As the texts readers encounter evolve from monomodal entities, dominated by written language, to multimodal ensembles, replete with visual images, sophisticated design elements and hypertextual features, the strategies readers draw upon to comprehend these complex texts will need to evolve as well (Serafini, in press; Unsworth, 2002). The competencies required of readers in today’s environment have expanded from cognitively-based reading comprehension skills (Pressley & Block, 2001) to include strategies for comprehending the visual images and design features presented within and across multimodal texts, for example postmodern picturebooks, magazines, news reports, and webpages (Serafini, 2011; Sipe & Pantaleo, 2008).

A multimodal ensemble is a text that draws on a variety or multiplicity of modes, for example painting, photography, written language, diagrams and visual design elements. Modes are socioculturally shaped resources for realizing, representing, interpreting and communicating meaning potentials in different ways and different contexts for different purposes. The meaning potentials realized through written language are different from photography, and photography offers different meaning potentials from sculpture and architecture.

Contemporary educators and literacy theorists have described a shift from the dominance of print-based texts that rely primarily on written language to the emergence of print and digitally-based, multimodal texts that contain visual images and other design elements (Anstey, 2002; Kress, 2003). In schools today, readers interact with traditional, print-based texts that contain multimodal elements, for example picturebooks, informational texts, magazines and newspapers, as well as digitally-based texts that contain visual images, hyper-text, video, music, and sound effects. Magazines, billboards, picturebooks, webpages, brochures, video games, advertisements and textbooks all contain a blend of visual images, design elements and written language. These digitally and print-based texts present challenges to novice readers as they work across multiple sign systems to construct meaning (Siegel, 2006).

As literacy educators working in schools today, we need to develop instructional approaches to help students navigate, interpret and analyze multimodal elements to ensure our students are capable of fully participating as literate citizens in our society, now and into the future (Serafini, 2009). This does not mean that we need to abandon the comprehension strategies focusing on print-based texts that have become an integral part of many classrooms, it means that we need to expand the repertoire of strategies we teach our students to address the changing dimensions of the texts they read and encounter.

As literacy educators working in schools today, we need to develop instructional approaches to help students navigate, interpret and analyze multimodal elements to ensure our students are capable of fully participating as literate citizens in our society, now and into the future (Serafini, 2009). The evolution of texts, from print-based, monomodal texts to digitally rendered, multimodal ensembles, has material, technological and social dimensions. Technologically, the most prominent shift in the past few decades has been from the medium of the page to the medium of the screen (Kress, 2003). As this shift has continued, the image has come to dominate written text and visual design elements have begun to play a more meaningful role in multimodal ensembles. As Kress (2003) asserts, we can no longer treat language as the primary means for representation or communication, and proficiency with written language alone cannot provide access to the meaning potentials of the multimodally constructed text. This assertion has important implications for literacy education. We as educators need to understand what a multimodal text is, how it is constituted and how mean-
As Kress (2003) asserts, we can no longer treat language as the primary means for representation or communication, and proficiency with written language alone cannot provide access to the meaning potentials of the multimodally constructed text.

In addition to the technological and material changes inherent in multimodal ensembles, we need to consider the potential shifts in power and cultural capital (Bourdieu, 1993) that occur between producers and consumers in a Web 2.0 environment (Lankshear & Knobel, 2006). All texts, whether print or digitally-based are social entities reflecting the interests of the producer and the social contexts in which they are produced (Kress & van Leeuwen, 2001). The meanings of multimodal texts are constantly shifting and responding to the dynamic social environments in which these texts are made and remade, reflecting the needs and interests of the producers and consumers of texts. Researchers and theorists working with multimodal texts are interested in how various elements and features of multimodal texts work together to provide cohesiveness and meaning potential for the reader (Baldry & Thibault, 2006; Jewitt, 2009). The question of what constitutes a multimodal text is not as important as the question of what multimodal texts can do for readers and writers, author-illustrators and publishers. How meaning is realized within and across different modes is an important aspect of recent studies in multimodality (Bourdieu, 1993; Machin, 2007). Texts that rely solely on written language are constrained by the temporal or sequential nature of written language (Kress, 2003). Visual images rely on spatial composition rather than temporal sequence. In other words, images contain objects and actors that must be realized in particular arrangements, where written text is not required to make such distinctions. The statement, the teacher is in the classroom, offers no specifics of where the teacher stands, it simply offers a state of being. A picture of a teacher in a classroom must place the teacher in a particular location. Visual Images must make different claims as to the location of the teacher beyond just the condition of being in the classroom, it must place the teacher in a specific location in the classroom. A picturebook spread that states the teacher is in the classroom offers the reader more information from a variety of modes than either illustrations or written text alone can offer. What the multimodal picturebook can do is offer spatial and temporal dimensions that allow the reader a fuller experience than written text alone. The meaning potentials offered through multimodal ensembles expand the possibilities of what single modes alone can offer.

