Creativity and multimodal literature

How to cite: Alison Gibbons. 08 Sep 2015, Creativity and multimodal literature from: The Routledge Handbook of Language and Creativity Routledge
Accessed on: 31 Oct 2023
https://www.routledgehandbooks.com/doi/10.4324/9781315694566.ch18

This Document PDF may be used for research, teaching and private study purposes. Any substantial or systematic reproductions, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The publisher shall not be liable for an loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.
Part III

Multimodal and multimedia creativity
Creativity and multimodal literature

Alison Gibbons

Introduction: Creativity, multimodal deviation, and conceptual shifts

Creativity is an inherent possibility of human communication. This is David Crystal’s argument in *Language Play* (1998), in which he understands linguistic creativity as a form of ludic play and a source of human pleasure:

We play with language when we manipulate it as a source of enjoyment either for ourselves or for the benefit of others. I mean ‘manipulate’ literally: we take some linguistic feature – such as a word, a phrase, a sentence, a part of a word, a group of sounds, a series of letters – and make it do things it does not normally do. We are, in effect, bending and breaking the rules of language. And if someone were to ask why we do it, the answer is simply: for fun.

(Crystal, 1998: 1)

Crystal argues that humans use language not merely to communicate, but often also as a means of self-expression and amusement. His reference to making ‘it do things it does not normally do’ points to linguistic deviation as a means of exploiting the compositional features of language – structure, sense, and sound – for creative purposes. Deviation is a device of foregrounding (van Peer, 1986), which attracts attention by disrupting textual patterns or by departing from expected norms in some way. It derives from the work of the Russian formalists – in particular, from the work of three key figures: Viktor Šklovsky (1917), Bohuslav Hávranek (1932), and Jan Mukařovsky (1932). Šklovsky proposed the concept of ‘defamiliarisation’, the essence of which is that art and literary language should not only draw attention to its own artfulness and constructedness, but also attract and hold that attention in the creation of literary meaning for the reader. Hávranek and Mukařovsky later built on the concept of defamiliarisation and it is from their work that the term ‘foregrounding’ stems. Although deviation is not the only creative technique available to writers, multimodal literature is heavily reliant upon creative deviations, both in terms of deviation from readers’ expectations about the literary conventions of the so-called traditional literary text and in terms of deviations from expected linguistic structures.

When used to describe literature, *multimodality* indicates that the literary work uses multiple semiotic modes for creative narrative purposes: a work of multimodal literature includes not only the verbal orthographic text in the form of printed type; it may also use varied
typography, unusual textual layouts and page designs (including the concrete realisation of type, as in concrete poetry), the inclusions of images such as photographs or illustrations, colour, and so on. Of course, there are no hard-and-fast rules or formulae for identifying or writing a multimodal fiction, just as there can be no checklist or formalised nomenclature of literary-linguistic creativity. Creativity is both context-bound and dependent upon the interpretations of sender and receiver of the communicative act. Indeed, Ronald Carter (2004: 66), in his discussion of literariness and creativity, states that it is:

important to isolate the different degrees of literariness, but even more important to understand that literariness and, by extension, literary language is socially and culturally relative. Because literary language is seen as creative, the same applies by extension to creativity.

Consequently, Carter (2004: 66) goes on to argue that ‘it may be more instructive to see literary and creative uses of language as existing along a cline or continuum rather than as discrete sets of features or as a language-intrinsic or unique “poetical” register’.

In relation to the multimodal literary work specifically, Wolfgang Hallet (2009) has argued that reading printed multimodal fiction involves five ‘conceptual shifts’, which we can think of as literary deviations. The first conceptual shift features the creative process itself – what Hallet (2009: 149) calls the shift from ‘writing to designing’. Creating the multimodal text involves more than just writing; it also involves considering the many different semiotic resources (words, image, colour, layout) available to the writer, who must decide how to use them in creative combinations for literary purposes. This, in turn, leads to a shift in the way in which the text itself is realised – or, as Hallet (2009: 150) puts it, a conceptual shift from ‘the monomodal (verbal) text to multimodal, multimedial texts’.

A third conceptual shift identified by Hallet (2009: 150) is that of the narrator becoming ‘narrator-presenter’, who ‘apart from delivering a story, searches, retrieves and “collects” documents and sources and eventually presents them to the reader, the process of narrating includes “showing” and “presentation”’.

