

Chapter 2

When image, text and other modes meet

Developing a metalanguage for understanding multimodal texts





Swimming in the ocean and finding yourself in or near a shoal of fish is an exhilarating but also slightly unnerving experience. In a flash, hundreds of colourful silvery bodies twist and turn, catching the light. Sometimes they move as one school, then divide into two before regrouping. And suddenly they have gone. If you are in the water, you catch only very brief glimpses of individual fish, who move as if one larger being. A photograph (Figure 2.1) catches a moment of the action, where eyes gaze at us, bodies are caught mid-turn before they dart away. We observe an array of individuals as well as the whole, even as it morphs and changes before our eyes.

Daily life is often an immersion into hundreds of sights and sounds as well as movement, even before we engage with a screen. As we learn to socialise from birth, we begin to understand oral language as well as gestures, movement, facial expressions, tone and volume as having different meanings. Adding screens to these already busy social contexts, our highly multimodal culture makes us adept at seeing many images and texts across our day. We may decide to focus on some while skipping across others. When we choose to look closely at an image, what types of 'tools' do we use to make sense of what we see? As educators, we are particularly interested in the types of meanings we make as readers and viewers. As well, having a language to talk about visual images – a metalanguage – means we can assist our learners to develop thoughtful and informed readings of all types of images. This flows into how we might also become creators of multimodal ensembles, as children and as adults.

Figure 2.1



Figure 2.2

Engaging with multimodal texts

Have you ever leapt into deep water from the end of a wharf? In Figure 2.2 from *Jetty jumping* (Rowe & Sommerville, 2021), we can see how Milla’s friends propel themselves with confidence into the water below, while a nervous Milla watches. Hannah Sommerville’s detailed drawings show how they each jump differently as their bodies twirl, twist and spin into the water. Even the shape of the written text reflects their movements. Each character draws our gaze, and we interpret the actions, even without having read the written text. The immediacy of images draws us in, as we take in a page or screen, even before we have time to reflect on what we see

Context shapes how we make meaning with all types of texts. David Wiesner’s wonderful and mysterious book *Flotsam* begins with a young boy finding an antique camera washed up on the beach (similar to the one in Figure 2.3). The adventures across time in the story fascinate all readers although young children reading the book are often confused when the main character takes photographic film out of the camera and has to wait a whole hour while the photos are developed!

Written in 2006, on the cusp of digital photography becoming the dominant way we record images, the story reminds us of the need to think about what contextual aspects are present in the books and texts we choose to share in our classrooms, and how we might introduce and scaffold such knowledge as



Figure 2.3

part of reading and viewing. We know the value of providing a rich range of quality literature and texts (McDonald, 2023), choosing books that engage, challenge and support each learner in our class. Similarly, we need to understand the context of our own learners – the knowledge they bring to a story, the skills they will need to access the text, and the social and cultural knowledge we may need to teach them as part of the reading and viewing process.

Context will impact engagement with any text – are we teaching a whole-class lesson, leading small-group discussions, hosting a family movie night or reading a bedtime story? Having chosen, and carefully read and considered the features of a text, we should then read or view the whole story with our students (or a meaningful part if it is a longer text) before we go back into it to learn about a specific feature or quality. Consider a teacher reading *Jetty jumping*. They may know that many students in their class don't know how to swim, and have limited experience of bays with wharves and jetties. So, the teacher may include time to discuss the cover (Figure 2.4) and find out how students might feel about jumping into deep water. An interactive reading – balancing reading with children's comments and teacher think-alouds – will begin to build students' enjoyment and knowledge of the story. The literary and visual features or specific reading skills can then be explored in subsequent activities with the book. Experienced teachers understand that literacy learning should be recursive, where knowledge and curriculum content should be introduced, revised and/or assessed across the weeks and terms of the school year.

When we select any story or text for our classroom, we may not initially use terminology such as 'cultural context', 'structural features' or 'grammatical elements'. However, as educators, this information is part of our professional knowledge that is embodied in curriculum content and then developed into appropriate literacy practices. As such, it's very useful to bring this knowledge to the fore when we examine the features of the texts we are choosing, particularly in terms of the multimodal dimensions.