In addition to expanding our understandings of multimodal texts, we have to develop a way of talking about the various features, structures and design elements contained in these texts. Without a theoretical and pedagogical framework and associated metalanguage or vocabulary for comprehending and analyzing multimodal ensembles, educators will struggle to prepare students to design and interpret these complex texts. A metalanguage refers to a set of terms for describing and analyzing a particular mode or system of meaning, for example photography or painting, and the various design and typographical features contained in multimodal ensembles. Before teachers can help support students as creators and interpreters of multimodal texts, they first have to become more familiar with these texts and develop a more extensive knowledge base from which to expand their literacy curriculum. In this article, I outline a series of con-
Considerations for teachers and literacy educators to support readers’ interpretation of visual images and multimodal ensembles. These considerations provide an initial metalanguage for addressing and discussing multimodal elements in contemporary picturebooks. I selected to use contemporary and postmodern picturebooks to illustrate these considerations because these texts are easily accessed by most classroom teachers, have been part of traditional literacy approaches and serve as a bridge between the texts of the past and those of the future. These considerations can also be used with digitally based texts as well, for example webpages and social media. With each consideration I offer some examples and some practical implications for classroom instructional practices.

**Considering Navigation**

In Freebody and Luke’s four resources model, readers were expected to draw on the four resources, namely, reader as decoder, meaning maker, user and analyst, to make sense of texts in new times (Freebody & Luke, 1990). As texts evolve, so do the strategies required for making sense of them. In considering the decoding of multimodal texts, it would serve literacy educators better if we expanded the concept of decoding to encompass how readers navigate texts, rather than simply decode them. The term reader as navigator is not new in research and discussions concerning hyper-text and on-line resources (Lawless & Schrader, 2008). When used in reference to reading print-based texts, the definition of the term navigator presented here subsumes several processes or abilities often associated with reading proficiency, such as decoding, concepts of print, directionality, and sequencing. However, when navigating multimodal texts readers are required to attend to the compositional arrangements of visual images and design features, in addition to the grammars and typography associated with written language.

For example, in navigating on-line resources, readers must understand how hypertextual links operate and how one’s reading path may no longer be a sequential, linear path, but may go in multiple directions depending on the choices made by the reader and their interests in the material presented. In fact, most websites are designed for visitors to have more choices in how they navigate each page than written language alone allows. Hyperlinks and choice buttons allow readers to move across as well as within pages, shifting the role of reader from follower to designer of the texts they navigate.

In much the same manner, postmodern picturebooks allow more choices for the reader, and require the reader to navigate the multimodal aspects of these texts in different ways from traditional picturebooks. Postmodern picturebooks challenge readers with metafictive and postmodern literary devices, for example multiple narrators, self-referentiality and non-linear sequences (Sipe & Pantaleo, 2008; Waugh, 1984). Picturebooks like *Voices in the Park* (Browne, 2001), *Black and White* (Macaulay, 1990), *The Three Pigs* (Wiesner, 2001), and *Wolves* (Gravett, 2006) require strategies for navigating the non-linear structures inherent in these texts, and must be navigated in unique ways in order to understand them. *Voices in the Park* (Browne, 2001) is a story told from four, unique perspectives, offering the reader four different narratives one after another with four different fonts used to demarcate each voice. Each of the four narrators tells their own version of a trip to a park, from their own perspectives, challenging the reader to make choices about the validity of each version. With no single, omniscient perspective offered, the multiple narrator structure requires readers to consider the perspectives of each character and contemplate the version of reality each character’s story represents.

