

MICHAŁ SZAWERNA  
University of Wrocław

## The Scope of the Pictorial and the Verbal in Multimodal Cognitive Linguistic Accounts of Visual Signification in Modal Ensembles Comprising Static Planar Signs

### Abstract

Since early 2000s, various modal ensembles (verbo-gestural utterances, feature films, animations, political cartoons, printed advertisements, television commercials, comics, picture books, computer games, pieces of music, corporate trademarks and logos, medieval textiles, *etc.*; *cf.* Forceville and Urios-Aparisi 2009; Pinar Sanz 2015) have been the subject of scholarly exploration in multimodal cognitive linguistics — a burgeoning research field situated at the intersection of multimodality studies and cognitive linguistics — with a view to explicating how general cognitive mechanisms shape meanings communicated across modalities and providing additional evidence for the psychological reality of various theoretical and descriptive notions put forward by cognitive linguists. Given the goals of multimodal cognitive linguistics, it is not surprising that research in this field has not only highlighted certain characteristics of the analyzed modal ensembles, but has also de-emphasized or altogether hidden others. In this article, an attempt is made to examine a number of contributions to the strand of multimodal cognitive linguistics that focuses on the cognitive underpinnings of various static planar ensembles (printed advertisements, cartoons, comics, corporate logos, *etc.*) with regard to what these contributions tacitly assume to be included in the scope of the pictorial and the verbal — two key concepts of multimodal cognitive linguistics that have never been explicitly characterized by researchers in this strand. This attempt is undertaken in order to demonstrate that in the examined contributions the two concepts are implicitly made to subsume representations that are qualitatively so different that there is little reason for bundling them together, which is in turn intended to alert researchers and readers alike to the problems caused by the imposition of a binary construal ('either verbal or pictorial') onto a continuum of qualitatively diverse static planar signs. This article further shows that an alternative characterization of this semiotic continuum — one that is free of the problems engendered by construing the continuum in binary terms — has for a long time existed in the field of comics scholarship. In conclusion, it is suggested that a more interdisciplinary approach may help multimodal cognitive linguists avoid unwarranted oversimplifications in the future.

**Keywords:** multimodality, cognitive linguistics, comics, modal ensembles, verbal representations, pictorial representations, semiotics

## 1. Introduction

For two and a half decades, a wide range of “modal ensembles” (Bezemer and Kress 2016: 6), that is, complex representations that either function as transitory means of face-to-face communication<sup>1</sup> or come in the form of more or less permanent “semiotic artefacts”<sup>2</sup> (Bateman 2008: 38), have been the subject of lively scholarly interest in a progressively expanding field of research that has come to be known as *multimodal cognitive linguistics*.<sup>3</sup> While the analytical scope of multimodal cognitive linguistics has become broader and broader, research in this field has always been unified by a methodological principle which consists in applying selected instruments of theories subsumed by cognitive linguistics<sup>4</sup> with a view to characterizing the meanings communicated by the analyzed modal ensembles. This principle is motivated by a central tenet of cognitive linguistics whereby the meanings communicated by linguistic expressions are shaped by the same mechanisms that guide non-linguistic cognition.<sup>5</sup> On the whole, then, the primary goal of multimodal cognitive linguistics has been to explicate how general cognitive mechanisms shape the meanings of a broad range of modal ensembles, but its secondary goal, determined by the above-mentioned tenet, has been to provide additional evidence for the psychological reality of various theoretical and descriptive notions proposed by cognitive linguists.

It stands to reason that the goals of multimodal cognitive linguistics have largely determined not only what research in this field has highlighted about the analyzed modal ensembles, but also what it is about them that it has hidden, de-emphasized, glossed over, or taken for granted. In the strand of multimodal cognitive linguistics that focuses on the cognitive underpinnings of modal ensembles made up of static planar signs (printed advertisements, cartoons, comics,<sup>6</sup> *etc.*), the scope of two concepts that are pivotal to any discussion of these static planar ensembles — the pictorial and the verbal — has never been explicitly delineated, with the result that in some cases the concepts are tacitly made to subsume signs that are qualitatively so different that there is little reason for bundling them together. Arguably, the principal reason why qualitatively diverse signs have been repeatedly bundled together in this strand of research — sometimes to the point of arbitrary oversimplification — is descriptive convenience. If the principal goal is to identify the cognitive underpinnings of a static planar ensemble, it is no doubt more convenient to discuss the ensemble’s form using two well-established, more or less intuitively understood

---

1 Such as the so-called “verbo-gestural utterances” discussed by Müller and Cienki (2009: 300).

2 Such as feature films, TV commercials, pieces of music, printed advertisements, political cartoons, comics, trademarks, logos, *etc.*

3 The term *multimodal cognitive linguistics* is relatively recent; it seems to have been used for the first time by Langlotz (2015: 55) and Moya Guijarro (2015: 117).

4 With noticeable emphasis on conceptual metaphor theory, initiated by Lakoff and Johnson (1980), image schema theory, originally formulated by Johnson (1987), and conceptual integration theory, also known as blending theory, articulated most fully by Fauconnier and Turner (2002).

5 Most notably, the interrelated mechanisms of conceptual metaphorization, conceptual metonymization, and conceptual integration.

6 In English, the form *comics* refers to the entire medium when it is used with a singular verb, as in *Comics is a fascinating medium* (Phoenix 2020: 2), or to more than one publication in this medium when it is used with a plural verb, as in *Comics are fun to read*.

concepts than to risk making the characterization less accessible by admitting that the form is not fully characterizable as a combination of pictorial and verbal signs.

In this article, an attempt is made to compare, contrast, and evaluate a number of contributions to the strand of multimodal cognitive linguistics that investigates the cognitive underpinnings of various static planar ensembles (most notably, Forceville 1996, 2005, 2011; Górska 2020; but also Caballero 2009; El Rafaie 2009; Koller 2009; Yus 2009) with regard to what these contributions tacitly assume to be included in the scope of the pictorial and the verbal. This attempt is made with a view to alerting researchers and readers alike to the problems caused by the imposition of a binary ('either-or') construal onto what is essentially a continuum of qualitatively diverse static planar signs. It is also shown here that an alternative characterization of this continuum — one that is largely free of the problems engendered by construing it in binary terms — has for a long time existed in the field of comics scholarship and suggested that a more interdisciplinary approach to analyzing static planar ensembles (or, indeed, modal ensembles of any kind) on the part of multimodal cognitive linguists may help them avoid unwarranted oversimplifications in the future.

Structurally, this introduction is followed by a section explaining why the everyday notion of pictoriality, whereby pictures are understood as visually mimetic paintings, drawings, and photographs, has been adopted here as the standard relative to which the contributions to multimodal cognitive linguistics included in the scope of this article are compared, contrasted, and evaluated (Section 2). This is in turn followed by a brief outline of Peirce's sign theory, which has been adopted here as a theoretical context in which the workings of modal ensembles will be discussed (Section 3). In the following section, pictorial and verbal representations are characterized in terms of Peirce's sign theory in order to facilitate the discussion of the scope of the pictorial and the verbal in multimodal cognitive explorations of static planar ensembles (Section 4). This is followed by a section discussing the scope of the pictorial and the verbal in the strand of multimodal cognitive linguistics that focuses on the cognitive underpinnings of static planar ensembles (printed advertisements, cartoons, comics, *etc.*) (Section 5). This section identifies problems caused by the imposition of a binary ('either-or') construal onto what is essentially a continuum of qualitatively diverse static planar signs. In the following section, an alternative characterization of this continuum is presented and its advantages over the binary construal espoused by multimodal cognitive linguists are discussed (Section 6). The final section recapitulates the points made in the main body of this article (Conclusion). It ends with a plea for a more interdisciplinary approach to analyzing static planar ensembles on the part of multimodal cognitive linguists, which may help them avoid unwarranted oversimplifications in the future.

## 2. The notion of pictoriality adopted in this article

At this juncture it is important to emphasize that none of the contributions to multimodal cognitive linguistics included in the scope of this examination subscribes, either explicitly or implicitly, to any particular theory of pictorial representation. Consequently, it would be unjustified to evaluate these contributions with reference to the tenets of one theory of pictorial representation or another, especially given that despite their unquestionable merits, none of the theories can successfully tackle all of the challenges posed by pictorial representations (Rollins 2001). Instead, it makes sense to evaluate these

contributions with reference to everyday understanding of what pictures are and how they work. This everyday understanding is perhaps best encapsulated in dictionary definitions of the word *picture*, which largely overlap in their content. Dictionaries of contemporary English concur that the word *picture* refers to three major kinds of static planar representations: paintings, drawings, and photographs. Dictionary entries for *picture* typically highlight the representational nature of pictures, whereby a picture is necessarily a picture of something (e.g. “a painting or drawing, *etc.* that shows a scene, a person or thing” OALD), and the visually mimetic nature of the representation, whereby pictures reproduce visual perception (e.g. “shapes, lines *etc.* painted or drawn on a surface, showing what someone or something looks like” LDOCE). These characteristics of pictures are evident in the examples of the word’s usage provided in dictionary entries, such as *Atticus drew/painted a picture of my dog* (CALD), *There’s a picture of his wife above the fireplace* (LDOCE), and *The picture shows the couple together on their yacht* (OALD).

The non-expert, everyday understanding of *picture* encapsulated in multiple dictionaries of contemporary English seems to have guided expert discussions of visual mimeticity in static planar representations. For example, in *Art and Illusion: A Study in the Psychology of Pictorial Representation* (1960) Gombrich discussed the interplay between visual perception and representational conventions in art on the basis of paintings, drawings, and photographs that are by and large visually mimetic, with the proviso that their visual mimeticity is a matter of degree as it is variously affected by the materials used, the techniques applied, and the styles adopted by the artists. In fact, visual mimeticity, whereby the viewer can recognize what a picture represents because the representation looks like the thing signified, seems to be regarded by experts of various theoretical persuasions as a relatively uncontroversial characteristic of pictures. In an overview of rival theories of pictorial representation, Rollins (2001) explains that the more sophisticated, perceptual accounts of pictorial representation<sup>7</sup> may vary in the degree of importance they place on the role of the viewer’s prior knowledge (of what physical objects look like, but also of pictures and the representational conventions they employ) in the process whereby a picture achieves visual mimeticity, but they seem to concur that visual mimeticity is a defining feature of pictures, especially when problematized as residing in correspondences established in the mind of the viewer between the mental representation of a picture and the mental representation of the thing signified.