When we are looking through a new picture book, we might see the front and back covers, and then read the book. We could then give a general review of, and response to, the overall story, sharing our personal views and opinions about it. We might do this verbally with colleagues, or more formally, as with the review below:

Jetty jumping is a delightful story of summer and swimming, reminding me of my own childhood. Milla's friends love diving and jumping off the end of the jetty while Milla watches, scared of the deep water. A lost bracelet leads to Milla conquering her fears. This warm-hearted story, with rich descriptions, is complemented by the bright pen-and-watercolour images. The storyline and the themes of swimming, friendship and bravery would suit ages five to seven.



Figure 2.4

Table 2.1 Textual features that inform teaching and learning

Genre and styles	Literary features including themes, vocabulary	Patterns across written text	Patterns across visuals	Multimodal features
<p>Text form: Picture book.</p> <p>Genre: Narrative with a clear orientation, complication and resolution.</p> <p>Illustrative style: The illustrations are light and delicate with a realistic yet whimsical feel.</p>	<p>Themes of swimming, friendship, overcoming fear, bravery.</p> <p>Richly descriptive noun groups and adjectives. Highly varied action verbs.</p> <p>Use of alliteration, e.g. 'scary seaweed', 'diamond drops'.</p>	<p>Short clauses with varied verb type repeated e.g. 'Bonny <i>twirls</i>', 'Clancy <i>corkscrews</i>'.</p> <p>Variety of circumstances, e.g. 'in Milla's town', 'on the edge', 'past the pylons'.</p>	<p>Bright watercolours with strong movement lines as the children jump.</p> <p>Some high- and low-angle pictures suggest fear of depth of water.</p>	<p>The pictures in this story work together to build similar meanings with some images enhancing the words, such as a high-angle shot showing a small Milla looking into the deep, blue water. The written text uses words such as 'mysterious', 'dark' and 'scary'.</p>

Going more deeply into the text, we might look more closely at specific features that could inform our teaching and learning (see Table 2.1).

At the word level, teachers can also investigate the text for word recognition opportunities, phonic knowledge and spelling patterns (see the PETAA unit of work by Sophie Honeybourne, 2022).

Using a social semiotic framework

Building on our initial observations, we can use a social semiotic framework to look more closely at how different meanings are made across each mode. Let's focus on Figure 2.2. We can ask three types of questions:

- What's happening?**

Milla, seated on the edge of the jetty, watches her friends as they jump enthusiastically into the water below.
- How do we interact with, and relate to, the image?**

We are viewing the scene from above, almost from a bird's-eye view. The jetty looks quite high, as the children drop into the bluey green water. Milla seems to look longingly at her friends as she twists her bracelet.
- How do design and layout build meaning?**

The children, twisting and twirling, draw our eye and are grouped together, while Milla is noticeably by herself on the right-hand side, framed by the safety of the jetty.

The written text can be explored the same way. Reading the whole book, we understand it to be part of a narrative genre or text type, set in a seaside location. The simple use of the

conjunction 'but' on this page signals the first complication in the narrative.

As with the visual image, we can ask three similar questions about the written text:

- **What's happening?**
The written text is mostly about actions – Milla's friends *twirl*, *dive*, *horsey* and *corkscrew*, but Milla just *dangles* her legs, *bites* her lip and *twists* her bracelet. This all happens *out on the end of the towering jetty*.
- **How do we interact with, and react to, the text?**
Using statements to tell about the events, the varied verbs build energy and excitement for the friends, while other evaluative language such as *dangles*, *bites* and the description of *the towering jetty* convey anxiety and worry for Milla.
- **How is the text organised and structured?**
Each short clause begins with a character's name, focusing our attention on them and what they are doing. The incomplete sentence 'But not Milla ...' introduces a contrasting conjunction and the ellipses leave the reader to decide what she is feeling.

Drawing out these three aspects from the image and text gives a clear introduction to some of the basic principles of a social semiotic framework. Both written and spoken language, as well as visual images allow us to convey ideas, feelings, thoughts and concepts. As forms of communication, they are functional – they allow us to get the work of communication done. This is not to exclude the aesthetic pleasure and enjoyment of engaging with texts, something

the semiotician Roland Barthes acknowledged and is central to this book (Barthes, 1990). Understanding how texts function should complement our enjoyment and understanding. The linguistic work of Michael Halliday (Halliday & Matthiessen, 2004) suggests that for each particular text we engage with (written, visual or multimodal), we can ask about three areas, referred to as register:

- **Field** – what is happening and what the topic or subject matter of the image or text is
- **Tenor** – who is interacting (speaker/listener, reader/viewer, writer/composer) and the roles and relationship developed by the text
- **Mode** – how the text is structured and logically organised according to the purpose and medium.