In the remaining two conceptual shifts, Hallet transfers his focus from the creator of the multimodal novel and the composition of the text to the receiver of the multimodal discourse. He speaks of a shift from ‘reading to the transmodal construction of narrative meaning’ (Hallet, 2009: 150). What he means by this is that readers must not only interpret meaning from the linguistic narrative, but also ‘integrate various literacies’ (Hallet, 2009: 150), construing meaning from the combination of modes that exist within the same textual space. Hallet’s (2009: 150) final conceptual shift moves the reader from ‘reader to user’, whereby the reader not only interacts with the multimodal novel through his or her imagination, but also may be required to ‘use’ the text in various ways, such as physically moving it or writing on it.

In this chapter, I explore some of the creative devices of multimodal literature. First, in the next section, I introduce two of the founding works within multimodality studies. These are then placed within the context of the larger field of multimodality studies, which also introduces stylistic approaches to multimodal literature. I go on to provide three examples of multimodal literature: extracts from Mark Z. Danielewski’s (2000) *House of Leaves*, Sherman Alexie’s (2007) *The Absolutely True Diary of a Part-time Indian*, and Joe Meno’s (2006) *The Boy Detective Fails*. Through these three literary examples, the final section illustrates the creative conceptual shifts at play within works of multimodal literature.
Historical development

The explicit study of multimodality stems from the work of Gunther Kress and Theo Van Leeuwen, whose 1996 book *Reading Images: The Grammar of Visual Design* brought multimodality to academic attention. In *Reading Images*, Kress and van Leeuwen set out to create a toolkit for the analysis of texts with visual elements. In their words, they seek to ‘provide inventories of the major compositional structures which have become established as conventions in the course of the history of visual semiotics, and to analyse how they are used to produce meaning by contemporary image-makers’ (Kress & van Leeuwen, 1996: 1). Kress and van Leeuwen (1996: 183, emphasis original) first use the term ‘multimodal’ in their discussion of textual composition – that is, of:

composite visuals, visuals which combine text and image, and, perhaps, other graphic elements, be it on a page or on a television or computer-screen. In the analysis of composite or multimodal texts (and any text whose meanings are realized through more than one semiotic mode is multimodal) the question arises whether the products of the various codes should be analysed separately or in an integrated way; whether the parts should be looked at as interacting with and affecting one another. It is the latter path we will pursue.

They argue both for the importance of composite multimodal texts as artefacts of analysis and for an integrated form of analysis that attends to the combination of modes within a single textual product. They clearly state their aim in the following manifesto:

We seek to be able to look at the whole page as an integrated text. Our insistence on drawing comparison between language and visual communication stems from this perspective. We seek to break down the boundaries between the study of languages and the study of images, and we seek as much as possible, to use compatible language, and compatible terminology in speaking about both, for in actual communication the two and indeed many others come together to form integrated texts.

(Kress & van Leeuwen, 1996: 183, emphasis original)

In doing so, they categorise their work as an extension of existing semiotic analyses of non-linguistic communications (such as Barthes, 1964; O’Toole, 1994), combined with the adoption of systemic functional linguistics (Halliday, 1978; Halliday & Matthiessen, 2004), an approach that they also take forward in later work (see Kress, 2010; Kress & van Leeuwen, 2001). Whilst Kress and van Leeuwen’s work in multimodality studies is important, it has been argued that their approach is founded upon linguistic categories, and therefore treats the devices of language and image inappropriately, as self-identical systems of meaning-making (see Bateman, 2008; Forceville, 1999; Gibbons, 2012).

Kress and van Leeuwen also touch briefly on the tension between the notion of a systematic analytical toolkit, such as the ‘grammar of visual design’ that they offer in *Reading Images*, and the very nature of creativity. They argue that such a toolkit ‘does not, of course, have to stand in the way of creativity’ (Kress & van Leeuwen, 1996: 3). Their discussion continues polemically in a way that accords with Carter’s discussion of literary language and creativity, mentioned above:

Teaching the rules of language has not meant the end of creative uses of language in literature and elsewhere, and teaching visual skills will not spell the end of the arts.
Yet, just as the grammar creatively employed by poets and novelists is, in the end, the same grammar we use when writing letters, memos, reports, so the ‘grammar of visual design’, creatively employed by artists is, in the end, the same grammar we need when producing attractive layouts, images, diagrams, for our course handouts, reports, brochures, communiques, and so on.