Occasionally, however, the word *picture* has its definitional boundaries redrawn by virtue of being contrasted with another word in a terminological system devised by a particular theorist. For example, in *Visual Thinking* Arnheim (1969: 136) characterized *pictures* in opposition to *symbols*, as contrasting “functions fulfilled by images.” In Arnheim’s (1969) terminological system, the contrast between *pictures* and *symbols* resides in the fact that while pictorial images “portray things located at a lower level of abstractness than they are themselves” (1969: 137), symbolic images portray “things which are at a higher level of abstractness” (1969: 138) than they are themselves. Arnheim (1969) explained further that images usually perform both functions simultaneously:

Holbein’s portrait of Henry VIII is a picture of the king, but it also serves as a symbol of kingship and of qualities such as brutality, strength, exuberance, which are located at a higher level of abstraction than is the painting. The painting, in turn, is more abstract than the visual appearance of the king in

---

<sup>7</sup> These accounts of pictorial representation reject the less sophisticated, non-perceptual accounts: both the strong version of the so-called conventionality view, whereby the link between pictures and what they represent is regarded as completely arbitrary, and the strong version of the so-called resemblance view, whereby pictures signify by virtue of an objectively existing similarity between them and the things they represent (for details, consult Rollins 2001).

flesh and blood because it sharpens the formal features of shape and color which are analogues of the symbolized qualities. (Arnheim 1969: 138–139)

One consequence of Arnheim's (1969) redefinition of *picture* is that an image may be entirely abstract, i.e. visually non-mimetic, and still qualify as a pictorial representation: "[A] totally non-mimetic geometrical pattern by Mondrian may be intended as a picture of the turmoil of New York's Broadway" (Arnheim 1969: 137).<sup>8</sup>

For reasons indicated above, the scope of the pictorial outlined tacitly in multimodal cognitive linguistic investigations of the cognitive underpinnings of static planar ensembles (printed advertisements, cartoons, comics, *etc.*) will be examined in the analytical part of this article with reference to the everyday understanding of what pictures are and how they work. This everyday understanding of pictures emphasizes their visual mimeticity, but visual mimeticity is a variable feature of pictorial representations, and it interrelates with other features exhibited by pictures: most notably, their conventionality. To facilitate the discussion of the scope of the pictorial in the analytical part of this article, the analytical part is preceded by a systematic, albeit concise, discussion of the interrelated features of pictorial representations. This discussion is framed in terms of Peirce's theory of signs — a comprehensive semiotic theory applicable to all signs (irrespective of their materiality and mode of signification) that has been shown to harmonize rather well with the tenets of cognitive linguistics (*cf.* Danaher 1998; Szawerna 2017: 101–107). A brief outline of Peirce's sign theory, with emphasis on the most important modes of signification: *icon*, *index*, and *symbol* (Peirce CP § 2.275), is offered below.

### 3. Peirce's sign theory and modes of representation

Seen from the Peircean perspective, signs "exist only in the mind of the interpreter" (Nöth [1990] 1995: 42), and they are pivotal to reasoning, or *ratiocination* (Peirce CP § 4.45),<sup>9</sup> that is, "the operation by which reason proceeds from the known to the unknown" (Peirce CP § 4.45), because "the only way we can acquire new knowledge is [...] by means of signs" (Jappy 2013: 2). This key function of the sign, whereby it is "something by knowing which we know something more" (Peirce CP § 8.332), is carried out through the action of *semiosis*, which involves the cooperation of three correlates: "a sign, its object, and its interpretant" (Peirce CP § 5.484). The three correlates, also known as, respectively, "sign, thing signified, cognition produced in the mind" (Peirce CP § 1.372; emphasis in the original), participate in semiosis in such a way that "the first term represents the second term to a third term or, more concretely,

8 Of course Arnheim's (1969) formulations are by no means immune to critique. In this case, for example, it is not at all obvious that the "geometrical pattern by Mondrian" (Arnheim 1969: 137) (presumably, a reference to Mondrian's 1943 painting *Broadway Boogie Woogie*) qualifies as "totally non-mimetic" (Arnheim 1969: 137). It may be difficult for viewers to recognize the painting as a visually mimetic representation of New York's Broadway, but it does look like an aerial view (albeit highly stylized) of the downtown area of a large American city, with streets running perpendicular to each other, forming a characteristic grid. And if Mondrian's geometrical pattern were indeed totally non-mimetic, as Arnheim (1969: 137) would have it, would it qualify as a picture only because it was "intended as a picture" (Arnheim 1969: 137) by the artist, irrespective of how it appeared to viewers? Is Arnheim's (1969: 137) suggestion that visual mimeticity is a matter of creative intention at all tenable?

9 The references to Peirce's *Collected Papers* (1931–1958) indicate volumes and paragraphs. For example, "CP § 2.308" refers to volume 2, paragraph 308.

the sign represents the object to the interpretant” (Jappy 2013: 3). According to Nöth ([1990] 1995: 43), Peirce (CP § 2.228) regarded the interpretant, variously characterized as “cognition produced in the mind” (Peirce CP § 1.372; emphasis in the original) and “the proper significate outcome of a sign” (Peirce CP § 5.473), as another sign. Nöth ([1990] 1995: 43) explains that because the action of the Peircean sign involves an interpretant which constitutes another sign, which in turn involves another interpretant, and so on, potentially “*ad infinitum*” (Peirce CP § 2.92; original emphasis), Peircean semiosis is *unlimited* (Peirce CP §§ 2.303, 2.92).<sup>10</sup> Nöth ([1990] 1995: 43) further points out that unlimited semiosis, which presupposes neither the very first nor the very last sign, does not imply a vicious circle, but instead exemplifies Peirce’s idea of the dialogic nature of the thought process.

Peirce (CP §§ 2.233–2.271) classified signs by relating the three correlates of the sign (the sign, the object, and the interpretant) to his three universal categories: *firstness*, which comprises qualities; *secondness*, which comprises relations; and *thirdness*, which comprises representations (Peirce CP §§ 1.300–1.353, 8.328–8.332). This resulted in the emergence of three trichotomies of signs:

Signs are divisible by three trichotomies; first, according as the sign in itself is a mere quality, is an actual existent, or is a general law; secondly, according as the relation of the sign to its object consists in the sign’s having some character in itself, or in some existential relation to that object, or in its relation to an interpretant; thirdly, according as its Interpretant represents it as a sign of possibility or as a sign of fact or a sign of reason. (Peirce CP § 2.243)

Of the three trichotomies, the second one is “the most fundamental” (Peirce CP § 2.275). It introduced the seminal “modes of representation” (Jappy 2013: 79) — icon, index, and symbol — which were distinguished on the basis of how the first correlate (the sign) is related to the second correlate (the object). Importantly, icons, indices, and symbols are not mutually exclusive categories of the Peircean sign, and “any single sign may display some combination of iconic, indexical and symbolic characteristics” (Atkin [2006] 2010).

Peirce’s icon belongs to firstness because it “signifies by its own quality” (Nöth [1990] 1995: 121). As an icon can only signify by sharing a signifying quality with its object, it is necessarily based on similarity (Peirce CP § 2.276). The problem, however, is that signifying qualities never exist alone, but are instead always embodied, so that a pure icon remains a “possibility alone” (Peirce CP 2.276). This means that an actual sign, that is, “an actual existent thing or event which is a sign” (Peirce CP § 2.245), cannot be a pure icon although it “may be *iconic*, that is, [it] may represent its object mainly by its similarity, no matter what its mode of being” (Peirce CP 2.276; original emphasis). In order to distinguish between pure icons and actual iconic signs, Peirce dubbed the latter *hypoicons* (CP 2.276). While Peirce repeatedly emphasized the iconic criterion of similarity between hypoicons and their objects, his understanding of similarity was so broad as to include “similarity of abstract relations or structural homologies” (Nöth [1990] 1995: 122).<sup>11</sup> Additionally, Peirce insisted on “referential openness” (Nöth [1990] 1995: 123) of hypoicons, whereby they do not necessarily presuppose the existence of their objects: “The Icon does not stand unequivocally for this or that existing thing [...]. Its Object may be a pure fiction, as to its existence” (CP § 4.531).

<sup>10</sup> Some semioticians, e.g. Jappy (2013: 20), prefer to describe Peircean semiosis as *continuous* rather than *unlimited*.

<sup>11</sup> This is evident not only in Peirce’s definitions of the icon (CP §§ 2.299, 3.362, 4.531), but also in what he considered prime examples of iconic signs: portraits (CP § 2.92), photographs (CP § 2.281), ideographs (CP § 2.280), diagrams, graphs (CP §§ 4.418–4.420), and algebraic formulas (CP § 2.279).

Peirce's broad conception of similarity led to his division of hypoicons into *images*, which involve qualitative similarity; *diagrams*, which involve relational similarity; and *metaphors*, which involve representational similarity:

Hypoicons may be roughly divided according to the mode of Firstness of which they partake. Those which partake in simple qualities, or First Firstnesses, are *images*; those which represent the relations, mainly dyadic, or so regarded, of the parts of one thing by analogous relations in their own parts, are *diagrams*; those which represent the representative character of a representamen by representing a parallelism in something else, are *metaphors*. (Peirce CP § 2.277; original emphasis)

It is a well-established fact that Peircean images, diagrams, and metaphors show up in language and elsewhere. Linguistic examples of Peircean images include countless onomatopoeic words, such as the ones representing the cry of a rooster: “*kukuruku* in Basque, *kukeleku* in Dutch, or *kokoriko* in Turkish” (Nöth [1990] 1995: 125). Outside language, Peircean images are instantiated by run-of-the-mill pictorial signs, such as the stick figure of a man stenciled on a bathroom door to indicate the bathroom's gender use. According to Farias and Queiroz (2006: 9–10), the stenciled man qualifies as a Peircean image to the degree that it shares its overall shape with that of a male human. Farias and Queiroz (2006: 9–10) further explain that the stenciled man also qualifies as a non-verbal Peircean diagram insofar as its own structural relations reproduce the relations among the major parts of the human body. In language, Peircean diagrams are instantiated by expressions conforming to *ordo naturalis*, whereby the sequence in which words are uttered mirrors the sequence of the events they describe, as in *Veni, vidi, vici* (Nöth [1990] 1995: 123). As to metaphors, their linguistic and non-linguistic manifestations have been explored for many years by cognitive linguists. Nearly three decades ago Lakoff (1993: 241) pointed out that metaphors are expressed linguistically and non-linguistically “in obvious imaginative products such as cartoons, literary works, dreams, visions, and myths” as well as “in physical symptoms, social institutions, social practices, laws, and even foreign policy and forms of discourse and of history.” Since then, multiple studies have been published on metaphoricity of gestures (Cienki 1998; Sweetser 1998), advertisements (Caballero 2009; Forceville 1996; Urios-Aparisi 2009; Yu 2009), cartoons (El Rafea 2009; Górska 2020; Teng 2009; Schilperoord and Maes 2009), comics (Abbott and Forceville 2011; Eerden 2009; Forceville 2005; Forceville, Veale and Feyaerts 2010; Potsch and Williams 2012; Shinohara and Matsunaka 2009; Szawerna 2017), films (Eggertson and Forceville 2009; Forceville [2013] 2015; Popa [2013] 2015; Rohdin 2009), and even medieval textiles (Díaz Vera [2013] 2015; Díaz Vera and Manrique-Antón 2015).

