type of calculated manipulation is the only viable methodology – i.e., the *scientific* method – by which the presence or absence of a grammar can be made. Only by doing these types of careful and explicit manipulations can a system truly be said to be like or unlike ‘language’.

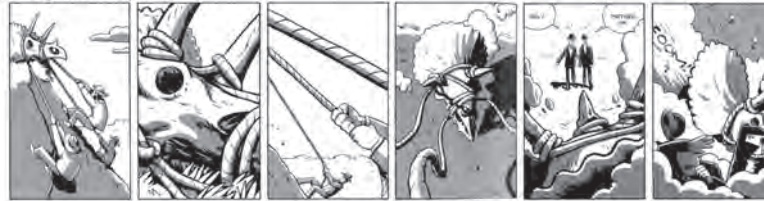
The most basic manipulations in linguistics research involve moving, deleting or substituting the elements of a sentence, and these techniques can easily be applied to sequential images. Let’s start by rearranging elements in a sequence. The most simple test would be to take a given sequence of narrative images found in a comic, and simply try out all possible orderings of the panels. That is, move around the orders of the panels and see if the scrambled orders make sense. We can call this a ‘movement test’. Figure 1a depicts an example sequence from *Bone Sharps, Cowboys, and Thunder Lizards* (2005) by Jim Ottaviani, Kevin Cannon and Zander Cannon. Here, several men pull at ropes attached to a triceratops skull, which finally comes loose and crashes into men standing at the bottom of the hill. This sequence should be easily comprehensible, and should not seem like a glaringly ‘exceptional’ example of the medium.

Figure 1b now rearranges the order of these images, resulting in a less understandable sequence. ‘Scrambling’ the order of all panels in a sequence has been used as a technique in several studies of sequential image understanding (Cohn et al. 2012; Gernsbacher et al. 1990; Nagai et al. 2007). Because the order of panels in Figure 1b is less comprehensible, it supports that there is some sort of rule system – a ‘grammar’ – underlying its understanding. Only if *every single* possible rearrangement of panels makes sense, would it provide evidence that Postema is correct that ‘panels in sequences can be combined in infinite ways, because they are not governed by such grammar’ (2013: 57). However, if *even one* sequence seems harder to understand, then Postema’s hypothesis is wrong, because it shows that some sort of *constraints* operate on the system to limit the way in which sequences are understood. A sequence that is ‘harder to understand’ – like Figure 1b – provides evidence for a grammar in the cognitive sense.

Another test is to *delete* elements, which can reveal how important an element might be to a sequence. For example, deleting the adjective in the sentence *The strong hero smacked the villain* gives us a perfectly understandable sentence (*The hero smacked the villain*), but deleting the noun is less acceptable (*The strong smacked the villain*). This tells us that adjectives are optional while nouns are not. Similarly, we can do this to panels in a sequence to test whether some panels are more or less important to a sequence. A theory of panel transitions like the one posed by McCloud (1993) would



a) Original sequence



b) Scrambled order of panels



c) Omission of non-essential panels



d) Omission of a more central panel



Figure 1: Example sequence (a) manipulated by (b) scrambling the order of panels, (c) omitting non-essential panels and (d) omitting a panel more central to the meaning. Bone Sharps, Cowboys, and Thunder Lizards is © 2005 Jim Ottaviani, Kevin Cannon, and Zander Cannon.

claim that omission of *any panel* should subsequently lead to inference (or ‘closure’) because all panel-to-panel relationships require inference anyhow. A deletion test could be used to see if this is true.

Let’s now apply the deletion test to our example sequence. Figure 1c omits the second and third panels from Figure 1a. This results in a sequence that is fully understandable, and these panels are hardly noticed as missing – does their omission require the inference that McCloud suggests is necessary? Compare this to the sequence in Figure 1d, where a single panel has been deleted, yet it results in a more choppy sequence, since the panel showing the initial pull of the skull is now missing. This omission certainly requires some inference, and likely feels less coherent as a sequence. Because these two deletions differ in their impact on the sequence, we must assume that they index some sort of system of comprehension in different ways.

A third test might be *substitution* of certain components for another. Because you can replace a pronoun for a noun or noun phrase (*She smacked the villain*) it tells us that these elements are distributed similarly. In comparison, if one panel (or a sequence of panels) could be taken from a sequence and replaced for another and the sequence could still be comprehensible, it might indicate that these panels share particular properties. On the other hand, if substituting one panel for another makes the sequence read like gobbledygook, it might also tell you something about the nature of the panel relative to its sequence context.