(Kress & van Leeuwen, 1996: 3)

Perhaps to evidence the false dichotomy that places analytical rigour in opposition to textual creativity, through the course of Reading Images, Kress and van Leeuwen investigate a number of different multimodal text types, from children’s drawings to magazine articles and film stills. Such diversity makes the book a pioneering work within the canon of multimodality studies.

Another valuable text within multimodality studies is Anthony Baldry and Paul Thibault’s (2006) Multimodal Transcription and Text Analysis, which can be seen as both a departure from, and an extension of, Kress and van Leeuwen’s work. Baldry and Thibault’s methodology is similarly drawn from systemic functional linguistics, although their approach is more overtly one of transcription – a method in which the composite text is broken down into its respective objects for analysis. A weakness of this method is therefore that it sometimes becomes a case of cataloguing the multimodal composition of the text, rather than considering how these elements work together to form integrated meanings. Nevertheless, in Multimodal Transcription and Text Analysis, Baldry and Thibault analyse a range of textual genres, including printed texts, web pages, and filmic works.

Kress and van Leeuwen’s (1996) Reading Images and Baldry and Thibault’s (2006) Multimodal Transcription and Text Analysis stand as foundational readings in multimodality studies, particularly in terms of providing early analytical approaches to multimodal texts. Both have been charged with suffering from an overreliance on linguistic methodology to analyse more than language, inevitable resulting from the systemic functional basis of their approaches. Jewitt (2009a: 26), in her introduction to The Routledge Handbook of Multimodal Analysis, summarises this by saying that the discipline has often suffered from ‘criticism that multimodality is a kind of “linguistic imperialism” that imports and imposes linguistic terms on everything. But these criticisms overlook the fact that much of the work on multimodality has its origins in a particular strand of linguistics’. Indeed, whilst ‘language is part of a multimodal ensemble’, Jewitt (2009a: 15) maintains that the study of multimodality ‘proceeds on the assumption that representation and communication always draw on a multiplicity of modes, all of which have the potential to contribute equally to meaning’.

Multimodality studies and stylistics

Kress and van Leeuwen’s (1996) Reading Images essentially launched multimodality studies as an academic discipline in its own right. That discipline has inevitably diversified as it has grown, leading Jewitt (2009b: 28) to speak of multimodality studies as ‘roughly categorized into three main approaches’. She identifies these as the social semiotic approach (seen most clearly in Kress and van Leeuwen’s work), a discourse analysis approach (typified by Baldry & Thibault, 2006; O’Halloran, 2004, 2005; O’Toole, 1994), and an interaction analysis approach (with principal works by Norris, 2004; Scollon & Scollon, 2003). To this, we can add a fourth domain within multimodality studies: the stylistic approach. The stylistic approach to multimodality studies essentially seeks to explore how multimodal literary works are composed and how the various modes interact with each other to produce literary meaning. There are two strands of multimodal stylistics: the social semiotic approach, and the cognitive approach.
The former clearly follows existing work, grounding itself in Halliday’s systemic functional principals and thus utilising many of the ‘tools’ put forward by Kress and van Leeuwen.


Foer’s novel is concerned with a young boy named Oskar who loses his father in the 9/11 terrorist attack on the World Trade Center. In the book’s conclusion, Foer provides a series of sequential photographs of a figure of a falling man. Essentially, this is a flipbook section in which the figure, conversely, moves upwards – a gesture that echoes the preceding linguistic narrative in which Oskar reverses the pages in his own scrapbook so that ‘the man was floating up through the sky’ (Foer, 2005: 325) as a way in which to imagine an alternate fate for his father. In her analysis, Nørgaard relates the falling man flipbook to Kress and van Leeuwen’s (1996: 186–206) deconstruction of visual composition along horizontal and vertical axes. The left side of an image is understood as ‘Given’ information, which is already known, and the right side of the image is understood to be ‘New’. Similarly, the bottom part of an image is ‘Real’ – again, known and existing – while the upper part of an image is ‘Ideal’ – that is, an idealised promise. This model is an adaptation by Kress and van Leeuwen of Halliday’s (1967) discussion of ‘given-and-new’ as a means of understanding linguistic knowledge and prosodic structure. Nørgaard (2010a: 123) uses this to conclude that the act of flipping the pages makes ‘the man float from Real to Ideal’, and thus argues that Oskar’s attempt to revise his father’s tragic fate is realised visually by the positioning of the figure of his father moving within the flipbook section from real to ideal textual space.