Peirce's index belongs to secondness because it “represents an object by virtue of its connection with it,” irrespective of “whether the connection is natural, or artificial, or merely mental” (CP § 8.368). Typically, an index signifies its object by virtue of being “affected by the Object” (Peirce CP § 2.248) in a physical way, so that “they make an organic pair, but the interpreting mind has nothing to do with this connection, except remarking it, after it is established” (Peirce CP § 2.299),<sup>12</sup> but “direct physical

12 Peirce's examples include a rap on the door as an index of the one who made the sound (CP § 2.92), a rise of the mercury in a thermometer as an index of the increase in temperature which caused the mercury to expand (CP § 5.473), and the orientation of a weathercock as an index of the direction of the wind, which oriented the weathercock parallel to that direction (CP § 2.286).

connection” (Peirce CP § 1.372) is not a necessary attribute of indices.<sup>13</sup> While many indices involve “a cause-effect relation” (Johansen and Larsen [1994] 2002: 32) between the object and the sign, this kind of relation is not a necessary attribute of indices either.<sup>14</sup> There is, however, one attribute that all Peircean indices seem to share. According to Nöth, an index invariably “focuses the interpreter’s attention on the object” ([1990] 1995: 114). In Peirce’s own words, the index “only says ‘There!’ It takes hold of our eyes, as it were, and forcibly directs them to a particular object, and there it stops” (CP § 3.361). Since an index is necessarily relational, it presupposes “the existence of the object as an individual entity” (Nöth [1990] 1995: 114), but according to Jappy (2013: 90), this necessarily existent object need not constitute a feature of the real world — it may well exist in an entirely fictive world (of a novel, film, comic, *etc.*).

Peirce’s symbol belongs to thirdness because unlike hypoicons, which signify by their own qualities, and indices, which signify by virtue of an inherent connection between the sign and its object, symbols, which are signs neither similar to their objects nor inherently connected to them, must rely on the interpretant for semiosis to take effect. According to Peirce, “[a] chalk mark is like a line though nobody uses it as a sign; a weather cock turns with the wind, whether anybody notices it or not. But the word ‘man’ has no particular relation to men unless it be recognized as being so related” (unpublished manuscript no. L75, 1902: 149). In Peirce’s view, then, the arbitrary association of a symbolic sign with its object “can only be realized by the aid of its Interpretant” (CP § 2.92), whereby this association becomes conventionalized to the degree that a symbol signifies its object “by virtue of a law” (CP § 2.249). Like hypoicons and indices, symbols are linguistic as well as non-linguistic. The former are exemplified by “[c]ommon nouns, verbs, adjectives, and adverbs” (Jappy 2013: 91) and the latter by traffic signs, logos, emblems, trademarks, *etc.*, to the extent that they are arbitrary and conventional signs.

#### 4. A Peircean characterization of the pictorial and the verbal

Above, pictures were characterized in their everyday, non-expert sense, as visually mimetic drawings, paintings, and photographs. From the perspective of Peirce’s sign theory, visual mimeticity of drawings, paintings, and photographs results from the fact that they function as iconic signs, and in particular as both imagetive and diagrammatic hypoicons of their objects (*cf.* Farias and Queiroz 2006: 9–10), with the proviso that the similarities between these hypoicons and their objects are of the visual kind. More specifically, drawings, paintings, and photographs qualify as imagetive hypoicons to the extent that they share visually perceptible “simple qualities” (Peirce CP § 2.277), such as shape, size, and color (*cf.* Elleström 2013: 105–106), with their objects, and they qualify as diagrammatic hypoicons to the extent that their own visually perceptible structural relations reproduce the visually perceptible structural relations of their objects. Because Peircean signs are mental entities (Nöth [1990] 1995: 42), the similarities between visually mimetic drawings, paintings, and photographs and their objects are established between mental representations of these pictures and mental representations of the things they show. For this reason, the similarities may be established at varying levels of specificity, depending on the viewer’s prior knowledge, which in turn may, but need not, include direct experience of the thing signified by a drawing, painting,

13 For example, there is no direct physical connection between the sign and its object when a spontaneous cry is produced at the sight of danger (Peirce CP §§ 2.256, 2.92).

14 For example, a proper name, characterized by Peirce as “a genuine Index” (CP § 2.329), is in no way caused by its object.

or photograph. It is quite often the case that the mental representation of a pictorial sign's object is a composite of previously seen representations — a kind of conceptual blend in the sense of Fauconnier and Turner (2002). For instance, most contemporary viewers have never met Albert Einstein, Mahatma Gandhi, or Adolf Hitler, and yet they will easily recognize those individuals in a newly encountered picture on the basis of previously seen representations (most notably, photographs and documentary footage). In contrast, a viewer who has no previous knowledge of the thing signified by a picture is not likely to uniquely identify this thing because the similarities between the sign and its object will be established at a relatively low level of specificity. For example, somebody who has no previous knowledge of, say, Mark Turner, is likely to interpret a current photograph of this cognitive scientist as a picture of a middle-aged Caucasian male. Similarly, someone who has no previous knowledge of the Willis Tower in Chicago is likely to interpret a drawing, painting, or photograph of this office building as a picture of a skyscraper.

There is, however, an important difference in how viewers interpret photographs and drawings or paintings of things they have no previous knowledge of, and it is to do with their indexicality. In particular, photographs are characterized by a “direct physical connection” (Peirce CP § 1.372) between the sign and its object, whereby the representation of the thing photographed is created when the light reflected from or emitted by the thing photographed passes through the camera's lens and physically interacts with a light-sensitive surface (a photographic plate or an electronic image sensor) inside the camera during exposure. The awareness of the direct physical cause-effect relation between the thing photographed and the photograph gives viewers “the undeniable certainty that the thing photographed really existed” (Lombardo [1989] 2010: 139). The certainty that the object of a pictorial sign really existed when the picture was created is obviously absent when the picture is drawn or painted, but due to the mental nature of the Peircean sign, drawings and paintings of imaginary things still qualify as both *imagetic* and *diagrammatic hypoicons* of their objects due to the referential openness of *hypoicons* (Peirce CP § 4.531). A drawing or painting of an imaginary thing can be recognized by the viewer on the basis of a conceptual blend of the thing's previously encountered representations — not unlike the above-mentioned photographs of Einstein, Gandhi, or Hitler. For example, a griffin may be uniquely recognized in a drawing or painting with reference to a mental composite of previously seen drawings, paintings, sculptures, embroideries, *etc.*, of these mythical creatures. In the absence of this kind of mental composite, the viewer is likely to interpret the drawn or painted griffin by recruiting mental representations of lions and eagles because the griffin combines elements of their appearance that are mentally available to the viewer from prior knowledge. In this way, the viewer will interpret the griffin as a visual hybrid of lion and eagle — another kind of conceptual blend in the sense of Fauconnier and Turner (2002).

But, as indicated above, pictures are not on a par in terms of their visual mimeticity, whose degree seems to be adversely affected by the degree to which a drawing, painting, or photograph is symbolic, and the variable symbolicity of pictures depends on how arbitrary the relation is between the mental representation of a sign's form and the mental representation of the thing signified. Some key characteristics of pictures that make this relation arbitrary are shared by all drawings, paintings, and photographs, which, due to their materiality (pencil on paper, oil on canvas, acrylic on glass, photographic paper, *etc.*), constitute largely two-dimensional (as opposed to sculptures) and static (as opposed to films) visual representations of three-dimensional objects that may or may not be in motion. Other characteristics of pictures that make the relation between the mental representation of a picture's form and the mental representation of the thing signified arbitrary emerge as a result of various conventions adopted more or less freely by

particular creators with regard to the materials they use, the techniques they apply, and the styles they follow. On the whole, the simpler the picture's form in terms of detail, color, perspectival depth cues, *etc.*, the more arbitrary the relation, and in this respect pictures run a whole gamut from the highly visually mimetic utilitarian photographs, through realistic paintings and variously stylized art photographs, to the kinds of idealized line drawings found in cartoons and comics and the highly schematic stick figures stenciled on bathroom doors. The conventional simplification of pictures that makes them symbolic additionally confers the status of indices on the "inherently incomplete" (Bordwell [1985] 1986: 101) drawings, paintings, and photographs insofar as they are automatically taken to refer to their much more specific semiotic objects. This is true not only of the necessarily existent objects of photographs, but also the existentially vague objects of drawings and paintings. For example, a viewer who comes across a small, fuzzy, black and white picture of Mark Turner's face will take it to represent an entire real-world human being — one that is maximally specific in every way. Similarly, a reader who sees drawings of, say, Charlie Brown in the successive panels<sup>15</sup> of Charles Schulz's comic strip will automatically take the drawings to refer to a fully formed 'walking and talking' individual inhabiting the imaginary world of the strip (Szawerna 2017: 197).

This semiotic characterization of pictures as predominantly iconic visual signs is far from complete, but it provides a sufficient springboard for further discussion of the scope of the pictorial outlined tacitly in multimodal cognitive linguistic investigations of the cognitive underpinnings of static planar ensembles (printed advertisements, cartoons, comics, *etc.*).