Movement, deletion, substitution, and other types of linguistic tests can – and have – worked to reveal aspects of the structure of the sequential images in comics. Methods like these should be used whenever claims about the structure used in comics are made. For example, these diagnostics can be useful might be in testing claims about the relationship of page layout and sequential image comprehension. Horstkotte (2013) argues that a particular page from Neil Gaiman’s *Sandman* belies a ‘linear’ understanding because of the particularly decorative and meaningful page layout. However, if one were to move these panels into another layout – say, into a straight linear order – the felicity of the sequence would not be damaged at all. However, the *layout* would no doubt be less decorative and ‘meaningful’. By manipulating the sequence, we can show that the system governing the creation of meaning across panels *is separate from* the system that guides how a person navigates across a page. They might interact, but they are separate systems, and thus do not quite provide the support that Horstkotte (2013) claims.

In closing, it is worth noting that these tests are not limited to simply seeing whether they damage sequences of images or not. By continuously manipulating a sequence in multiple ways, such permutations can lead to generalizing a theory about how the system of sequential images works – its ‘grammar’. Enterprising theorists then might want to test this theory by using its predictions to manipulate other sequences and then gather judgments from other people, in an experiment. These types of *scientific* methods offer the best way to investigate how comics and sequential images are structured and conceived in a cognitive sense.

## Generalizations vs specificity

One advantage of the described scientific method for studying sequential images is that it requires that the scope of enquiry remains constrained and specific. Many approaches to comic theory are so broad as to be mostly uninformative. Simply pointing out that certain parts of the medium exist and behave in 'unexpected' ways does not contribute to any real understanding of how they operate. Such cases provide analysis that is so general as to not be able to provide evidence for or against any particular theory. By comparison, papers in linguistics or psychology almost never try to generalize to 'all of language' or 'all of cognition', but rather pinpoint a specific structure and then rigorously detail it. Scholars aiming to discuss the structure used in comics would be better served by following a similar model.

A related practice in comics theory is the proclamation of various 'principles' that operate across the structure or comprehension of sequential images. Ideas such as Groensteen's (1999) braiding or arthrology or Postema's (2013) invocation of 'filling of gaps' are extremely broad and general. Just how do these concepts allow us to understand comics? What are the mechanisms that allow them to work? How could we study if these claims are true? What are the alternatives to them such that they can be studied as falsifiable hypotheses? Unless you can say what it is, how it works and how it offers predictions about other examples, it is not a useful piece of theory.

Let's look at a few examples. McCloud's (1993) idea of closure was offered with decent structure around it. At the very least, he claimed that closure was facilitated by various types of panel transitions, and that these transitions created 'more' or 'less' closure. This provided at least a basis for measurement – predictions for more or less of this process. Nevertheless, the idea remains a little vague. Closure is supposed to be a 'cognitive' process – so how exactly is the mind/brain supposed to be doing this? And, is the word 'closure' anything more than just a hand-waving term to mask a lot of complexity that McCloud simply cannot provide? Subsequent authors offered further descriptive value to this idea. For example, Saraceni (2000, 2001) operationalized McCloud's transitions through the contributions of repeated elements and/or maintenance of a common semantic theme between panels. He then argued that the degree to which these components changed across panels motivates the differing amounts of inference (i.e. closure), thereby providing a measurable account for McCloud's idea. And, in fact, similar notions to this operationalized account of 'transitions' exist within cognitive science (e.g. Zwaan et al. 1998), and have been applied to studying visual narratives in film (e.g. Magliano and Zacks 2011, Magliano et al. 2001).

By comparison, let's consider Groensteen's (1999, 2007) notion of 'arthrology', which he uses to describe the linear relations of panels to each other (restrained arthrology) or non-linear relations (general arthrology). Groensteen states that 'every panel exists, potentially if not actually, in relation with each of the others' (Groensteen 2007: 146). How is this notion useful to understanding how

sequential images are comprehended if it is so unrestrained that there are no limits to it? Where do the relations stop? Do the relations end at a book's finale? For episodic comics, do all the panels of one issue then relate to all subsequent issues too? What if they are all collected into one volume? What then stops relations from forming between two completely unrelated comics? Second, such an unrestrained notion of panel relations would be completely unfeasible, and would overload the working memory of the human mind. As argued in Cohn (2010), an average 24-page monthly comic book with six panels per page would have 144 total panels, yielding 10,296 possible panel relationships! (This is calculated as  $144!/(2! \cdot 142!)$ ). Without some sort of system to constrain these relations (i.e., a grammar!) keeping track of all of these connections between panels (whether they are 'active' or not) would overwhelm human working memory.