Elsewhere, I have questioned the value of using Kress and van Leeuwen’s model of textual layout as the most accurate model of understanding how readers process this narrative event (see Gibbons, 2012: 161–2), since the application is not unproblematic as a result of the fact that visual images do not operate with the same structural or temporal logic as verbal language. Nevertheless, Nørgaard’s social semiotic stylistic analysis clearly demonstrates that meaning-making in Foer’s novel is multimodal and that, in order to interpret Oskar’s narrative hopes, readers must attend to both textual and visual resources.

The second strand of multimodal stylistics is, as mentioned previously, a cognitive approach. This can be seen in work by Masako K. Hiraga (2005) and Alison Gibbons (2008, 2009, 2010, 2012, 2013). Hiraga restricts his focus to the analysis of metaphors in multimodal literature. In Metaphor and Iconicity, he devotes two chapters to the analysis of texts with a visual element using both English and Japanese examples (Hiraga, 2005: 57–126). Of most interest is his analysis of Roger McGough’s (1971) poem ‘40-LOVE’, a poem that undoubtedly relies upon its typographical arrangement for the conception of its meaning. With words positioned in two columns (each one syllable wide) to be read alternately, the poem is both literally about a tennis match and metaphorically about the relationship between a middle-aged couple. Hiraga (2005: 106) makes this clear, stating that the poem ‘can be seen as a metaphor in which a love relationship is understood in terms of a game of tennis’. While Hiraga does not explicitly name the conceptual metaphor at work here, comprehension of the poem depends upon it nevertheless: LOVE IS A GAME.

In looking at the linguistic details of ‘40-LOVE’, Hiraga (2005: 106) articulates, ‘small lexical elements are repeated in pairs 12 times in a row over the two pages as if they were the monotonous movement of a tennis ball across a net’. Actually, this monotonous movement is built into the text’s structure only implicitly. The iconic arrangement of text serves to
generate the illusion of a tennis ball’s movement by guiding the optical path that the reader must take in order to read and comprehend the poem’s semantic content. Consequently, readers have a performatory role to play in the construction and actualisation of metaphoric meaning. What Hiraga’s analysis shows is that meaning, including metaphorical meaning, emerges in ‘40-LOVE’ from the linguistic content, imagistic design, and the reader’s relationship with the text.

My own work in cognitive poetics and multimodal literature offers a critical synthesis of tools within multimodality studies and within stylistics. As such, it draws on techniques suggested by the likes of Kress and van Leeuwen, and Baldry and Thibault, as well as attending to literary-linguistic devices, particularly foregrounding, as mentioned at the start of this chapter, as well as frameworks from the cognitive study of literature (see Gavins & Steen, 2003; Stockwell, 2002). In Multimodality, Cognition, and Experimental Literature (Gibbons, 2012), I present cognitive-poetic analyses of four twenty-first-century multimodal novels, of which one example will suffice here.

In a narrative episode from Mark Z. Danielewski’s (2000) House of Leaves, one of the central characters, Navidson, is exploring a labyrinthine dark space that has appeared in the centre of his home and finds himself climbing what seems to be an endless ladder. The words of the text are arranged so that they form a visual ladder on the page. Thus, in the course of reading, the reader’s eyes must travel upwards, climbing the ladder too. At the top of the ladder is an arrow symbol (↓), which transpires to be a footnote directing the reader to a smaller ladder of words that the reader must read downwards and the words of which are arranged in the opposite direction from that of the first ladder (for example upside-down). After rotating the book, the ladder can be read as follows:

↓ Erich Kästner in Öberge Weinberge (Frankfurt, 1960, p.95)
   comments on the force of vertical meanings:
   The climbing of a mountain reflects redemption. That is due to the force of the word ‘above,’
   and the power of the word ‘up.’ Even those who have long ceased to believe in Heaven
   and Hell, cannot exchange the words ‘above’ and ‘below.’

