Choosing the Easy Way Out – Neil Cohn and his Visual Lexicons

by Fred Andersson

Cohn, Neil. *The Visual Language of Comics: Introduction to the Structure and Cognition of Visual Images*. Bloomsbury, Bloomsbury Academic Press, 2013. ISBN 978-1-4411-7054-5. 221 pages.

Neil Cohn's theory of visual "language" in comics is rather well known by now. *The Visual Language of Comics: Introduction to the Structure and Cognition of Visual Images* is the most unified and accessible explanation of the theory this far. It was published in 2013 and received considerable attention, even in "mainstream" media such as *The Guardian*. There are some unresolved contradictions in the book. Already in the introduction, Cohn repeatedly insists that he does not claim that comics are "a language." As for myself, I basically agree with Cohn's definition that visual language is "a biological and cognitive capacity that humans have for conveying concepts in the visual-graphic modality" and that comics are "a socio-cultural context in which this visual language appears" (p. 2). However, the title of the book is still *The Visual Language of Comics*, which says exactly the opposite of what the author claims. *Visual Language in Comics* would have been a less problematic title. Interestingly, Cohn claims that there are regional "visual languages" such as AVL (American Visual Language) and JVL (Japanese Visual Language). In the last chapter, he also analyzes the ancient "language" CAVL (Central Australian Visual language). This is problematic. On the one hand, Cohn claims that there is a universal "biological and cognitive capacity" called visual language. On the other hand, he claims that there are different visual languages. What are the exact criteria for separating the latter from the former? This is never clarified in the book.

Another salient contradiction of Cohn's theory is his attitude to linguistic analogy. Obviously his whole project is based on analogies between words and pictures – even to the extent that they are not treated as analogies but simply stated as objective facts. Still, he rejects the dominant tenet of linguistic analogy in pictorial semiotics, i.e. the one developed within the tradition of Structuralist and Saussurean "semiology." According to Cohn, the "semiological" conception of signs and Language is erroneous and outdated. As outlined by Ferdinand de Saussure and developed by Louis Hjelmslev, structural linguistics upholds a strict separation between Signifier and Signified, Denotation and Connotation, Form and Substance, System (or "paradigmatics") and Process (or "syntagmatics"). This means that certain operations must be made in order for the analysis to hold. For example, distinctive features must be identified. It must be shown that distinctive features such as "B" and "F" can be substituted, and that this results in a change of semantic content. Example: "BAR" changes into "FAR." If such "commutation" (i.e. simultaneous mutation) cannot be proved, the message cannot be a sign according to the structural definition, but merely an element of "symbolic language." Linguistic analogy in generalized "semiology" has often

consisted in identifying certain visual or "plastic" features in pictures, pretending that these features are genuinely distinctive, and then claiming this as a fact.

Chomskyan linguistics rejects the structural dogma of arbitrary selection. Instead of seeing the "lexeme" (i.e. the "word" or "term") as a fixed bundle of semantic and phonetic features, a Chomskyan analysis treats both "semantics" and "grammar" as collections of cognitive categories, which can be systematized as "lexicons". Consequently, the Latin alphabet is seen as a grammatical lexicon, which every child must learn in order to write English. One semantic example would be that all associations between color terms in a language could be regarded as a semantic lexicon. The grammatical and semantic dimensions are seen as independent, *malgré* Structuralist priorities, as famously exemplified by Chomsky's phrase "colourless green ideas sleep furiously." Here, the phrase is grammatically correct – and therefore a normal example of the "lexicon" of English syntax – but semantically nonsensical. With such wide definitions of the terms grammar, semantics, and lexicon, a scholar of visual communication can easily find "visual lexicons" at various "grammatical" and "syntactic" levels, ignoring the analysis of Signifier and Signified. In this rather loose manner, Cohn seems to conceive of visual language as a variety of Chomskyan universal grammar, thus embracing linguistic analogy while at the same time rejecting it.

The general argument and disposition of Cohn's book follows his theoretical orientation. He sets out to prove that panels in comics are grammatical units, comparable to words and phrases. If they are grammatical units in a Chomskyan sense, it must be proved that readers can register the systematic order between panels as normal even if the semantic content is scrambled and nonsensical. This is attempted empirically, by means of EEG (electroencephalogram) and VEP (visually evoked potentials) in an experimental setup known from psycholinguistic studies. Before launching these tests in Chapter Six, however, Cohn spends the first half of the book explaining his theories. Most of the speculations in these chapters are not relevant for the empirical question. In the first chapter, Cohn defines the general structure of visual language, especially with regard of "navigational structure" within the strip or page in a comic. The second and third chapters deal with two levels of "the visual lexicon," i.e. the morphological level and the level of panels, respectively. Here, Cohn also reviews the contrasting semiotic theories of Ferdinand de Saussure and Charles Sanders Peirce, and rejects two hypothetical objections against the idea that panels can be lexical units. Cohn rejects the first objection – "Panels are not arbitrary signs" – simply by stating that the whole notion of arbitrary signs is an outdated dogma. The second objection – "There is no systematic lexicon of panels" – is rejected with reference to cognitive linguistics. Cohn admits that the number of conventionalized panels is small, but adds that "we find even more systematic patterns in the components that make up panels and the broader

sequences in which they are placed" (p. 22). For Cohn, a visual lexical unit is essentially a pattern which is stored in visual memory and therefore repeated.