Unlike the elusive scope of the pictorial, the scope of the verbal seems somewhat easier to circumscribe. In simplest terms, the scope of the verbal includes meanings encoded linguistically, that is, with the use of the lexical and grammatical resources of a natural language. Importantly, this formulation does not specify the materiality of verbal signs because the latter is quite variable — in large part due to the symbolicity of verbal signs, whereby the association between their form and the things they represent is to a great extent arbitrary, albeit with some notable exceptions, such as the previously referenced onomatopoeic words signifying the cry of a rooster: *kukuruku*, *kukeleku*, and *kokoriko* (Nöth [1990] 1995: 125). While verbal signs primarily materialize as sounds made in the vocal tract, they may also take other forms: ink marks on paper (as is the case with written or printed letters, syllabograms, or logograms), manual articulations (extensively used in various signed languages), and a few others, such as the raised dots of Braille. Importantly, none of these materialities is uniquely associated with the use of language. Humans use their vocal tracts to produce non-verbal sounds every time they hum or whistle a tune, they draw pictures by placing ink marks on paper with a pen or a brush, they often use gestures independently of language, *etc.* Insofar as the linguistic component of a verbal sign can be divorced from the sign's materiality, there are grounds for considering verbal signs as being multimodal, or at least bimodal (Szawerna 2020: 191).

---

15 Recognized by many (e.g. Eisner [1985] 2008: 26; Groensteen [1999] 2007: 4–5; Wartenberg 2012: 95) as the basic unit of expression in comics, the panel is an individual vignette, aptly characterized by Cohn as "a demarcated frame of image content put into discrete sequences" (2007: 35).

## 5. The pictorial and the verbal in multimodal cognitive linguistic explorations of static planar ensembles

In the strand of multimodal cognitive linguistics that investigates the cognitive underpinnings of static planar ensembles, the words *verbal* and *pictorial* are often combined into *verbo-pictorial*, which is meant to identify the modalities<sup>16</sup> characterizing the ensembles included in the scope of the analysis. The term *verbo-pictorial* has its origin in Charles Forceville's *Pictorial Metaphor in Advertising* (1996), a seminal book that paved the way for what later became known as multimodal cognitive linguistics. In that book, the term *verbo-pictorial* was used to refer to metaphorical ensembles "in which one of the terms is rendered pictorially and one is rendered verbally" (Forceville 1996: 148). In relation to what was said above about the scope of the pictorial and the scope of the verbal, the term *verbo-pictorial* rather felicitously characterizes virtually the entire corpus of modal ensembles analyzed by Forceville (1996),<sup>17</sup> in which verbal signs come in the form of printed alphabetic writing and pictorial signs come in the form of decidedly mimetic visual representations — in most cases photographic images.<sup>18</sup> Sometimes these images show non-existent (if not impossible) things, such as, for example, an automobile with life buoys in place of tires (Figure 6.9 in Forceville 1996: 123) or the human spine topped with a parking meter instead of a skull (Figure 6.16 in Forceville 1996: 135). These images, optically or digitally composited from actual photographs, are not unlike the previously discussed pictorial representations of griffins except that the non-existent things they show cannot be uniquely recognized by viewers with reference to mental composites of their previously seen representations for the reason that the latter are simply not available. Instead, composite images showing novel non-existing things are likely interpreted with reference to mental representations of the things that are visually integrated in these images (automobiles, life buoys, spines, parking meters, etc.) because these mental representations can be effortlessly recruited by viewers from prior knowledge for purposes of interpretation.

Importantly, despite the fact that Forceville's (1996) corpus of static planar ensembles includes non-verbal signs representing existent as well as non-existent objects, all these signs qualify as pictures because they exhibit the unique combination of imagetic iconicity, diagrammatic iconicity, indexicality, and symbolicity ascribed previously to uncontroversially pictorial representations. In his later publications, however, Forceville (2005, 2011) used the term *pictorial* with reference to visual signs that are clearly non-mimetic in that they do not look like the thing they represent. It seems that Forceville's broadening of the term *pictorial* resulted at least in part from his adoption of the term *pictorial rune*, attributed by Forceville (2005: 73) to the perception psychologist Kennedy (1982), who first used it with reference to non-mimetic signs found in cartoons and comics, such as spirals drawn in place of a character's eyes to represent the character's anxiety, wavy lines drawn above the picture of a smelly object to represent the

16 In multimodality studies, the thorny concept of modality (a.k.a. mode) has so far been theorized the most aptly by Bateman, Wildfeuer, and Hiippala (2017: 113–117), who suggested that modalities be characterized as combinations of values along three parameters: materiality, semiotics, and discoursivity (cf. Szawerna 2020: 192–194).

17 Minor exceptions include a couple of logotypes, which contain visual elements that are neither clearly verbal nor clearly pictorial (such as the Lassale logotype, reproduced by Forceville 1996: 139 in Figure 6.18, or the BMW logotype, reproduced by Forceville 1996: 151 in Figure 6.21).

18 Except for several hand-drawn pictorial signs, such as the ones reproduced in Figure 6.5 (Forceville 1996: 117) and Figure 6.11 (Forceville 1996: 128).

object's unpleasant odor, straight lines signifying the sound coming from a pictorially represented source or radial lines drawn next to the picture of an injured body part to represent the pain associated with the injury (Kennedy 1982: 600). On the whole, Kennedy's (1982) pictorial runes are static planar signs that represent non-visual concepts (anxiety, odor, sound, pain, *etc.*) in visual terms. It seems that Kennedy referred to such signs as *pictorial* despite their lack of visual mimeticity for the simple reason that they "are in common use today in pictures" (1982: 603). In other words, Kennedy (1982: 603) seems to have classified his so-called runes as pictorial on the basis of their co-occurrence with other kinds of static planar signs in what he referred to as pictures.

While co-occurrence may no doubt constitute an important factor in categorization (as evidenced by such linguistic categories as *body parts, desert animals, furniture, garbage, etc.*), it inevitably engenders problems if taken as the sole factor unifying the semiotically related, yet very different kinds of representations found in cartoons and comics. While Kennedy's (1982: 603) assertion that pictorial runes are commonly found in pictures is relatively uncontroversial,<sup>19</sup> it needs to be emphasized that the kinds of static planar signs they regularly co-occur with include not only unequivocally mimetic drawings of people, places, various everyday objects, *etc.*, but also various unequivocally non-mimetic visual representations, such as elements of manuscript writing (for example, the letters of the Latin alphabet), elements of numeral systems (for example, Arabic or Roman numerals), and elements of musical notation (for example, symbols used in modern staff notation). If regular co-occurrence were taken as the sole factor unifying these semiotically diverse static planar signs, they should all be referred to as *pictorial*, but this would certainly be both counterintuitive and counterproductive as it would ignore the irreconcilable differences among these signs concerning the manner in which they become meaningful. This is especially evident with regard to imagetive iconicity and diagrammatic iconicity, which, as previously indicated, are jointly conducive to visual mimeticity achieved by representational drawings, paintings, and photographs, but are not really exhibited by predominantly symbolic signs, such as written or printed language, numbers written in Arabic or Roman numerals, music written in staff notation, *etc.*

But perhaps the modifier *pictorial* in *pictorial runes* could be understood in a different way: as 'being functionally dependent on visually mimetic representations of people, animals, objects, *etc.*, to which pictorial runes are spatially proximal, but from which they are visually separable.' This is what Cohn meant when referring to such static planar signs as "bound" (2007: 48) and "parasitic" (2007: 49). The functional dependence of pictorial runes on the visual signs they co-occur with was also recognized by Forceville, who characterized pictorial runes as "very simple, abstract-looking flourishes, which would have little meaning if we were to encounter them in isolated, decontextualized form" (2011: 876). The problem with this understanding of *pictorial* is that, on the one hand, pictorial runes do not have to depend for their interpretation on visually mimetic signs,<sup>20</sup> and, on the other hand, pictorial runes do not have to be visually separable from the signs they do depend on for their interpretation. Both of these reservations

19 With the caveat that pictorial runes are certainly *not* in common use in pictures of just about any kind. For example, they do not seem to be found at all in paintings or photographs, which are commonly referred to as *pictures*. In fact, Kennedy's (1982) pictorial runes are only used with any regularity in cartoons and comics. While Kennedy (1982) did not make it explicit, the examples he discussed do indicate that what he meant by *pictures* were hand-made line drawings of the kind that are regularly found in cartoons and comics.

20 Forceville (2011: 876–877) himself points out that "[a] rune's meaning reveals itself in combination with one or more of the following: (1) other runes; (2) pictograms; (3) balloonic features; (4) iconic information, including facial expressions and hand/arm postures [...]; (5) panel form, lay-out and orientation; and (6) verbal text."

pertain to Forceville's so-called "bold face" (2005: 77), whereby oversized, boldface letters making up a written representation of an angry character's utterance are used to signify the excessive loudness of the character's speech. This pictorial rune, created by an increase in the size and thickness of a letter, is functionally dependent on a visually non-mimetic sign, i.e. the letter, with which it is visually integrated. Both reservations also pertain to Forceville's so-called "jagged lines" (2005: 77), whereby the indexical projection of a speech balloon, referred to by comics scholars as the *tail* (e.g. Khordoc 2001: 159; Cohn 2007: 48, 2013: 35; Forceville, Veale and Feyaerts 2010: 56), is drawn in a jagged line to indicate angry speech. This pictorial rune, created by a modification of the shape of the visually non-mimetic speech balloon tail, is neither bound to a visually mimetic sign, nor is it visually separated off from its functional "root" (Cohn 2007: 48). On the whole, then, pictorial runes do not qualify as being pictorial by virtue of being dependent on visually mimetic signs from which they are visually separable; at least some of them are visually integrated with visually non-mimetic signs, such as letters and balloon tails.