(Danielewski, 2000: 441)

This reflection upon the words ‘up’, ‘above’, and ‘below’ is essentially a meditation founded on two conceptual metaphors, \textit{good is up} and \textit{bad is down}. In this context, these conceptual
metaphors would be present regardless of textual design, since they arise from linguistic content. However, the reader’s own optical journey – ascending the first ladder and then descending the second – has two effects. First, it adds intensity to their meaning, since the physical act of reading this part of *House of Leaves* can be connected to the philosophical musing. Secondly, it also refracts the meaning of the conceptual metaphors back onto the experience of the narrative. In the first ladder, Navidson’s climb is related as an endless and exhausting journey, raising questions about whether he will succeed or whether his fatigue will overwhelm him. He does, though, reach the top, with readers informed that he now finds himself ‘standing inside a very’ (Danielewski, 2000: 441). Since the arrow leads readers to the second ladder, Danielewski’s topsy-turvy page design has a disorientating effect for the reader and the narrative delay as to Navidson’s location creates a cliff-hanger, whereby readers must turn the page for the reveal. As such, readers’ understandings of **GOOD IS UP** and **BAD IS DOWN**, and the dual directionality of the ladder’s reading paths (up–down), may raise anxiety as to the nature of the page-turning location overleaf: Will it be good or bad?

As with the examples from Hiraga and Nørgaard, my own reading of *House of Leaves* (see Gibbons, 2010: 71–5, for a more comprehensive study of this passage) shows that, within analyses of multimodal literature, language is a central feature. However, multimodal stylistic analysis involves interpreting linguistic structures within the context of other modal systems. Unusual textual layouts, for instance, are deviant in terms of readers’ expectations about how novels tend to look, and consequently create unusual reading paths. Moreover, reader involvement is often a crucial and creative element of multimodal literature, and affects the ways in which readers experience narratives and their story worlds.

**Multimodal literature in practice**

This chapter has, thus far, sought to introduce the field of multimodal literature and the different critical approaches that have emerged within multimodality studies. The creative deviations and conceptual shifts brought about by multimodal literature – in comparison to what might be considered more conventionally designed literary works – have necessitated critical approaches that attend to both language in literature and to the other semiotics modes at play. In the previous section, in order to demonstrate the social semiotic approach and the cognitive approach to multimodal literature within stylistics, I reviewed examples from analytical works, in the process also providing literary illustrations. This section will offer three further examples of multimodal literature, paying particular attention to the creative deviations that they employ. Each extract has been selected in order to demonstrate the conceptual shifts of multimodal literature suggested by Hallet (2009) and discussed in the opening to this chapter.

**Example 1: Writing to designing and mono- to multimodal**

The first example comes from Mark Z. Danielewski’s (2000) *House of Leaves*, an extract from which was discussed in the previous section of this chapter. Not only did the novel notoriously take its author ten years to write, but Danielewski also had to typeset the manuscript himself, since, at that time, the unusual textual layouts that the author had in mind were far from publishing norm. At 709 pages, *House of Leaves* features an array of concrete-poetic designs, the layouts of verbal text alternating from one page to the next. Thus, whilst the predominant mode in *House of Leaves* is written or typed verbal language, Danielewski exploits the visual dimension of language to shift the novel from the monomodal to a multimodal fiction.
At the heart of the novel is the story of Will Navidson and his family, who have just moved into a house on Ash Tree Lane. The house initially appears to be unremarkable – until an unfamiliar door appears in the master bedroom. The door opens onto a dimly lit corridor and a series of disconcerting explorations into the dark space begin to reveal limitless proportions to the house’s dark interior. The cavernous, labyrinthine, dark space becomes increasingly difficult to navigate. In one episode, Navidson, who has just found himself alone on a stairway, is equipped with a rope tied to the bottom of the final banister. On page 288, the text consists of three lines, all positioned at the bottom of the page:

Then as the stairway starts getting darker and darker and as that faintly illuminated circle above – the proverbial light at the end of the tunnel – starts getting smaller and smaller, the answer becomes clear:

(Danielewski, 2000: 288)

Whilst the blank white space of the page above the text might be considered deviant in terms of its departure from the conventional norms of printed books, the visual layout of the verbal narrative here is relatively unchallenging for readers. Danielewski’s language is also relatively straightforward, although readers might notice the foregrounding alliteration in ‘stairway starts’, and the lexical repetition within ‘darker and darker’ and ‘smaller and smaller’, as well as the parallel structures that these two expressions employ.