Using register as a starting point, we can then look in more detail at what specific semiotic resources are used to describe field, tenor and mode (see Table 2.2). These meaning-making functions or metafunctions can be applied to spoken, written and visual texts (as well as sound, music and movement) and provide us with a rich and detailed way of understanding, discussing and appreciating all types of multimodal ensembles.

Engaging with the text as a multimodal ensemble is a crucial factor. If we analyse the visual and the written text separately, we are likely to lose the combination of image and text that creates various levels of meaning. Table 2.3 shows how we can bring together an analysis of both the written and visual elements of a page, as well as take aspects of the audience, purpose and text type into consideration.

Table 2.2 Semiotic resources used to describe register

Register	Metafunction	Meaning-making focus
FIELD	What's happening?	Resources for expressing actions and ideas, presenting characters or participants and showing circumstances using written, visual, spoken, aural and gestural modes
TENOR	How do we interact and relate?	Resources for interacting with others, showing feelings, attitudes, credibility and power relationships using written, visual, spoken, aural and gestural modes
MODE	How do design and layout build meaning?	Resources for organising logical and cohesive texts using written, visual, spoken, aural and gestural modes

Table 2.3 Analysis of written and visual elements of Figure 2.2

Jetty jumping is a picture book for young children, drawing on **cultural knowledge** about swimming and diving off a jetty. Written for the **social purpose** of entertainment and enjoyment, the **narrative text type** presents a simple complication and resolution, as Milla finds herself conquering her fear of jumping. The story develops the broader themes of bravery, overcoming fear of the unknown and friendship. The **specific context** of this page is the complication where we find Milla unable to join her friends as they play.

VISUAL	MULTIMODAL MEANINGS	WRITTEN
<p>What's happening?</p> <p>Milla is seated on the edge of the jetty, watching as her friends do all types of jumps into the water. Their body shape and visible action lines show the friends' movement.</p>	<p>How do image and text work together to create meaning?</p> <p>The ensemble of words and pictures here show a fearful Milla, along with a group of her energetic friends, jumping and spinning into the water, their bodies twisting, accentuated with visible action lines (vectors). Looking from above at a high angle, Milla appears small and vulnerable, sitting on the right while all her friends are in close proximity on the left. Compared to the busy description of the friends, Milla is left to <i>dangle</i> her legs and <i>bite</i> her lip.</p>	<p>What's happening?</p> <p>The written text uses a range of actions verbs – Milla's friends <i>twirl</i>, <i>dive</i>, <i>horsey</i> and <i>corkscrew</i>, but Milla just <i>dangles</i> her legs, <i>bites</i> her lip and <i>twists</i> her bracelet. The circumstances happen <i>out on the end of the towering jetty</i>.</p>
<p>How do we interact with and relate to the image?</p> <p>We are viewing the scene from above, almost from a bird's-eye view. The jetty looks quite high, as the children drop into the bluey green water, making it seem exciting. Milla seems to look longingly at her friends as she twists her bracelet.</p>	<p>How do image and text work together to create meaning?</p> <p>The ensemble of words and pictures here show a fearful Milla, along with a group of her energetic friends, jumping and spinning into the water, their bodies twisting, accentuated with visible action lines (vectors). Looking from above at a high angle, Milla appears small and vulnerable, sitting on the right while all her friends are in close proximity on the left. Compared to the busy description of the friends, Milla is left to <i>dangle</i> her legs and <i>bite</i> her lip.</p>	<p>How do we interact with and react to the text?</p> <p>Using statements to tell about the events, the varied verbs build energy and excitement for the friends, while other evaluative language, such as <i>dangles</i>, <i>bites</i> and the description of <i>the towering jetty</i>, convey anxiety and worry for Milla.</p>
<p>How do design and layout build meaning?</p> <p>The children, twisting and twirling, draw our eye and are grouped together, while Milla is noticeably by herself on the right-hand side, framed by the safety of the jetty.</p>	<p>How do image and text work together to create meaning?</p> <p>The jetty, described as <i>towering</i>, is adorned with shells and seaweed. It is quite salient on the right, and the legs of the jetty lead our eye to Milla, as does her placement on the corner of it.</p>	<p>How is the text organised and structured?</p> <p>Each short clause begins with a character's name, focusing our attention on them and what they are doing. The incomplete sentence 'But not Milla ...' introduces a contrasting conjunction and the ellipses leave the reader to decide what she is feeling.</p>

YOUR
TURN

Look at the image and text in Figure 2.5. Use the questions in the table to think about the various ways meaning is constructed.