In order to say something substantive, comic theory should be directed towards specific issues that are more than just vague 'principles' that mask deeper probing and specificity. At the least, we should ask for theories to be verifiable through some sort of testing. Without this, we risk the perception that different theories provide alternate yet equally valid perspectives on how comics are structured or comprehended in the mind. This is simply untrue. There are many ways to show that theories are not viable, and all involve testing them against the ways in which human brains are able to comprehend sequences of images.

Thus, the advocacy here is that works on comic theory restrict themselves to constrained and specific domains of interest. Most papers cannot hope to cover *all* of the structure of sequential images, let alone cover how multiple components interact. Useful and informative research should target specific and particular domains of structure and rigorously investigate them (unless of course a paper is a review, which would naturally require a broader scope summarizing the more narrow scope of precedents). This is especially true because there is *so much* going on within comics that we do not understand, and only by doing careful analysis will we discover how such minutia operate.

We might characterize research across two dimensions, each with two methods of going about them. First, research can be either *observation driven* or *theory driven*. Observation-driven research involves seeing a phenomenon in a comic, then rigorously analysing how it operates in other books (if it does) and exploring the theory behind why it might occur. This is the method taken by Abbott and Forceville (2011), who noticed that characters in the manga *Azumanga Daioh* sometimes have hands that turn into stumps. They carried out a rigorous analysis of the conditions by which this occurs, and then detailed why they think this phenomenon arises. Observation-driven research can occur across all levels of comics' structure, from parts of images to whole panels, to sequential image comprehension and page layouts. Again though, this research is most effective when it targets specific issues rather than broad generalizations.

Theory-driven research is motivated by ideas about the structure found in comics. This type of paper would be how new theories of an aspect of the structure of comics would be proposed. This

could also be in support of or against a position. For example, Saraceni's (2001) paper on panel transitions lays out explicit and testable ways that McCloud's (1993) panel transitions could operate. This paper did not attempt to discuss every aspect of structure in comics; it just focused on meaningful panel-to-panel relationships. An opposite case is the critique of panel transitions by Cohn (2010), which argues that such panel-to-panel relationships are not enough to account for all of sequential image comprehension. Both of these papers had very specific topics (panel transitions) using targeted examples.

This brings us to a second dimension of methodology. Besides providing several examples in support of or against a position that might be taken, there are two predominant methods that observation or theory-driven research can use. They can either manipulate the structure to see how it works, or catalogue and describe how a phenomenon works within a single comic or many comics. The first method of manipulation was described previously, using techniques such as movement, deletion and substitution. This technique can be motivated either by observation (you find something interesting in a comic, then manipulate it to see how it works) or by theory (you have an idea about how something works, then manipulate it to see if you are right or wrong). This methodology is the standard for psychological experimentation.

An aside here is worth mentioning. Subjective experiences that someone has while 'reading' a comic do not adequately inform how *actual* comprehension works. For example, comic theorists sometimes describe how their eyes tend to move across a comic page – or how they assume a readers' eyes might move – and they use this as support for a particular theoretical viewpoint (Abbott 1986; Miodrag 2013; Postema 2013; Pratt 2009). However, the psychological literature has established quite well that eye-movements differ substantially from our conscious awareness of what they are doing, and that only measurements such as eye-tracking can truly detail such behaviours. Similar caveats extend to other subjective aspects of understanding garnered through the reading process (but notably, not the 'ugh, that doesn't make sense' response used in Figure 1, which involves direct manipulation to examples and can always be checked by empirical measures). The creation of experimental methods makes such self-reports largely unneeded.

Cataloguing a phenomenon is also a useful methodology. Here, someone might observe a phenomenon in a comic, then rigorously detail how it is used in that and/or other comics. Again, a good example of this comes from Abbott and Forceville's (2011) study of manga where people's hands turn to stumps. Rather than observing one instance and describe it, they carefully and quantitatively tabulated all the instances of this and similar phenomenon throughout the book, which allowed them to get a handle on the scope and magnitude of the phenomenon. By doing this, they also discovered examples where legs turned to stumps, along with other related examples.