Cohn then defines the elements of visual morphology and syntax, inventing a number of idiosyncratic neologisms. Cohn's rejection of the notion of arbitrary signs means that certain criteria for identifying morphemes in Language are abandoned. This is obviously the only justification for Cohn's broad claim that there *are* indeed morphemes in visual language, and that there are both "open-class" and "closed-class" morphemes. Open linguistic classes are those which are relatively flexible and variable, as when new synonyms are added to a vocabulary. Closed classes, such as markers for case and gender, will only change over long spans of time. When Cohn simplistically tries to demonstrate that "graphic schemas" of "basic objects" are open-class morphemes, he does not see the fuzziness of pictorial categories in comparison to linguistic ones as a problem. Instead, he tries to show that pictures are not very different from spoken languages in this respect, as there are both "analytic" languages like English and "synthetic" languages like Finnish and Inuit. For example, an Englishman needs three words in order to ask the same thing that a Finn asks in one: "ymmärrätkö?" ("do you understand?"). That such merging or "agglutination" may be observed both in language and in pictures does not however presuppose that the constituent elements are in any way similar. Thus, Cohn again justifies his brand of linguistic analogy by ignoring analytical levels favored by other linguistic analogies.

As regards closed-class morphemes, Cohn thinks that elements such as speech bubbles and speed lines are equivalent to prepositions and affixes in Language. Speech bubbles are called "carriers", their connectors are called "tails" and the speakers/thinkers are called "roots." If someone in a comic says something that other characters can hear, it is said to be a combination of "root awareness" and "adjacent awareness." For example, a question mark or a light bulb appearing above a character's head is called "upfix"; speed indicators replacing body parts or the whole body are called "umlaut" or "suppletion"; a repetition of body parts or contours to indicate movement is called "reduplication." As regards whole panels, Cohn does not locate them specifically at the morphological level or the syntactical level, but regards them as "constructions." They are treated as "attentional units," framing certain objects, characters and events at the same time as providing orientation within the spatial setting of a scene. Thus, they are categorized as to whether they are "macro" (several characters), "mono" (one character), "micro" (detail of character), or "amorphic" (inanimate detail). As "filmic shots," they are categorized as Long Shot, Full Shot, Medium Shot, Close Shot, and (extreme) Close-Up.

For Cohn, narrative structure is the site of "Visual Language Grammar" proper. It is defined in Chapter Four. Lay-out, designated with the acronym ECS (External Compositional Structure), is

the topic of Chapter Five. It is odd that whilst Cohn defines and analyzes five categories of whole page ECS, the only cases considered in the chapter on narrative structure are those of single strips. The importance of page breakdown and braiding for narrative structure, as convincingly shown by Thierry Groensteen, is completely ignored by Cohn. The problem of reading order is addressed in the chapter on ECS, but this chapter remains isolated from the main argument.

Cohn's conclusions are based on very simple cases of narrative visual sequences, cases which are also used as stimuli for the tests described in Chapter Six. The cases are actually pure "constructions," drawn by Cohn himself. Examples from the work of other artists are scarcely used and comparative approaches are absent. Applying the diagrammatic approach of Chomskyan syntactic analysis, Cohn states that "visual narrative grammar" is structured as an "arc" involving all or some of the following panel types in the following order: Orienter, Establisher, Initial, Prolongation, Peak, and Release. Sometimes, modifications or Refiners will also appear. As the terms indicate, this taxonomy best describes straightforward event structures in which an action or conflict is anticipated, launched, continued, reaching climax, and resolved. None of the test cases are longer than six panels. In the tests, Cohn proves with some reliability that panels showing stages of the same event but in the wrong or fragmentary "grammatical" order produce a positive signaling change 600 milliseconds after the onset of the stimulus. This is the so called "P600" effect known from violations of grammar in psycholinguistic experiments. Similarly, it is shown that panels in correct "grammatical" order but taken from different events produce the "N400" effect known from violations of semantic order.

These tests may be thoroughly controlled and the results may be fully reliable. But what do they actually prove? They obviously prove that we expect simple things to happen in a certain order and that it would be highly confusing if persons and objects were suddenly replaced during an event. This is nothing new, and there is nothing here that specifically characterizes drawn events as opposed to real ones, not to mention "Visual Language Grammar." What Cohn manages to prove, then, is already known by common experience.

The rest of his theory of "Visual Language" remains purely speculative and, indeed, arbitrary. This is not least true about the last three chapters and their analyses of the regional "languages" AVL, JVL and CAVL. Here, Cohn enumerates standardized graphical features and figurative "morphemes" in a way strangely reminiscent of the Structuralist analyses which he believes himself to have rejected. As for CAVL, the ancient Australian practice of accompanying a verbal narrative with markings in the sand has a multimodal structure very different from modern comics. If Cohn wants to define universal traits of all visual Language, his arguments are constantly disproved by the empirical facts of many different languages. In the end, the narrow

significance of Cohn's psychological experiments and his reluctance to consider other examples than those invented by himself gives his project a very limited scholarly value.