There are, however, other problems with the scope of the pictorial in Forceville's (2005) article on visual representations of the concept of anger in Goscinny and Uderzo's *Asterix and the Roman Agent* ([1970] 1972). In that article, Forceville distinguished a number of static planar signs characterized as "[p]ictorial signals of anger" (2005: 75). While all of them constitute unequivocally non-mimetic visual representations of anger in the sense that they represent a non-visual concept in visual terms, they do not make up a semiotically homogenous set. The static planar signs of anger grouped by Forceville (2005: 77) in Category I ("bulging eyes," "tightly closed eyes," "wide mouth," "tightly closed mouth," "pink/red face," "arm/hand position," and "shaking") may not represent anger itself mimetically, but they are no doubt visually mimetic representations of the bodily symptoms of anger recognizable on the basis of everyday experience. More specifically, the semiosis of the static planar signs of anger in Forceville's (2005: 77) Category I involves two consecutive cycles. In the first cycle certain combinations of ink marks on paper are interpreted as visually mimetic representations of bodily configurations,<sup>21</sup> and in the second cycle these bodily configurations are recognized as typical symptoms of anger. Due to its visual mimeticity, the first cycle, but not the second, can be characterized as pictorial. In contrast, the static planar signs of anger grouped by Forceville (2005: 77) in Category II ("spirals," "ex-mouth," "smoke," "bold face," and "jagged line"), explicitly labelled "pictorial runes" (Forceville 2005: 77), differ from the ones in Category I in that they do not qualify as visually mimetic representations of typical bodily symptoms of anger. The problem is that they do not make up a semiotically homogenous set either. As I pointed out elsewhere (Szawerna 2017: 15), some of the static planar signs of anger in Forceville's (2005: 77) Category II — in particular, the so-called "smoke," a sign which comprises multiple visually mimetic representations of smoke puffs placed above the picture of an angry character's head, and the previously referenced "bold face," which consists of oversized, boldface letters making up a written representation of an angry character's loud speech — are aptly characterized as each other's semiotic converses. The converse relationship between the two pictorial runes consists in the fact that "smoke" qualifies as a *realistic representation of an unrealistic index*, insofar as this visual sign mimetically depicts puffs of smoke in one semiotic cycle, and in the next the puffs are interpreted as the unrealistic effect of anger conceived of as fire (by virtue of the *ANGER IS FIRE* metaphor, invoked by Forceville 2005: 82), whereas "bold face" qualifies as an *unrealistic representation of a realistic index*, insofar as in one semiotic cycle it is an unrealistic, metaphorical sign of loud speech (by

21 With the caveat that mimeticity is a matter of degree. The signs in Forceville's (2005: 77) Category I are no doubt visually mimetic, albeit considerably simplified and exaggerated, representations of human bodily configurations.

virtue of the MORE OF FORM IS MORE OF CONTENT metaphor, invoked by Forceville 2005: 82), and in the next the loud speech is recognized as a realistic symptom of anger known from everyday life.

On the whole, Forceville's (2005) corpus of static planar signs of anger, referred to as "[p]ictorial signals of anger" (Forceville 2005: 75), comprises three kinds of representations distinguishable on the basis of their visual mimeticity. These three kinds are alike in that they are all visually non-mimetic representations of anger, and they represent this emotion indexically. Where they differ is in the way this indexical connection is established. The first kind is exemplified by the static planar signs of anger included in Forceville's (2005) Category I, that is, visually mimetic representations of the bodily symptoms of anger recognizable on the basis of everyday experience. These signs constitute *realistic representations of realistic indices*. The second kind is exemplified by signs like "smoke," characterized above as *realistic representations of unrealistic indices*, which employ visually mimetic signs as representations of unrealistic (specifically, metaphorical) symptoms of anger. The third kind is exemplified by signs like "bold face," characterized above as *unrealistic representations of realistic indices*, which employ visually non-mimetic (specifically, metaphorical) signs as representations of realistic (in this case, aural) symptoms of anger. To conclude: Forceville's (2005) corpus of static planar signs of anger, referred to as "[p]ictorial signals of anger" (Forceville 2005: 75), is made up of representations that can only be referred to as *pictorial* in the sense of Kennedy (1982), that is, as visual signs co-occurring with other visual signs in the kind of pictures found in cartoons and comics. Forceville's (2005) static planar signs of anger are definitely not pictorial in the sense that they are visually non-mimetic, indexical representations of anger, with the caveat that two out of the three kinds distinguished above establish their indexical connection with anger via visually mimetic signs of the (realistic or unrealistic) symptoms of this emotion.

The use of the term *pictorial* with reference to static planar signs that are visually non-mimetic has continued to this day in the strand of multimodal cognitive linguistics that investigates the cognitive underpinnings of static planar ensembles. Koller (2009: 47) characterized the visually non-mimetic "abstract design elements" of the corporate logos included in the scope of her analysis as pictorial signs. A couple of years later, Forceville (2011) identified a number of pictorial runes in Hergé's *bande dessinée* album *Tintin and the Picaros* ([1976] 1976). Visually, these pictorial runes are completely non-mimetic, with the exception of "droplets" (Forceville 2011: 877), which look like "little units of liquid, for instance in the form of water, sweat, tears, and spit" (Forceville 2011: 879), and, less obviously, "the twirl" (Forceville 2011: 877), which looks like "a miniature tornado, or a small spring" (Forceville 2011: 877). In terms of their semiotics, these two pictorial runes resemble the previously discussed pictorial rune dubbed "smoke" (Forceville 2005) in that they constitute (more or less) *realistic representations of unrealistic indices*. More recently, the term *pictorial* was used with reference to non-mimetic visual signs by Górska (2020), who referred to the cartoons created by the artist Kapusta as "verbo-pictorial aphorisms" Górska (2020: 9). As I pointed out elsewhere (Szawerna 2020), the characterization of the non-verbal static planar signs in Kapusta's artwork is problematic because these signs are semiotically quite diverse. While some of them are visually mimetic,<sup>22</sup> others are definitely not. In particular, Kapusta's cartoons analyzed by Górska (2020) contain various arrows and lines (short, long, straight, bent, squiggly, *etc.*) which are visually non-

22 Kapusta's visually mimetic signs discussed by Górska (2020) include drawings of his pawn-like protagonist (Example 1.1, Górska 2020: 13; Example 1.2, Górska 2020: 14; *etc.*), blazing flames (e.g. Example 2.6, Górska 2020: 35), a signpost (e.g. Example 3.1, Górska 2020: 49), a stretcher bar (e.g. Example 3.12, Górska 2020: 65), and a tolling bell (e.g. Example 3.13, Górska 2020: 66).

mimetic signs, meant to function as Peircean diagrammatic hypoicons, metaphorical hypoicons, indices of one kind or another, symbols, or combinations thereof. For instance, Kapusta's cartoon reproduced as Example 3.2 (Górska 2020: 51) features his pawn-like protagonist situated on a line from which a number of other lines fan out. If this visual representation were interpreted in its entirety as a visually mimetic sign, the cartoon would likely be taken to show the protagonist flying on a broom, *à la* Harry Potter, but the Polish verbal anchor (in the sense of Barthes [1964] 1986: 29) "Chociaż mamy nieskończoną ilość dróg do wyboru i tak pójdziemy tylko jedną"<sup>23</sup> effectively blocks this interpretation by suggesting that the lines be understood as components of a diagrammatic as well as symbolic representation of a metaphorical path that forks into numerous other metaphorical paths.

In the strand of multimodal cognitive linguistics investigating the cognitive underpinnings of static planar ensembles, the term *pictorial* is sometimes replaced with *visual*, so that the pair *visual* and *verbal* is used instead of *pictorial* and *verbal* to identify the modalities characterizing the ensembles included in the scope of the analysis (Caballero 2009; El Rafaie 2009; Koller 2009; Yus 2009). Unfortunately, this terminological substitution does not rectify the problems discussed above; instead, it adds to the confusion for the reason that all verbal representations in static planar ensembles (printed advertisements, cartoons, comics, *etc.*) also qualify as visual representations. Consequently, instead of referring to two distinct modalities that characterize static planar signs belonging to complementary sets — as was no doubt intended by the above-mentioned multimodal cognitive linguists — the pair *visual* and *verbal* in fact characterizes two sets of static planar signs in an asymmetrical relation whereby the latter (verbal) set is in its entirety included in the former (visual) set. As a result, the term *visual*, which pertains to all signs making up the analyzed static planar ensembles (printed advertisements, cartoons, comics, *etc.*) even though the signs are tremendously varied in terms of their semiotic make-up, ceases to be analytically useful. At this juncture, it is worth invoking the example of written representations of onomatopoeias, such as *skroww*, *baroom*, *kerashh*, *choom*, or *tschrak*,<sup>24</sup> which are easy enough to come across in action comics published anywhere in the world. As conclusively demonstrated by Kowalewski (2015: 31), these written representations are anything but verbal. They do not qualify as words in written form because they do not correspond to conventional linguistic units stored in the lexicon, they cannot be assigned to particular lexical categories, and they are not subject to the rules of the grammar (Kowalewski 2015: 31). At the same time, these written representations of onomatopoeias are clearly non-pictorial insofar as they are visual representations of non-visual entities (sounds), and for this reason they cannot be regarded as visually mimetic signs. Referring to written representations of onomatopoeias in comics as visual rather than pictorial signs seems counterproductive because despite their semiotic diversity, all signs are visual in the static planar medium of comics.

23 This has been rendered in English as "Even though we have an infinite number of paths to choose from, we will follow only one" (Górska 2020: 51).

24 These one-off written forms have been reproduced from a single comic-book anthology, titled *X-Men: Supernovas* (Carey et al. 2007).

## 6. An alternative account of the meaning-making potential of static planar ensembles

The issue of the semiotic resources of the comics medium has been taken up on many occasions in the field of comics studies, and in particular in the so-called “new comics scholarship” (Heer and Worcester 2009: xiv), which took off in the mid-1980s, following the publication of seminal explorations of the meaning-making potential of comics written by practicing cartoonists like Eisner (1985), McCloud ([1993] 1994), and Harvey (1994, 1996), who paid “special attention to the formal aspects of comics” (Heer and Worcester 2009: xiv). In a previous study (Szawerna 2012), I analyzed 12 expert definitions of comics proposed by prominent comics scholars, most of whom represent the new comics scholarship: Waugh (1947), Kunzle (1973), Eisner ([1985] 2008), Barker (1989), Inge (1990), McCloud ([1993] 1994), Sabin (1993), Harvey (1994, 1996), Groensteen ([1999] 2007), Carrier (2000), Saraceni (2003), and Duncan and Smith (2009). That analysis enabled me to compile a list of 22 attributes characterizing comics understood as, on the one hand, a (predominantly narrative) medium and, on the other hand, a corpus of publications instantiating the conventions prescribed by the medium (Szawerna 2012: 65). If the attributes are weighed in terms of how many times they appear in the analyzed definitions of comics, it turns out that the second weightiest attribute on that list, right after the rather uncontroversial attribute which says that comics consist of panels, is the one whereby comics combine pictorial and verbal signs, understood respectively as hand-drawn pictures and hand-written words. This attribute is part and parcel of 10 out of the 12 analyzed definitions: the ones proposed by Waugh (1947: 13-14), Eisner ([1985] 2008: 2), Inge (1990: xi), McCloud ([1993] 1994: 9), Sabin (1993: 5), Harvey (1994: 9, 1996: 3), Groensteen ([1999] 2007: 7), Carrier (2000: 7), Saraceni (2003: 35), and Duncan and Smith (2009: 4). The view of comics as combinations of hand-drawn pictures and hand-written words informs many other scholarly treatments of comics (for example, Bongco 2000; di Liddo 2009; Miodrag 2013; Phoenix 2020; Postema 2013; Wartenberg 2012), and multiple comics scholars have put considerable effort into elucidating the relations between pictorial and verbal signs in the comics medium.<sup>25</sup>