Theory can also motivate this type of research. In this case, a person might have a theoretical idea for how comics operate, and then use that theory to quantitatively describe how it manifests across

different comics. This was the method used by McCloud (1993) after he proposed his theory of panel transitions. He first made an observation about the differences between what he saw occurring in comics from America, Europe and Japan. He then built a theory to account for these differences: panel transitions. He then used this theory to see if his observations were correct, that American and Japanese comics use transitions in different ways. This analysis then allowed him to make data-driven claims about the ways in which different comics might be structured across the world.

In all of these cases, the research was limited to specific and constrained domains of interest. This is also what makes them compelling studies: they are able to say interesting things backed up by data that is verifiable and does not overreach. This is a curious aspect of the legacy of McCloud's (1993) *Understanding Comics*. Throughout this book, McCloud tested his theories using explicit methodologies like those described here. He used a corpus analysis to test panel transitions across cultures, and directly manipulated sequences to examine how time passes across panels. Though the book is widely referenced and has inspired many subsequent theories, few scholars have taken McCloud's lead in terms of methodology. Granted, there are significant problems with McCloud's ideas, many of which have been raised by me (as in Cohn 2010). However, the creation of theory that requires careful testing and quantitative analysis is a sound insight, and one that is rarely followed by the majority of those who study the 'theory' of comics.

### **What is 'comic theory'?**

Given this preceding discussion, it is worth asking what 'comic theory' means to people, and to clarify the intent by which people seek to understand 'how comics work'. In some sense, the critiques raised here may fall on disciplinary divides. 'Theory' within comics studies has often straddled the line between how comics are 'understood' in a cognitive sense and how they are 'understood' in an artistic/literary, aesthetic sense. This tension has contributed to the challenges outlined above, and addressing which aspect of 'understanding' the field is concerned with (or at least, each paper) seems important for establishing a path forward.

If 'comic theory' is meant as a branch of literary/artistic theory, then it assumes that scholars are trying to understand how the structure of comics might be used to create some sort of interpretable expression in the service of appreciating various works as literature/art. Here, it would be quite understandable to find a paper waxing poetic about exceptional examples and providing interpretations of what various expressions 'mean' in the sense of their aesthetic 'message'. However, literary theory does *not* seek to explain how the structure of the medium works or is comprehended, just as it does not seek to understand how *language* works.

It is a separate set of questions if 'comic theory' seeks to identify how the structure of sequential images works generally and across possibly diverse cultural conventions and/or how that structure is

comprehended by people's minds. These questions should be studied within the realm of the cognitive and psychological sciences, just like the study of the structure of language or any other human behaviour. Given this, the *methods* of linguistics and cognitive science should be held as a standard for research done on these questions.

Unfortunately, this distinction between basic cognitive understanding and aesthetic interpretation often becomes muddled, especially in studies from the 'semiotic' tradition (e.g. Miodrag 2013; Postema 2013). Such confusion is damaging to the understanding of 'theory' and 'how comics work' – both to people within and outside of this field. Certainly, notions from cognitive science can frame ways to appreciate the aesthetics of comics as literary, artistic, or cultural objects. As discussed, Kukkonen's (2013b) paper does this quite well. However, the reverse is less true: Groensteen's concepts of arthrology and braiding, for example, may be useful for people wishing to discuss the aesthetics of comics as literature, though they contribute little to understanding how the mind/brain comprehends sequences of images in any cognitive sense.

Realizing this distinction – and openly acknowledging it – is important for establishing the sub-domains of how people study comics. Neither of these approaches is necessarily 'better' in an absolute sense, but each approach can be better or worse given whatever particular topic a scholar is aiming to address. Ultimately, the questions that a researcher asks should motivate how they go about studying that topic. The study of comics has always been interdisciplinary, and that broad scope will naturally invite appropriation and discussion of ideas from various different perspectives. This is a good thing, and at its best it should lead to a deeper understanding of comics as a literary, artistic, cultural, *and* cognitive phenomenon. Yet, along with this interdisciplinary nature is the need for recognizing which questions can be addressed by which methods, and the demand for negotiating how to interpret findings across numerous disciplines. These challenges need to be acknowledged as the study of comics now progresses beyond its infancy.

## Acknowledgements

Kent Worcester is thanked for helpful feedback on previous drafts.

## References

- Abbott, L. L. (1986), 'Comic art: Characteristics and potentialities of a narrative medium', *The Journal of Popular Culture*, 19: 4, pp. 155–76.
- Abbott, M. and Forceville, C. (2011), 'Visual representation of emotion in manga: Loss of control is Loss of hands in *Azumanga Daioh* Volume 4', *Language and Literature*, 20: 2, pp. 91–112.