Table 2.4 Analysis of written and visual elements of Figure 2.5

Context and purpose – What context could this photo be from? A friend’s social media post or a child’s story? Does the language sound more like a formal narrative or a spoken comment? Does the photo seem too professional to be just a personal photo?.

	Visual	Written
Happenings	What actions can you see? Are there any vectors that show fast movement? Is the white dog looking in a certain direction?	How do you know what the two dogs are doing? Find the verbs that make up the three clauses. Are there any details that show where or how they are running?
Interactions	Are we positioned at eye level or are we lower than each dog? Does it make them look weak or strong?	Are the clauses statements or questions? What words build a sense of the dogs’ personalities?
Design and layout	What attracts your attention first in the photo?	What is the focus of the written text, based on what is first in each clause?



Figure 2.5

Different modes making meaning in different ways

Words have an exactness that differs from images. In 1989, the UN adopted the United Nations Convention on the Rights of the Child (UNCRC). Each of the 54 articles were carefully written with attention to the formal and legal details needed. In Figure 2.6, we can see an abridged version of the rights, with an illustration of each article.



Use the QR code or visit: www.savethechildren.org.au/our-stories/rights-of-the-child



Figure 2.6

Figure 2.7



If we look at each article, some of them are clear examples of where the written mode and the illustration are similar in meaning. Figure 2.7 shows someone giving money to a family, reflecting a similar meaning to the written text.

Figure 2.8



While the written text in article 31 uses the term 'a wide range of activities', the illustration shows a child playing a guitar (Figure 2.8). Here the image complements the written text, enhancing the meaning with a specific example.

Figure 2.9



Article 16 states that children have a right to privacy, yet the illustration (Figure 2.9) shows a child without privacy, surrounded by cameras. Here the illustration uses a contrasting image to make the point of the written text.

A significant point here is that in a multimodal text, different modes have different affordances or potentials for meaning-making. Each mode won't express exactly the same meanings, which of course reflects the diversity and creativity we have when communicating. The inclusion of music, audio or movement adds further richness as well as complexity. More detailed descriptions of what are termed intermodal relations allow for more focused analysis (Mills & Unsworth, 2017; Painter & Martin, 2011). For the purposes of this text, the following three terms are useful for reflecting on multimodal meanings:

- **Similar meanings:** Each mode in a text may communicate similar meanings across the multimodal ensemble.
- **Complementary meanings:** Each mode may add extra meanings, enhancing another mode in the overall ensemble.
- **Contrasting meanings:** Some modes may present a contrasting or even contradictory meaning as part of the overall purpose of a multimodal ensemble.

As with the Rights of the Child poster, sometimes an image may represent similar, complementary or contrasting meanings. Written text can, of course, elaborate with more detail, using various linguistic structures and forms, but the meanings that images create are never just the exact linguistic equivalent. The visual mode and other modes such as music, speech and performance create meanings that inform, move and challenge us, in very powerful and sublime ways. Theorist Nicholas Mirzoeff argues that 'concentrating solely on linguistic meaning, such readings deny the very element that makes visual imagery of all kinds distinct from texts, that is to say, its sensual immediacy. This is not at all the same thing as simplicity but there is an undeniable impact on first sight that a written text cannot replicate' (Mirzoeff, 1999, p. 15). Or as artist Edward Hopper said, 'If you could say it in words, there'd be no reason to paint'.