A more dramatic layout occurs on the following page. The text is printed upside-down, so to speak, meaning that the reader must rotate the book 180° in order to read the printed text. The text itself can be read as ‘Navidson is sinking. Or the stairway is stretching, expanding.’ The arrangement of text on the page, though, is particularly striking:

(Danielewski, 2000: 289)
The first line, ‘Navidson is’, is easy to read, conforming as it does to traditional Western reading patterns – that is, left to right, and from the top line to the bottom line, on the page. The verb ‘sinking’, however, has been divided into five fragments, or clusters, each a little lower and positioned a little further right than the preceding cluster. The effect of this is that Danielewski’s arrangement of the word is iconic: the form reiterates the meaning with readers’ eye movements in reading descending down the clusters in a motion that might be considered akin to sinking itself. The arrangement of text therefore creates a vector for the reading path by directing the eyes down the word ‘sinking’ and the succeeding punctuation, before making the reader begin the coordinated clause ‘Or the stairway is’ back on the left-hand side. Danielewski performs a similar trick with the verb ‘stretching’, which is arranged vertically down the right-hand side of the page, and the verb ‘expanding’, which spans the bottom of the page horizontally. In both, the spacing between letters grows through the word in order to iconically create the effects of ‘stretching’ and ‘expanding’, respectively. It should also be noted that the ‘s’ in ‘is’ doubles up as the ‘s’ in ‘stretching’, whilst ‘stretching,’ and ‘expanding,’ share the closing ‘g.’. Combined with Danielewski’s use of continuous tense, this acrostic technique allows the visual layout of text on the page to represent the effect of visual space distorting, just as Navidson experiences the shifting architecture in the dark stairway within the narrative world of *House of Leaves*.

Over the next few pages, the stairway continues to grow in proportion, placing increased stress on the rope. Finally, Danielewski (2000: 293–6) tells his readers, ‘the rope snaps’, although the realisation of the text on the page again takes an unusual design. ‘the rope’ sits on page 293, with the letters in the noun ‘rope’ arranged vertically down the right-hand side of the page in order to visually represent the length of rope. The word ‘snaps’ is separated into three parts: ‘sn–’ towards the bottom of page 294, ‘–a–’ which sits close to the middle of page 295, and ‘–ps . . .’ at the top of page 296. Divided in this way, the word is a visual embodiment of its meaning. Furthermore, as Danielewski explains in an interview (McCaffery & Gregory, 2003), the fragmented word takes on new significance for readers who realise that, backwards, it spells ‘spans’, an action that it also fulfils since each segment is placed on a separate page. Whilst the word is, as Danielewski (quoted in McCaffery & Gregory, 2003: 122) puts it, ‘a literal, thematic, and semantic representation of all that’s happening at that moment in the novel’, it also embodies two opposing meanings simultaneously, thus exposing the relationship between words.

As this discussion of *House of Leaves* demonstrates, design plays an important role within multimodal fiction. In the case of *House of Leaves*, Danielewski uses textual concrete-poetic layouts in ways that exploit the semantic meaning of textual content. For readers, this means that their interpretations of the novel are based on both the described actions of the narrative world and their experiential relationship to the text itself.

**Example 2: Narrator to narrator-presenter**

The conceptual shift in which the narrator becomes a presenter of visual information is a shift from telling to telling-and-showing. The ease and more economic cost of printing images has meant that there are many published contemporary novels that feature narrator-presenters. Indeed, this is even the case in novels that feature a relatively low level of multimodality. What I mean by this is that the narrator might present visual images to the reader, but for the most part these images do not visually integrate with the written content – for instance as in the sort of concrete-poetic designs of *House of Leaves* (Danielewski, 2000). Nevertheless, it is a shift that is worth discussing through recourse to a short example.
One novel that features a narrator-presenter is *The Absolutely True Diary of a Part-time Indian* by Sherman Alexie (2007), illustrated by Ellen Forney. The novel’s protagonist and narrator is Junior, a native American boy from the Spokane Indian Reservation who considers himself to be a budding cartoonist. Junior’s cartoonic drawings feature as images within the novel. What makes Junior a narrator-presenter is that his narrative explicitly references these cartoons. He is their creator. As such, the images themselves do not function only as illustrations that sit alongside the text, but also have a place within the context of the narrative world. An example will clarify this.