- Bloomfield, L. (1933), *Language*, New York, NY: Holt, Rinehart and Winston.
- Branigan, E. (1992), *Narrative Comprehension and Film*, London, UK: Routledge.
- Carroll, J. M. (1980), *Toward a Structural Psychology of Cinema*, The Hague: Mouton.
- Cohn, N. (2003), *Early Writings on Visual Language*, Carlsbad, CA: Emaki Productions.
- (2010), 'The limits of time and transitions: Challenges to theories of sequential image comprehension', *Studies in Comics*, 1: 1, pp. 127–47.
- Cohn, N., Paczynski, M., Jackendoff, R., Holcomb, P. J. and Kuperberg, G. R. (2012), '(Pea)nuts and bolts of visual narrative: Structure and meaning in sequential image comprehension', *Cognitive Psychology*, 65: 1, pp. 1–38.
- de Saussure, F. (1972), *Course in General Linguistics* (trans. R. Harris), Chicago, IL: Open Court Classics.
- Gernsbacher, M. A., Varner, K. R. and Faust, M. (1990), 'Investigating differences in general comprehension skill', *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16: 3, pp. 430–45.
- Groensteen, T. (1999), *Système de la bande dessinée/The System of Comics*, Paris: Presses Universitaires de France.
- (2007), *The System of Comics* (trans. B. Beaty and N. Nguyen), Jackson: University of Mississippi Press.
- Gubern, R. (1972), *El lenguaje de los Comics/The language of the comics*, Barcelona: Peninsula.
- Harris, R. A. (1993), *The Linguistics Wars*, New York, NY: Oxford University Press.
- Hick, D. H. (2012), 'The language of comics', in A. Meskin and R. T. Cook (eds), *The Art of Comics: A Philosophical Approach*, West Sussex, UK: Wiley-Blackwell. pp. 125–44.
- Hockett, C. F. (1960), 'Logical considerations in the study of animal communication', in W. E. Lanyon and W. N. Tavolga (eds), *Animal Sounds and Communication*, Washington, DC: American Institute of Biological Sciences, Symposium Series Number 7, pp. 392–430.
- Horstkotte, S. (2013), 'Zooming in and out: Panels, frames, sequences, and the building of graphic storyworlds', in D. Stein and J.-N. Thon (eds), *From Comic Strips to Graphic Novels: Contributions to the Theory and History of Graphic Narrative*, Berlin: de Gruyter, pp. 27–48.
- Hünig, W. K. (1974), *Strukturen des Comic Strip/Structure of comic strips*, Hildensheim: Olms.
- Hutto, D. D. (2013), 'Action understanding: How low can you go?', *Consciousness and Cognition*, 22: 3, pp. 1142–51.
- Jackendoff, R. (1994), *Patterns in the Mind: Language and Human Nature*, New York: Basic Books.



- (2002), *Foundations of Language: Brain, Meaning, Grammar, Evolution*, Oxford: Oxford University Press.
- Johnson, M. (1987), *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason*, Chicago: University of Chicago Press.
- Johnson-Laird, P. N. (1983), *Mental Models*, Cambridge, MA: Harvard University Press.
- Koch, W. A. (1971), *Varia Semiotica*, Hildensheim: Olms.
- Kukkonen, K. (2008), 'Beyond language: Metaphor and metonymy in comics storytelling', *English Language Notes*, 46: 2, pp. 89–98.
- (2013a), 'Space, time, and causality in graphic narratives: An embodied approach', in D. Stein and J.-N. Thon (eds), *From Comic Strips to Graphic Novels: Contributions to the Theory and History of Graphic Narrative*, Berlin: de Gruyter, pp. 49–66.
- (2013b), 'Navigating infinite earths: Readers, mental models, and the multiverse of superhero comics', in C. Hatfield, J. Heer and K. Worcester (eds), *The Superhero Reader*, Jackson: University of Mississippi Press, pp. 155–69.
- Liddell, S. K. (2003), *Grammar, Gesture, and Meaning in American Sign Language*, Cambridge: Cambridge University Press.
- Magliano, J. P. and Zacks, J. M. (2011), 'The impact of continuity editing in narrative film on event segmentation', *Cognitive Science*, 35: 8, pp. 1489–517.
- Magliano, J. P., Miller, J. and Zwaan, R. A. (2001), 'Indexing space and time in film understanding', *Applied Cognitive Psychology*, 15: 5, pp. 533–45.
- Mallia, G. (2007), 'Learning from the sequence: The use of comics in instruction', *ImageText*, 3: 3, [http://www.english.ufl.edu/imagetext/archives/v3\\_3/mallia/](http://www.english.ufl.edu/imagetext/archives/v3_3/mallia/).
- McCloud, S. (1993), *Understanding Comics: The Invisible Art*, New York, NY: Harper Collins.
- Medley, S. (2010), 'Discerning pictures: How we look at and understand images in comics', *Studies in Comics*, 1: 1, pp. 53–70.
- Miodrag, H. (2013), *Comics and Language: Reimagining Critical Discourse on the Form*, Jackson: University Press of Mississippi.
- Mitchell, W. J. T. (1986), *Iconology: Image, Text, Ideology*, Chicago, IL: University of Chicago Press.
- Nagai, M., Endo, N. and Takatsune, K. (2007), 'Measuring brain activities related to understanding using near-infrared spectroscopy (NIRS)', in M. J. Smith and G. Salvendy (eds), *Human Interface and the Management of Information: Methods, Techniques and Tools in Information Design*, vol. 4557, Heidelberg: Springer Berlin, pp. 884–93.