In view of the above, it seems fair to say that there is a widespread consensus among comics scholars to problematize comics as combinations of hand-drawn pictures and hand-written words. Arguably, this is the case because there are good reasons for doing so. After all, some comics do indeed comprise nothing but pictorial and linguistic images.<sup>26</sup> The problem, however, is that the view of comics as ensembles of pictorial and verbal signs does not seem to do the medium of comics full justice as for a long time now the medium has had at its disposal visual signs that are, strictly speaking, neither pictorial nor verbal, such as the previously discussed written representations of onomatopoeias or the visually non-mimetic pictorial runes discussed by Forceville (2005, 2011). But comics theorists seem to be well aware of this fact, and some of them have recognized it in their explorations of the medium. In his definition of comics, McCloud spoke of “juxtaposed pictorial and other images in deliberate sequence” ([1993] 1994: 9), thereby allowing for the occurrence of not only verbal signs in comics, but also of signs that are neither pictorial nor verbal. Additionally, McCloud ([1993] 1994: 49) and Saraceni (2003: 18–27) postulated the existence of a sign continuum situated between the two extremes constituted by uncontroversially

25 For a comprehensive overview, consult Chapter 4 in Miodrag (2013: 83–107).

26 Strictly speaking, however, comics can do quite well without written words: Groensteen ([1999] 2007: 14–15) shows that wordless comics are not particularly difficult to come across.

pictorial and uncontroversially verbal signs, and Szyłak (2000: 21) acknowledged the occurrence of signs that are neither pictorial nor verbal in static planar ensembles by discussing an example of visually non-mimetic visual representations in a cartoon by Saul Steinberg (Figure 15 in Szyłak 2000: 22). More recently, Cohn (2007: 47–50, 2013: 34–48) discussed a number of visually non-mimetic signs that are neither pictorial nor verbal in the framework of his highly original theory of visual languages. This group of static planar signs includes, among others, such immediately recognizable elements of comics form as *speech balloons*, which represent externalized speech of characters; *thought bubbles*, which represent their inner speech; *motion/speed lines*, which show the trajectory traversed by a moving object and indicate the object's speed; and *impact stars*, which pinpoint the exact location of forceful contact between objects. The lack of visual mimeticity on the part of these signs follows from the fact that they “represent unseen aspects of the visual representation” (Cohn 2007: 48) and “are ‘invisible’ in any ‘realistic’ visual sense” (Cohn 2007: 48).

It seems, however, that the most comprehensive model of the meaning-making potential of the comics medium — one that allows the analyst to make a principled distinction between any two static planar signs found in comics — was formulated even more recently by Kowalewski (2015). This comics scholar proposed to characterize the process whereby static planar signs become meaningful in terms of Peirce's sign theory as relying on a number of semiotic *moments*, which consist in the establishment of indexical, iconic, and symbolic relations between a sign and its objects in consecutive semiotic cycles. In Szawerna (2017: 17–19), I described the semiosis of written representations of onomatopoeias on the example of the written form *blam*, which is sometimes used in comics to represent the sound of a gunshot (*cf.* Figure 55 in Szawerna 2017: 333), in terms of Kowalewski's (2015) model, and the following two paragraphs summarize that description.

The semiosis of the written form *blam*, conceived of as a representation of the sound of a gunshot, presupposes a number of semiotic cycles involving a complex interaction of Peirce's modes of signification. While some of these cycles are obligatory, others are optional. In one obligatory cycle, the sound of a gunshot is transformed into the so-called “proxy sound” (Kowalewski 2015: 32): a representation made up of phonological units belonging to the linguistic repertoire of the comic's creator and its readers, such as /blæm/. In this semiotic cycle, imagetic iconicity constitutes the dominant mode of signification due to the similarity observable between the proxy sound and the gunshot sound, which depends on the aural properties (pitch, duration, loudness, *etc.*) shared by the sign and the thing signified in this semiotic cycle. As pointed out by Kowalewski (2015: 33), it is in this cycle that the sign acquires its intuitively recognizable onomatopoeic quality, but in addition to its imagetic iconicity, the proxy sound /blæm/ constitutes a symbol because it consists of highly conventionalized linguistic units. In another obligatory cycle, the proxy sound is transformed into its orthographic counterpart *blam*. With relation to the phonological units of the proxy sound /blæm/, the letters of the orthographic representation function as symbols because their form is in no way motivated by the phonological characteristics of the signified units. Additionally, the individual letters making up *blam* qualify as metaphorical hypoicons of the corresponding phonemes in the proxy sound /blæm/ insofar as they visually objectify these aural units as spatially discrete ink marks on paper. In its entirety, the orthographic form *blam* exhibits both metaphorical and diagrammatic iconicity with relation to the signified proxy sound. On the one hand, *blam* qualifies as a metaphorical hypoicon of the proxy sound /blæm/ because it visually objectifies this higher-order aural event in a spatially bounded static planar form. On the other hand, *blam* qualifies as

a diagrammatic hypoicon of the signified proxy sound due to the one-to-one correspondence between the spatial left-to-right arrangement of the letters in *blam* and the temporal progression of the phonological units making up /blæm/. Additionally, the left-to-right vectorization of *blam*, determined by the reading protocol prescribed for texts written in the Latin alphabet, confers symbolic status onto this orthographic representation. In yet another obligatory semiotic cycle, the orthographic form *blam* exhibits two kinds of indexicality. On the one hand, *blam* qualifies as an index of the signified gunshot sound because the presence of this orthographic form inside a panel indicates the occurrence of the corresponding sound event within the situation depicted by the panel. On the other hand, *blam* qualifies as an index of the gun that produced the visually objectified sound because the presence of this orthographic form inside a panel presupposes the existence of the source of the gunshot sound within the narrative situation represented by the panel, irrespective of whether or not a drawing of the gun is placed inside the panel's frame.

In non-obligatory cycles, the semiotics of *blam* may be further developed by variations in the typography of the letters making up this orthographic representation. For example, increased size (height as well as thickness) of the letters making up *blam* may be used to signify the deafening loudness of the visually objectified gunshot sound, in accordance with a visual convention of comics motivated by conceptual metaphors capturing quantity in terms of size and vertical elevation (cf. Grady 1997: 285). Also, the shape of the letters making up *blam* may be modified in such a way as to indicate certain aural properties of the gunshot sound it represents. In comics, various environmental sounds that are unpleasant to the ear are sometimes signified by onomatopoeic representations written in sharp-edged, angular letters (McCloud 2006: 147). If *blam* were written in this way, the sharp-edged angularity of the letters would function as a metaphorical hypoicon of the perceptual unpleasantness of the visually objectified gunshot sound, in accordance with a synaesthetic metaphor whereby an aural property of a sound is metaphorized as a tactile property of a physical object (cf. Shen 2008: 302–305).

## 7. Conclusion

In this article, an attempt has been made to examine a number of contributions to the strand of multimodal cognitive linguistics investigating the cognitive underpinnings of various static planar ensembles (printed advertisements, cartoons, comics, etc.) with regard to what these contributions tacitly assume to be included in the scope of two concepts that are pivotal to any discussion of such ensembles: the pictorial and the verbal. As the contributions included in the analytical scope of this article were not based on any particular theory of pictorial representation, the everyday notion of pictoriality, whereby pictures are understood as visually mimetic paintings, drawings, and photographs, was adopted as the standard relative to which the contributions have been compared, contrasted, and evaluated. In order to facilitate the discussion of the scope of the pictorial in multimodal cognitive explorations of static planar ensembles, pictures have been characterized in terms of Peirce's sign theory as a category of predominantly iconic signs which additionally exhibit a certain degree of symbolicity and display certain indexical traits. In particular, the visual mimeticity of pictures was said to result from a combination of their imagetic and diagrammatic iconicity, whose variable degree was in turn described as being inversely proportionate to the degree of symbolicity exhibited by these static planar signs. It was also said that as a category, pictures exhibit indexicality on account of their inherent incompleteness (Bordwell [1985] 1986: 101), whereby

pictures are invariably less specific than their semiotic objects. As for verbal representations, they have been characterized for the purposes of this examination as predominantly symbolic signs whose meanings are encoded linguistically, that is, with the use of the lexical and grammatical resources of a natural language.

The examination of the scope of the pictorial tacitly assumed in multimodal cognitive linguistic explorations of static planar ensembles has demonstrated that it is usually circumscribed much too broadly and, as a result, the category of pictorial representations is implicitly made to include static planar signs that do not only lack visual mimeticity, but also differ from each other in their semiotic make-up to the degree that there is little reason for treating them as members of a single category. For example, it was pointed out that Forceville's (2005: 75) "[p]ictorial signals of anger" fail to exhibit the combination of imagetic and diagrammatic iconicity that makes pictures visually mimetic due to the fact that these "signals" (Forceville 2005: 75) represent a non-visual concept (anger) in visual terms. For this reason, they should not have been referred to by Forceville (2005) as *pictorial*. Additionally, Forceville's (2005) visually non-mimetic representations of anger were characterized as multicyclic combinations of Peirce's modes of signification, and they were shown to comprise three very different kinds of static planar signs: (1) realistic representations of realistic indices ("bulging eyes," "tightly closed eyes," "wide mouth," *etc.*), (2) realistic representations of unrealistic indices ("smoke"), and (3) unrealistic representations of realistic indices ("bold face"). It was concluded that the static planar signs of anger discussed by Forceville (2005), which do not qualify as pictures in the everyday sense of the word, are best characterized as visually non-mimetic, indexical representations of anger, with the proviso that the first two out of the three kinds of static planar signs listed above establish their indexical connection with anger via visually mimetic representations of the (realistic or unrealistic) symptoms of this emotion. As regards the scope of the verbal, the examination of multimodal cognitive linguistic explorations of static planar ensembles included in the scope of this article has shown that it is also circumscribed too broadly, as including any written forms whatsoever. This is problematic because written forms are not necessarily verbal. As shown by Kowalewski (2015: 31), written representations of onomatopoeias found in comics (*skroww*, *baroom*, *kerashh*, *etc.*) do not qualify as words in written form because they do not correspond to conventional linguistic units stored in the lexicon, they cannot be assigned to particular lexical categories, and they are not subject to the rules of the grammar.