In the first chapter, Junior tells the reader:

> I mostly hang out alone in my bedroom and read books and draw cartoons. Here’s one of me:

![Junior's cartoon](image)

(Alexie, 2007: 4–5)

At this point, the image of Junior’s cartoon appears. In the book, it is designed to look like a scrap of slightly crumpled paper, with tape sticking it down onto the page on top-left and bottom-right corners. In the middle of the image is a gawky-looking boy, with a speech bubble occupying the space to his left. On the right, in a cartoon font, is the objective case of the first-person pronoun – ‘Me’ – and in smaller, handwritten text underneath are the words ‘in all my glory’. The verbal text then continues:

> I draw all the time.
> I draw cartoons of my mother and father; my sister and grandmother; my best friend, Rowdy; and everybody else on the rez.
> I draw because words are too unpredictable.
> I draw because words are too limited.

(Alexie, 2007: 5)

The transmodal construction of meaning must take place here in order for readers to understand the image as being produced by the narrator. Junior’s words of introduction to his picture are an aid in this respect, with the proximal reference of the locative adverb ‘Here’ directly referencing the surface of the page and therefore the picture that is presented. Moreover, in the text that follows, the repetition of ‘I draw’ at the start of four sentences foregrounds Junior as the active subject (as the narrative ‘I’) undertaking the action of the verb ‘draw’.

This extract from *The Absolutely True Diary of a Part-time Indian* exemplifies the narrator-as-presenter that is often to be found in multimodal literature. As the discussion shows, it is important that the verbal text works in conjunction with the images, to anchor them within the context of the narrative world and to ensure that their function is not illustrative, but part of integrative multimodal meaning-making.

**Example 3: Reading to transmodal construction and reader to user**

In the examples from *House of Leaves* (Danielewski, 2000) discussed in this chapter, readers needed to rotate the book in order to read deviantly positioned text, as well as to reconstruct words from fragments often arranged across visual space or even pages. These are, of course, examples of the transmodal construction of narrative meaning, as was the combination of word and image in the discussion of the extract from *The Absolutely True*
Diary of a Part-time Indian (Alexie, 2007). However, acts that require the reader to physically engage with the book as artefact are also examples of readers having to ‘use’ the text in some way. The conceptual shift in which the reader becomes a user of the text sees the reader take on a performative role.

Perhaps a more overt case of the reader-as-user occurs when fictional texts contain hidden messages – ciphers that the reader must decode in order to gain additional textual information. A novel that frequently incorporates secret codes into its narrative is Joe Meno’s (2006) The Boy Detective Fails, which tells the story of the boy detective in question – Billy Argo – who is not, in fact, much of a boy. When Billy leaves for college, devastating events tarnish the detective fantasies of his youth, the most tragic being the death of his sister Caroline, who supposedly commits suicide. After ten years in an asylum unable to deal with the loss, Billy eventually returns to society and takes a telemarketing job. Yet the world is not as it seems; rather, it is a surreal place in which buildings suddenly disappear and Billy, the now not-so-boyish detective, is determined to find out what really happened to his sister.

There are at least seven hidden messages within The Boy Detective Fails, the first of which appears in chapter 22:

The boy detective is surprised to receive a mysterious letter in the mail when he returns to his room. There is no postmark or address. The envelope, in small black handwriting, simply says: To the boy detective. Billy slowly opens the letter, slipping his finger beneath the fold of paper and tearing. Inside is a single piece of yellowed paper that reads:

XI: 5–12–15–15–2

It is a secret code, but sent from who? The boy detective thinks Caroline immediately, and then admonishes himself for thinking that at all. He stares at it for a few moments more before giving up, hiding it under his bed, afraid of what it may be.

Perhaps dear reader, you might help him. Match the code XI with the decoder wheel found on the back flap of this volume.