Self-referentiality is a literary concept where the author, illustrator or publisher may use visual images, design elements and written text to break the secondary story world of the narrative and speak directly to the reader. In *The Stinky Cheeseman and Other Fairly Stupid Tales* (Scieszka, 1992) the title page falls in the middle of the book, and the red hen complains about her role in the story directly to the reader. These techniques call attention to the fictional nature of the picturebook being read. The reader is forced to forego their immersion in the story world and realize they are reading a fictionalized narrative. This playful technique keeps the reader from losing sight of the fact that they are reading a constructed textual artifact and secondary story world. In one particularly interesting image in *The Three Pigs* (Wiesner, 2001), one of the pig characters looks directly at the reader and states he thinks there is someone out there, referring to the world of the reader outside of the book. This comment breaks the wall between the fictional world of the picturebook and the actual world of the reader challenging the reader to consider the distinction between story world and reality.
In *Black and White* (Macaulay, 1990) the story is told in four sections or quadrants on each page. A warning is given on the title page that there may be four separate stories or only one, and the reader is advised to read very carefully. This groundbreaking work disrupted the traditional linear sequence of the narrative and requires the reader to read across and within each page to construct a story for him or herself. Simply decoding the words of these texts or expecting a linear narrative would leave the reader far short of the meaning potentials offered in the visual images and design elements of these unique picturebooks. Teachers need to demonstrate how to navigate through the structures of these texts, for example how to read multiple narratives and narratives that do not adhere to the traditional beginning, middle, and ending structures of traditional stories. Discussing how the four stories are presented in *Voices in the Park*, or helping students navigate the non-linear narrative structure of *Black and White* are important lessons for young readers (Serafini, 2009).

**Considering Paratextual Features**

Title pages, covers, endpapers, dedications, author notes, book jackets, advertisements, promotional materials, associated websites and such, make up what is known as a book’s paratextual features (Genette, 1997). The paratext is divided into those features that are contained within the covers of a book (peritext) and those that are outside the book itself (epitext). In many contemporary picturebooks, the endpapers and other paratextual features are used by the authors, artists and book designers to extend the narrative beyond the written text itself (Sipe & McGuire, 2006). By-passing these peritextual elements to skip to the beginning of the written narrative detracts from the reading experience and limits the meaning potential of these multimodal ensembles.

Spending time exploring the paratextual features before beginning to read the narrative itself expands readers’ experience of picturebooks and provides extended opportunities and resources for interpretation. An important strategy that can be shared with students during this exploration is talking about what is significant on the cover, back cover, title page and endpapers. For example, in the book *Once Upon an Ordinary School Day* (McNaughton, 2004) the endpapers in the front of the book show a boy with his eyes closed and birds sleeping. On the endpapers that close the book, the boy is soaring with the birds suggesting a metamorphosis has taken place in the boy’s outlook on school. Another example is the red hat presented on the back cover, title page and book jacket of the book *Voices in the Park* (Browne, 2001). We may not be sure what the hat symbolizes before we begin reading the book, but by noticing its presence throughout the peritext, we can be sure that it is not a literal image, but refers to some aspect of the story. This is what is meant by the term determining significance (Serafini, 2009). Calling students’ attention to the symbols and motifs presented on the endpapers and other paratextual features is an important lesson for approaching a multimodal text.

In today’s digital environment, there are numerous epertextual resources associated with most published, print-based books. Author websites, publisher resources, fan fiction writing sites, advertisements, on-line book sharing sites (e.g. librarything.com and goodreads.com) and commercial sites like Amazon.com all offer reviews, critiques, information and resources that can extend and enhance one’s reading experience. Readers can now read a book, publish a review, read other reviews, watch a book trailer, read extended versions of stories and share their favorite books with other readers from their classroom or homes through the internet. Taking advantage of these resources requires teachers to become more familiar with them and allow students time to explore them on their own.