Summing up, it seems that the broadening of the scope of the pictorial and the verbal observable in the studies of static planar ensembles examined in the analytical part of this article, which may likely have been motivated by descriptive convenience, has resulted in analyses that are inherently flawed insofar as they present a vastly simplified view of the intricate process of multicyclic semiosis whereby structurally complex static planar ensembles become meaningful. Granted, the visual structure of some static planar ensembles consists of little more than a relatively balanced combination of uncontroversially pictorial signs and uncontroversially verbal signs, as is the case with nearly all printed advertisements discussed by Forceville (1996), but in other cases the overall meaning of static planar ensembles (cartoons, comics, corporate logos, *etc.*) emerges from an interaction of uncontroversially pictorial signs (for example, visually mimetic drawings), uncontroversially verbal signs (for example, written language), and a whole range of visual representations that are neither pictorial nor verbal (written onomatopoeias, speech balloons, thought bubbles, motion/speed lines, *etc.*). The view that static planar ensembles "achieve meaning" (Heer and Worcester 2009: xiii; emphasis in the original) through an interaction of signs occupying various locations along a continuum between pictures and words has — for a long time now — been espoused

by researchers working in the field of comics scholarship (notably, McCloud [1993] 1994; Saraceni 2003; Szyłak 2000; Cohn 2007, 2013). Only several years ago a comprehensive model of the meaning-making potential of the comics medium, which can also be productively applied in explorations of static planar signs other than comics, was formulated by Kowalewski (2015), who proposed to characterize the semiosis of comics in terms of Peirce's sign theory as relying on the establishment of indexical, iconic, and symbolic relations between signs and their objects in consecutive semiotic cycles. The accomplishments of comics scholars have been reviewed in this article, with particular emphasis on Kowalewski's (2015) comprehensive model, with the hope that in a truly interdisciplinary spirit they will inform further research in the area of multimodal cognitive linguistics because what they offer is a considerably more realistic view of the workings of static planar ensembles than the one emerging from the imposition of a binary construal onto a range of qualitatively diverse static planar signs.

### References

- Abbott, Michael, Charles Forceville (2011) "Visual Representations of Emotion in Manga: LOSS OF CONTROL IS LOSS OF HANDS in *Azumanga Daioh* Volume 4." [In:] *Language and Literature* 20/2; 1–22.
- Arnheim, Rudolf (1969) *Visual Thinking*. Berkeley, CA, Los Angeles, London: University of California Press.
- Atkin, Albert ([2006] 2010) "Peirce's Theory of Signs." [In:] Edward N. Zalta (ed.) (2015) *Stanford Encyclopedia of Philosophy*. Summer 2015 edition. [At:] <<http://plato.stanford.edu/entries/peirce-semiotics/>> ED: 17 October 2020.
- Barker, Martin (1989) *Comics: Ideology, Power, and the Critics*. Manchester: Manchester University Press.
- Barthes, Roland ([1964] 1986) ["Rhétorique de l'image." In: *Communications* 4; 40–51.] Translated into English by Richard Howard. "Rhetoric of the Image." [In:] Roland Barthes *The Responsibility of Forms*. Oxford: Blackwell; 21–40.
- Bateman, John A (2008) *Multimodality and Genre: A Foundation for the Systematic Analysis of Multimodal Documents*. Basingstoke, New York: Palgrave Macmillan.
- Bateman John A., Janina Wildfeuer, Tuomo Hiippala (2017) *Multimodality: Foundations, Research and Analysis: A Problem-Oriented Introduction*. Berlin, Boston: Walter de Gruyter.
- Bezemer, Jeff, Gunther Kress (2016) *Multimodality, Learning and Communication: A Social Semiotic Frame*. London, New York: Routledge.
- Bongco, Mila ([2000] 2013). *Reading Comics: Language, Culture, and the Concept of the Superhero in Comic Books*. London, New York: Routledge.
- Bordwell, David ([1985] 1986). *Narration in the Fiction Film*. London: Routledge.
- Caballero, Rosario (2009) "Cutting across the Senses: Imagery in Winespeak and Audiovisual Promotion." [In:] Charles Forceville, Eduardo Urios-Aparisi (eds.) *Multimodal Metaphor*. Berlin, New York: Mouton de Gruyter; 73–94.
- CALD: *Cambridge Advanced Learner's Dictionary*. Fourth edition ([2003] 2013) Edited by Colin McIntosh. Cambridge: Cambridge University Press.
- Carey, Mike, Chris Bachalo, Humberto Ramos, Mark Brooks (2007) *X-Men: Supernovas*. New York: Marvel Publishing.
- Carrier, David (2000) *The Aesthetics of Comics*. University Park: Pennsylvania State University Press.

- Cienki, Alan (1998) "Metaphoric Gestures and Some of Their Relations to Verbal Metaphoric Expressions." [In:] Jean-Pierre Koenig (ed.) *Discourse and Cognition: Bridging the Gap*. Stanford: Center for the Study of Language and Information; 189–204.
- Cohn, Neil (2007) "A Visual Lexicon." [In:] *Public Journal of Semiotics* 1/1; 53–84.
- Cohn, Neil (2013) *The Visual Language of Comics: Introduction to the Structure and Cognition of Sequential Images*. London: Bloomsbury Publishing.
- Danaher, David (1998) "Peirce's Semiotic and Conceptual Metaphor Theory." [In:] *Semiotica* 119/1–2; 171–207.
- Díaz-Vera, Javier E. ([2013] 2015) "Woven Emotions: Visual Representations of Emotions in Medieval English Textiles." [In:] María Jesús Pinar Sanz (ed.) *Multimodality and Cognitive Linguistics*. Amsterdam/Philadelphia: John Benjamins; 45–60.
- Díaz-Vera, Javier E., Teodoro Manrique-Antón (2015) "'Better Shamed before One than Shamed before All': Shaping Shame in Old English and Old Norse Texts." [In:] Javier E. Díaz-Vera (ed.) *Metaphor and Metonymy across Time and Cultures: Perspectives on the Sociohistorical Linguistics of Figurative Language*. Berlin, New York: Mouton de Gruyter, 225–264.
- Di Liddo, Annalisa (2009) *Alan Moore: Comics as Performance, Fiction as Scalpel*. Jackson: University Press of Mississippi.
- Duncan, Randy, Matthew J. Smith (2009) *The Power of Comics: History, Form and Culture*. New York, London: Continuum.
- Eerden, Bart (2009) "Anger in Asterix: The Metaphorical Representation of Anger in Comics and Animated Films." [In:] Charles Forceville, Eduardo Urios-Aparisi (eds.) *Multimodal Metaphor*. Berlin, New York: Mouton de Gruyter; 243–264.
- Eggertsson, Gunnar Theodór, Charles Forceville (2009) "Multimodal Expressions of the HUMAN VICTIM IS ANIMAL Metaphor in Horror Films." [In:] Charles Forceville, Eduardo Urios-Aparisi (eds.) *Multimodal Metaphor*. Berlin, New York: Mouton de Gruyter; 429–449.
- Eisner, Will ([1985] 2008) *Comics and Sequential Art*. Reprint. New York: W. W. Norton and Company, Inc.
- Elleström, Lars (2013) "Spatiotemporal Aspects of Iconicity." [In:] Lars Elleström, Olga Fischer, Christina Ljungberg (eds.) *Iconic Investigations*. Amsterdam/Philadelphia: John Benjamins; 95–118.
- El Rafaie, Elisabeth (2009) "Metaphor in Political Cartoons: Exploring Audience Responses." [In:] Charles Forceville, Eduardo Urios-Aparisi (eds.) *Multimodal Metaphor*. Berlin, New York: Mouton de Gruyter; 173–196.
- Farias, Priscila, João Queiroz (2006) "Images, Diagrams, and Metaphors: Hypoicons in the Context of Peirce's Sixty-six Fold Classification of Signs." [In:] *Semiotica* 1/4; 1–21.
- Fauconnier, Gilles, Mark Turner (2002) *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities*. New York: Basic Books.
- Forceville, Charles (1996) *Pictorial Metaphor in Advertising*. London, New York: Routledge.
- Forceville, Charles (2005) "Visual Representations of the Idealized Cognitive Model of Anger in the Asterix Album *La Zizanie*." [In:] *Journal of Pragmatics* 37/1; 69–88.
- Forceville, Charles (2011) "Pictorial Runes in *Tintin and the Picaros*." [In:] *Journal of Pragmatics* 43/3; 875–890.
- Forceville, Charles ([2013] 2015) "Metaphor and Symbol: SEARCHING FOR ONE'S IDENTITY IS LOOKING FOR A HOME in Animation Film." [In:] María Jesús Pinar Sanz (ed.) *Multimodality and Cognitive Linguistics*. Amsterdam/Philadelphia: John Benjamins; 27–44.
- Forceville, Charles, Eduardo Urios-Aparisi (eds.) (2009) *Multimodal Metaphor*. Berlin, New York: Mouton de Gruyter.