- Nakazawa, J. (2005), 'Development of manga (comic book) literacy in children', in D. W. Shwalb, J. Nakazawa and B. J. Shwalb (eds), *Applied Developmental Psychology: Theory, Practice, and Research from Japan*, Greenwich, CT: Information Age Publishing, pp. 23–42.
- Nöth, W. (1990), *Comics Handbook of Semiotics*, Indianapolis, IN: University of Indiana Press, pp. 472–75.
- Pinker, S. (1994), *The Language Instinct: How the Mind Creates Language*, New York: HarperCollins.
- Postema, B. (2013), *Narrative Structure in Comics: Making Sense of Fragments*, Rochester, NY: RIT Press.
- Pratt, H. J. (2009), 'Narrative in comics', *Journal of Aesthetics and Art Criticism*, 67: 1, pp. 107–17.
- Rizzolatti, G. and Arbib, M. A. (1998), 'Language within our grasp', *Trends in Cognitive Neuroscience*, 21: 5, pp. 188–94.
- Saraceni, M. (2000), 'Language beyond language: Comics as verbo-visual texts', dissertation, Nottingham: University of Nottingham.
- (2001), 'Relatedness: Aspects of textual connectivity in comics', in J. Baetens (ed.), *The Graphic Novel*, Leuven: Leuven University Press, pp. 167–79.
- Trout, J. D. (2008), 'Seduction without cause: Uncovering explanatory neuroophilia', *Trends in Cognitive Sciences*, 12: 8, pp. 281–82.
- Turella, L., Pierno, A. C., Tubaldi, F. and Castiello, U. (2009), 'Mirror neurons in humans: Consisting or confounding evidence?', *Brain and Language*, 108: 1, pp. 10–21.
- Walsh, R. (2006), 'The narrative imagination across media', *MFS Modern Fiction Studies*, 52: 4, pp. 855–68.
- Zwaan, R. A. and Radvansky, G. A. (1998), 'Situation models in language comprehension and memory', *Psychological Bulletin*, 123: 2, pp. 162–85.

### **Suggested citation**

Cohn, N. (2014), 'Building a better 'comic theory': Shortcomings of theoretical research on comics and how to overcome them', *Studies in Comics* 5: 1, pp. 57–75, doi: 10.1386/stic.5.1.57\_1

### **Contributor details**

Neil Cohn is an American cognitive scientist whose research investigates the overlap of the structure and cognition of sequential images and language. He received his doctorate in Psychology from Tufts University, and his most recent book, *The Visual Language of Comics* (Bloomsbury, 2013),

introduces a broad framework for studying comics and visual narratives in the linguistic and cognitive sciences. His work is online at [www.visuallanguagelab.com](http://www.visuallanguagelab.com).

Contact: Center for Research in Language, University of California, San Diego, 9500 Gilman Dr. Dept. 0526, La Jolla, CA 92093-0526, USA.

E-mail: [neilcohn@visuallanguagelab.com](mailto:neilcohn@visuallanguagelab.com)

Neil Cohn has asserted his right under the Copyright, Designs and Patents Act, 1988, to be identified as the author of this work in the format that was submitted to Intellect Ltd.

---