**Considering Typography**

The typography and design elements associated with written language have taken on new forms and new roles in contemporary publications (van Leeuwen, 2006). Many aspect of typography have been naturalized to the point of being invisible (Iedema, 2003) allowing readers to focus past the typographical representation to the meaning potentials offered in the written narrative. Rather than acting as a conduit for the communication of a verbal narrative, fonts have become an integral part of the story itself, a meaningful resource that adds to the potential meanings of the picturebook. Written language is presented through a particular typography, and design elements are included to add to the cohesive and coherent nature of a picturebook (Nodelman, 1988).

Some typographical features teachers can call students attention to in-
Chester has gained power as the editor of Watt’s original story. Chester is shown holding a marker in several illustrations as if he is literally taking a red pen to Watt’s story. Students need to recognize that Chester’s comments are offered from a voice distinct from the original author to understand this wonderful book.

In the book Crazy Hair (Gaiman, 2009) the designer has chosen to use color in a very unique way on the cover. The word crazy is written in bright colors: green, blue, red, and orange. Below the word crazy, the word hair is rendered in green, pink, yellow and blue. Rather than being vibrant, the colors of the word hair are muted, darkened and neutral. The contrast between the word crazy and the word hair makes the vibrancy of the colors used for the font of crazy even more obvious, making it the most prominent element on the cover.

The use of color in Chester (Watt, 2009) is a key element in how the multiple narratives are told in this hilarious picturebook. Everywhere in the book where the character of Chester writes, his comments are rendered in bright red marker, overlaid on the story written by the actual author Melanie Watt. Chester’s disruption of the narrative is presented through the use of color (red) and the playful written script. This vying for power between the author and the character of Chester is the foundation of the humor in this book.

Often, red is interpreted as a symbol of power and confidence. The use of a red font in this case suggests Chester has gained power as the editor of Watt’s original story. Chester is shown holding a marker in several illustrations as if he is literally taking a red pen to Watt’s story. Students need to recognize that Chester’s comments are offered from a voice distinct from the original author to understand this wonderful book.

**Considering Framing and Borders**

The two most common types of illustration included in picturebooks are the full bleed illustration, where the illustration extends to the edges of the page, and bordered images where white space or black lines are used to frame the illustration. Full bleed illustrations, especially when presented across two pages or double page spread, are more intimate presentations, allowing the viewer to feel immersed in the scene, drawing the reader into the world depicted. Border images create more distance between the viewer and the scene, and offer the viewer a window onto the scene providing a more detached view of the events and characters. It is like looking through a window onto the secondary world of the story.

In addition, Illustrators have used the borders of an image as a narrative element of the story. For example, in the book Freefall (Wiesner, 1988), the opening spread has a border on three sides, leaving the right edge open to suggest a continuation of the illustrations throughout the rest of the book. The borders on the ensuing pages are left open on the left and right edges until the final page where the right edge is bordered suggesting a closing to the narrative. Wiesner masterfully uses borders in a different way in his award-winning book Flotsam (Wiesner, 2006). He uses borders like comic book illustrators do to suggest a series of sequential events (McCloud, 1994). What happens between the sequentially bordered images is just as important as what happens within the images. Varying lengths of time and changes in setting are suggested in the gutter between the sequentially rendered images.

*Once Upon A Cool Motorcycle Dude (O’Malley, 2005)* is an excellent example of the use of a formal versus an informal font. The title of the book is written using two fonts, Old English and a standard sans serif font. Old English, often associated with calligraphy, is much more traditional and formal than the sans serif used in the second part of the title. Throughout the body of the book, the first half of the story is written from the perspective of a little girl, using a comic sans font, a very round and curvy font used mostly in informal writing. The motorcycle dude’s story is rendered in a rounded and curved font, but it features a heavier weighted font, suggesting a more powerful voice than the female voice in the book. These are just a few examples of the way fonts play a role in the interpretation of stories. The fonts selected for websites, magazines and brochures signal different meanings and are worth exploring as a design element of multimodal ensembles.
In the book *Animalia* (Base, 1986), the author-illustrator breaks the borders of images to create a more dramatic effect, as animals seem to leap out of the images into the world of the reader. In *Wolves* (Gravett, 2006) the author-illustrator breaks the borders of the images as well as the distinction between the story world and the world outside the book, blurring the distinction between narrative and reality. In this book, multiple story worlds are visually depicted as the rabbit in the story is duped by a wolf into checking out a book on wolves from the library, only to be pursued by a wolf from another level of the story. The wolf resides in the story read by the rabbit and the world the fictional rabbit inhabits. A similar technique is used in *The Three Pigs* (Wiesner, 2001), to navigate the pigs in and out of different stories and story worlds in their attempt to find their way home. In these instances, the borders of the visual images have evolved beyond a simple design element, to being a semiotic resource readers and illustrators use to convey and construct meanings. The borders of various images included in contemporary picturebooks become part of the narrative itself, offering potential meanings to those readers willing to consider them as a semiotic resource.