(Meno, 2006: 113–14)

The letter that Billy finds contains a code, which is then presented to the reader by an omniscient narrator (rather than the more explicit first-person narrator of The Absolutely True Diary of a Part-time Indian). The code itself is foregrounded by means of its visual separation from the surrounding text and its realisation as emboldened type. Whilst some readers might be tempted to skip this in the hope that the hidden message will, in fact, be revealed later in the narrative, the italicised font at the end of the extract explicitly addresses the reader and asks him or her to undertake the task of decoding the message. It does this through the rather archaic literary address of ‘dear reader’, which is immediately followed by the apostrophic ‘you’. The suggestion that ‘you’, the reader, should actively engage in the task of decoding the message is initially hedged with the epistemic modal in the verb phase ‘might help’. However, the composition of the succeeding sentence eliminates the sense of possibility and readerly choice by issuing a directive: ‘Match the code . . . ’

The Boy Detective Fails does indeed come with a decoder wheel that the reader can use. The code in this extract uses numbers to represent letters. The most obvious number-to-letter code would be to number the letters A–Z as 1–26. This has not quite happened here:
the number for each letter has been shifted upward by three places, so that A is represented by 4, B by 5, and so on. Decoded, the message here reads: *Billy/Why have you forgotten me?* In the context of the story, this message is still somewhat ambiguous. Nevertheless, it provides additional narrative content and, more importantly, necessitates the reader’s active physical engagement with the text.

**Further directions**

Future practice for research on creativity and multimodal literature should test the empirical validity of readers’ experience of such creative literary texts. For instance, Gibbons (2012: 183–97) undertook an empirical study of readers’ responses to an extract from multimodal fiction – namely, *Woman’s World* by Graham Rawle (2005). Gibbons divided her participants into two groups: one group read the extract in its original multimodal form, whilst another group of readers functioned as a control group, reading a conventional-looking mock-up of the extract. This allowed Gibbons to compare the different textual experiences of the two groups. The study also consisted of two parts. First, readers were asked to write notes on the extract, so that their instant online reactions in reading could be accessed. Secondly, readers were asked to answer a short questionnaire about their reading experience. The questionnaire asked research-driven questions, such as whether readers experienced mental imagery whilst reading the extract. Ultimately, the empirical study provided insight into readers’ reflections on the experiential dimension of reading multimodal literature, and showed that there is indeed a difference between reading multimodal text and reading more conventionally type-set texts. Further investigations should be conducted to extend and build upon Gibbons’ study.

Another empirical approach to multimodal literature that may produce useful insight is the use of eye-tracking studies in order to provide data on reading paths, as well as how readers attend to the different multimodal elements within texts. For instance, Holsanova, Holmberg, and Holmqvist (2009: 1224) conducted an eye-tracking study of newspaper layouts and discovered that ‘a difference in the spatial layout has a significant effect on readers’ behaviour’. Specifically, when text and images are clearly separated, readers ‘treat them as two independent units, and almost no integration occurs’, whereas when text and images are arranged in a more integrated format by means of closer positioning, reading itself appears more integrated because ‘it makes it easier for the reader to find the correspondences between referents in the text and in the illustration’ (Holsanova, Holmberg, & Holmqvist, 2009: 1224). An eye-tracking study of a piece of multimodal literature would show the reading path taken by a reader, as well as (perhaps) show which parts of the text were the most deviant and deautomatised.

**Related topics**

creativity and digital text; creativity and technology; language, creativity, and cognition; literary narrative; metaphor and metonymy; stylistics; vernacular creativity in urban textual landscapes

**Further reading**

This is the first book-length study devoted to exploring the literary-linguistic fabric of multimodal fiction and develops a cognitive-poetic approach to such visual texts.


This article explores conceptual metaphors in a range of multimodal literary works (poetry, fiction, and graphic novel) and, through stylistic analysis, argues that such metaphors are the creative products of a combination of verbal and visual modes, as well as readers’ performative engagements.


This book chapter presents a social semiotic stylistic analysis of Foer’s multimodal novel, paying particular attention to typography, layout, and photographic images.


In this book chapter, Nørgaard offers an excellent overview of work in stylistics and multimodality studies, focusing particularly on the social semiotic approach.

**References**