- Forceville, Charles, Tony Veale, Kurt Feyaerts (2010) "Balloonics: The Visuals of Balloons in Comics." [In:] Joyce Goggin, Dan Hassler-Forest (eds.) *The Rise and Reason of Comics and Graphic Literature: Critical Essays on the Form*. Jefferson: McFarland; 56–73.
- Gombrich, Ernst Hans ([1960] 1961) *Art and Illusion: A Study in the Psychology of Pictorial Representation*. Princeton: Princeton University Press.
- Górska, Elżbieta (2020) *Understanding Abstract Concepts across Modes in Multimodal Discourse: A Cognitive-Linguistic Approach*. New York: Routledge.
- Gosciny, René (writer), Albert Uderzo (artist) ([1970] 1972) [*La Zizanie. Pilote #531–552*] Translated into English by Anthea Bell, Derek Hockridge. *Asterix and the Roman Agent*. Leicester: Brockhampton Press.
- Grady, Joseph E (1997) *Foundations of Meaning: Primary Metaphors and Primary Scenes*. Unpublished PhD dissertation. University of California at Berkeley.
- Groensteen, Thierry ([1999] 2007) [*Système de la bande dessinée*. Paris: Presses Universitaires de France] Translated into English by Bart Beaty, Nick Nguyen. *The System of Comics*. Jackson: University Press of Mississippi.
- Harvey, Robert C. (1994) *The Art of the Funnies: An Aesthetic History*. Jackson: University Press of Mississippi.
- Harvey, Robert C. (1996) *The Art of the Comic Book: An Aesthetic History*. Jackson: University Press of Mississippi.
- Heer, Jeet, Kent Worcester (2009) "Introduction." [In:] Jeet Heer, Kent Worcester (eds.) *A Comics Studies Reader*. Jackson: University Press of Mississippi; xi–xv.
- Hergé (Georges Remi) ([1976] 1976) [*Tintin et les Picaros*. Paris and Tournai: Éditions Casterman] Translated into English by Leslie Lonsdale-Cooper, Michael Turner. *Tintin and the Picaros*. London: Methuen Children's Books.
- Inge, M. Thomas (1990) *Comics as Culture*. Jackson, London: University Press of Mississippi.
- Jappy, Tony (2013) *Introduction to Peircean Visual Semiotics*. London: Bloomsbury Publishing.
- Johansen, Jørgen Dines, Svend Erik Larsen ([1994] 2002) [*Tegn i Brug*. Copenhagen: Amanda/Dansklærerforeningens Forlag] Translated into English by Dinda L. Gorfée, John Irons. *Signs in Use: An Introduction to Semiotics*. London, New York: Routledge.
- Johnson, Mark (1987) *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason*. London, Chicago: The University of Chicago Press.
- Kennedy, John M. (1982) "Metaphor in Pictures." [In:] *Perception* 11/5; 589–605.
- Khordoc, Catherine (2001) "The Comic Book's Soundtrack: Visual Sound Effects in *Asterix*." [In:] Robin Varnum, Christina T. Gibbons (eds.) *The Language of Comics: Word and Image*. Jackson: The University Press of Mississippi; 156–173.
- Koller, Veronika (2009) "Brand Images: Multimodal Metaphor in Corporate Branding Messages." [In:] Charles Forceville, Eduardo Urios-Aparisi (eds.) *Multimodal Metaphor*. Berlin, New York: Mouton de Gruyter; 45–71.
- Kowalewski, Hubert (2015) "From Icono-linguistic Unity to Semiotic Continuity: An Alternative Description of Semiotic Repertoire of Comics." [In:] *International Journal of Comic Art* 17/1; 24–44.
- Kunzle, David (1973) *The Early Comic Strip: Narrative Strips and Picture Stories in the European Broadsheet from c.1450 to 1825*. Berkeley: University of California Press.
- Lakoff, George (1993) "The Contemporary Theory of Metaphor." [In:] Andrew Ortony (ed.) ([1979] 1993) *Metaphor and Thought*. Second edition. Cambridge, New York, Melbourne: Cambridge University Press; 202–251.

- Lakoff, George, Mark Johnson (1980) *Metaphors We Live By*. London, Chicago: The University of Chicago Press.
- Langlotz, Andreas (2015) "Language, Creativity, and Cognition." [In:] Rodney H. Jones (ed.) *The Routledge Handbook of Language and Creativity*. London, New York: Routledge, 40–60.
- LDOCE: *Longman Dictionary of Contemporary English*. Fifth edition ([1978] 2009) Edited by Michael Mayor. Harlow: Pearson Education Limited.
- Lombardo, Patrizia ([1989] 2010). *The Three Paradoxes of Roland Barthes*. Athens, London: University of Georgia Press.
- McCloud, Scott ([1993] 1994). *Understanding Comics: The Invisible Art*. Reprint. New York: Harper Perennial.
- McCloud, Scott (2006) *Making Comics: Storytelling Secrets of Comics, Manga and Graphic Novels*. New York: Harper Collins.
- Miodrag, Hannah (2013) *Comics and Language: Reimagining Critical Discourse on the Form*. Jackson: University Press of Mississippi.
- Moya Guijarro, A. Jesús ([2013] 2015) "Visual Metonymy in Children's Picture Books." [In:] María Jesús Pinar Sanz (ed.) *Multimodality and Cognitive Linguistics*. Amsterdam/Philadelphia: John Benjamins; 115–130.
- Müller, Cornelia, Alan Cienki (2009) "Words, Gestures, and Beyond: Forms of Multimodal Metaphor in the Use of Spoken Language." [In:] Charles Forceville, Eduardo Urios-Aparisi (eds.) *Multimodal Metaphor*. Berlin, New York: Mouton de Gruyter; 297–328.
- Nöth, Winfried ([1990] 1995) *Handbook of Semiotics*. Bloomington: Indiana University Press.
- OALD: *Oxford Advanced Learner's Dictionary*. Sixth edition ([1948] 2000) Edited by Albert Sidney Hornby, Sally Wehmeier. Oxford: Oxford University Press.
- Peirce, Charles Sanders (n.d.) Unpublished manuscripts. Indiana University. (References to MS with number, year and pagination as in the files of the library.)
- Peirce, Charles Sanders (1931–1958) *Collected Papers of Charles Sanders Peirce*. Volumes 1–6 edited by Charles Hartshorne, Paul Weiss; volumes 7–8 edited by Arthur W. Burks. Cambridge, MA: Harvard University Press.
- Phoenix, Jack (2020) *Maximizing the Impact of Comics in Your Library: Graphic Novels, Manga, and More*. Santa Barbara: Libraries Unlimited.
- Pinar Sanz, María Jesús (ed.) (2015) *Multimodality and Cognitive Linguistics*. Amsterdam/Philadelphia: John Benjamins.
- Popa, Diana E. ([2013] 2015) "Multimodal Metaphors in Political Entertainment." [In:] María Jesús Pinar Sanz (ed.) *Multimodality and Cognitive Linguistics*. Amsterdam/Philadelphia: John Benjamins; 79–95.
- Postema, Barbara (2013) *Narrative Structure in Comics: Making Sense of Fragments*. Rochester: RIT Press.
- Potsch, Elizabeth, Robert F. Williams (2012) "Image Schemas and Conceptual Metaphor in Action Comics." [In:] Frank Bramlett (ed.) *Linguistics and the Study of Comics*. Basingstoke, New York: Palgrave Macmillan; 13–36.
- Rohdin, Mats (2009) "Multimodal Metaphor in Classical Film Theory from the 1920s to 1950s." [In:] Charles Forceville, Eduardo Urios-Aparisi (eds.) *Multimodal Metaphor*. Berlin, New York: Mouton de Gruyter; 403–428.
- Rollins, Mark (2001) "Pictorial Representation." [In:] Gaut, Berys, Dominic McIver Lopes (eds.) *The Routledge Companion to Aesthetics*. London, New York: Routledge; 297–312.
- Sabin, Roger (1993) *Adult Comics: An Introduction*. London, New York: Routledge.
- Saraceni, Mario (2003) *The Language of Comics*. London, New York: Routledge.

- Schilperoord, Joost, Alfons Maes (2009) "Visual Metaphoric Conceptualization in Editorial Cartoons." [In:] Charles Forceville, Eduardo Urios-Aparisi (eds.) *Multimodal Metaphor*. Berlin, New York: Mouton de Gruyter; 213–240.
- Shen, Yeshayahu (2008) "Metaphor and Poetic Figures." [In:] Raymond W. Gibbs, Jr. (ed.) *The Cambridge Handbook of Metaphor and Thought*. Cambridge: Cambridge University Press; 295–307.
- Shinohara, Kazuko, Yoshihiro Matsunaka (2009) "Pictorial Metaphors of Emotion in Japanese Comics." [In:] Charles Forceville, Eduardo Urios-Aparisi (eds.) *Multimodal Metaphor*. Berlin, New York: Mouton de Gruyter; 265–293.
- Sweetser, Eve (1998) "Regular Metaphoricity in Gesture: Bodily-based Models of Speech Interaction." *Actes du 16e Congrès International des Linguistes* [CD-ROM]. Oxford: Elsevier Sciences.
- Szawerna, Michał (2012) "Toward a Cognitive Approach to Analyzing Comics: The Contrasting Construals of Comics in Similarity-based Theories of Conceptual Representation." [In:] Zdzisław Wąsik, Piotr Czajka, Michał Szawerna (eds.) *Alternate Construals in Language and Linguistics*. Wrocław: Wydawnictwo Wyższej Szkoły Filologicznej we Wrocławiu, 59–74.
- Szawerna, Michał (2017) *Metaphoricity of Conventionalized Diegetic Images in Comics: A Study in Multimodal Cognitive Linguistics*. Frankfurt am Main: Peter Lang.
- Szawerna, Michał (2020) "Modality, Medium, and More: A Toolkit for the Multimodal Cognitive Linguist. Sidenotes on *Understanding Abstract Concepts across Modes in Multimodal Discourse: A Cognitive-Linguistic Approach* (2020) by Elżbieta Górska." [In:] *Anglica Wratislaviensia* 58; 187–200.
- Szyłak, Jerzy (2000) *Poetyka komiksu. Warstwa ikoniczna i językowa*. Gdańsk: Wydawnictwo Słowo/Obraz Terytoria.
- Teng, Norman Y. (2009) "Image Alignment in Multimodal Metaphor." [In:] Charles Forceville, Eduardo Urios-Aparisi (eds.) *Multimodal Metaphor*. Berlin, New York: Mouton de Gruyter; 197–211.
- Urios-Aparisi, Eduardo (2009) "Interaction of Multimodal Metaphor and Metonymy in TV Commercials: Four Case Studies." [In:] Charles Forceville, Eduardo Urios-Aparisi (eds.) *Multimodal Metaphor*. Berlin, New York: Mouton de Gruyter; 95–117.
- Wartenberg, Thomas E. (2012) "Wordy Pictures: Theorizing the Relationship between Image and Text in Comics." [In:] Aaron Meskin, Roy T. Cook (eds.) *The Art of Comics: A Philosophical Approach*. Malden, Oxford: Wiley-Blackwell; 87–104.
- Waugh, Coulton ([1947] 1991) *The Comics*. Reprint. Jackson, London: University Press of Mississippi.
- Yu, Ning (2009) "Nonverbal and Multimodal Manifestation of Metaphors and Metonymies: A Case Study." [In:] Charles Forceville, Eduardo Urios-Aparisi (eds.) *Multimodal Metaphor*. Berlin, New York: Mouton de Gruyter; 119–143.
- Yus, Francisco (2009) "Visual Metaphor Versus Verbal Metaphor: A Unified Account." [In:] Charles Forceville, Eduardo Urios-Aparisi (eds.) *Multimodal Metaphor*. Berlin, New York: Mouton de Gruyter; 147–172.