**Considering Symbols and Motifs**

Many illustrators use common and uncommon objects to symbolize meaning potentials beyond the literal presentation of the object. Many common symbols have been used throughout society to connote specific and general meanings; a rose for passion, the color green for envy, a human skull for death, and a cross for religious faith. Literal objects and figurative meanings are connected through specific uses in the cultures in which they are adopted. There is no direct or universal connection between an object and its meaning, rather symbols and metaphors offer meaning potentials to readers and author-illustrators through their use and contexts (Lakoff & Johnson, 2003).

Many visual images used in picturebooks offer meaning potential beyond the literal object depicted. Browne’s use of the red hat in *Voices in the Park* may represent the overbearing nature of Charles’ mother. Browne also uses bars and street lamps as symbols throughout his books to offer more sophisticated meaning through his illustrations (Serafini, 2009). Some of these may be readily identified, like the pigs used to represent male chauvinism in *Piggybook* (Browne, 1986), while others are more subtle and require some investigation, like his use of Frans Hals’ painting *The Laughing Cavalier* in the background of the same book (Serafini, 2010).

**Concluding Remarks**

As multimodal texts become the norm rather than the exception in today’s schools and society in general, educators need to expand their own knowledge to support students’ ability to design, interpret and utilize multimodal texts in a variety of settings. We need to accept the evolution of the texts our students encounter, and can no longer hide our heads in the sand and focus our literacy instructional practices exclusively on decoding written language. We should embrace the multimodal ensembles and pop cultural artifacts our students’ experience and expand our own knowledge base concerning literacy and multimodality if we expect to expand the literate lives of our students.

In contemporary classrooms, we need to develop a metalanguage for discussing the visual and design elements of multimodal ensembles, to call students’ attention to the meaning potentials offered in these semiotic resources. Our instructional approaches need to expand beyond traditional comprehension strategies to include strategies for making sense of the texts experienced by our students. Expanding our students’ interpretive repertoires beyond strategies for written language is essential as the world told becomes the world shown (Kress, 2003).

---

**FRANK SERAFINI** is an author, illustrator, photographer, educator, musician, and an Associate Professor of Literacy Education and Children’s Literature at Arizona State University. In addition, Frank was an elementary school teacher and literacy specialist for twelve years in Phoenix, AZ.

Frank received his Masters degree in Elementary Education and his Ph.D in Reading Education at Arizona State University. In 2003, he received the Distinguished New Faculty Award, and in 2007 the Outstanding Teacher Award from the University of Nevada, Las Vegas. In 2009, Frank received a Foundation Award at ASU for his Excellence in Teaching.

Frank has published six books with Heinemann, including: *The Reading Workshop, Reading Aloud and Beyond, Lessons in Comprehension, Around the Reading Workshop in 180 Days* and *Classroom Reading Assessments*. In addition, Frank has recently published Interactive Comprehension Strategies: Fostering Meaningful Talk About Texts with Scholastic. Frank is currently working on a project with Teachers College Press focusing on Visual Literacies.
NOTES
Color Cover image from *The Three Pigs* by David Wiesner with permission from Clarion Books
Color Cover image from *Crazy Hair* by Neil Gaiman & Dave McKean with permission by HarperCollins Publishers
Color Cover image from *Once Upon a Cool Motorcycle Dude* by Kevin O’Malley with permission by Bloomsbury/Walker Books for Young Readers

REFERENCES