Table 2.5 Summary of multimodal semiotic features

VISUAL	MULTIMODAL MEANINGS	WRITTEN
<p>Happenings</p> <p>Action qualities – how line and shape represent people, objects, events and places Vectors (visible or inferred) showing action, expression, gestures, speech and thought, cycles and processes Setting – detailed or limited settings Conceptual and symbolic qualities – images that may classify, analyse or symbolise. Forms include graphs, timelines, diagrams, cross sections and infographics.</p>	<p>How do image and text work together to create meaning?</p> <p>Do they present:</p> <ul style="list-style-type: none"> • similar meanings? • some complementary meanings? • contrasting meanings? <p>What meanings are dominant: image, text or a balance?</p>	<p>Happenings</p> <p>Noun groups and adjectives represent participants, people, places, things and ideas Verbs show actions and ideas Adverbs and adverbials show the manner, surroundings and circumstances of actions Clauses express a message, using verbs, nouns and other word groups Nominalisation builds abstract concepts from verbs, e.g. evaporate > evaporation</p>
<p>Interacting and relating</p> <p>Gaze to viewer and point of view – demand viewer attention or offer observation; viewing from outside the text or inside (with character or over their shoulder) Shot distance – close, mid or long shot suggests level of intimacy with viewer Angles – high, low or eye level; frontal or oblique. Each suggests a power relation with viewers Gaze, proximity and angle between participants – reflects relationship between each participant Colours – create moods, reflect symbolism Authenticity and colour – suggests natural, scientific or abstract codes</p>	<p>How do image and text shape our feelings and interactions?</p> <p>Do the images and written text create similar or different affective responses?</p> <p>How are feelings, power and personal connection created through image and text?</p> <p>How does the use of colour in the image complement or contrast the written text meanings?</p>	<p>Interacting and relating</p> <p>Interacting – using statements, questions, commands, offers and exclamations Evaluative vocabulary – expresses affect between or about characters, offering appreciation or judgement Literary language – using alliteration, repetition, rhyme, rhythm, metaphor etc. to create mood and feeling Modality – degrees of certainty, probability using modal verbs, adverbials or adjectives</p>
<p>Design and layout</p> <p>Salience – what feature attracts a viewer's attention? Reading paths – movement of gaze around page, often directed by salience or vectors Placement – proximity, symmetry, balance and grouping of different elements Layout – top/bottom, left/right, centre/margin positions of each element Framing – extent to which text and image frame each other</p>	<p>How does the layout of the image and text guide the reader/viewer?</p> <p>How might the layout privilege particular visual or written elements?</p> <p>How might the meaning of the written text complement or contrast the visual placement and design?</p>	<p>Text design</p> <p>Text, paragraph and sentence openers – introductions, topic sentences, sentence themes Reference – flow of ideas using nouns and pronouns Text connectives – joining ideas across paragraphs and sentences Cohesion – sets of associated vocabulary and terms Clause themes – choice of clause beginnings using noun groups, adverbials or verbs</p>

Table 2.5, with its columns, headings and questions, will be the organising framework for each chapter. While developing knowledge about written grammar is not the main focus of this book, it clearly plays a central role in understanding multimodal texts. The texts *Grammar and meaning* by Humphrey, Droga and Feez (2012) and *A new grammar companion for teachers* by Derewianka (2022) provide an extensive introduction to written grammar. Across this text, there will also be some discussion of how other modes, such as video and music, play a role in multimodal texts.

Making meaning beyond the page or screen

Reading and viewing is not simply about decoding words or images. Words and pictures build patterns across sentences, pages and books, working together to create meaning. Books and stories are created as part of different cultures and contexts, appealing to different audiences by using various genres. Being mindful of these features can enrich and extend the way we teach using all types of text. We choose books for our classroom differently when we have a diverse group of students who arrive at school with English as an additional language. We work to extend students who are enthusiastic readers while at the same time bringing in texts that will engage our reluctant readers.

Reflect on how you might use fairytales in your classroom. If your students have had very little exposure to the Western canon of fairytales, you may explain (in age-appropriate ways) where these stories are from, and why they have been shared down through the ages. For older readers, fairytales are often used for satire or parody. Reading the classic picture book *The stinky cheese man and other fairly stupid tales* (Scieszka & Smith, 1993) still requires students to know the original fairy story characters, as well as understand the devices that Jon Scieszka and Lane Smith use to entertain and engage us. Bringing this level of text knowledge shapes how you might use each text in your classroom.

When we draw attention to the various techniques that are used across texts, we are working at the genre and text level, rather than at the specific word or image level. Words

and pictures of course make up the creation of different literary devices, from personification, motif and metaphor, to metafictional devices that use both image and text. Metafictional devices draw attention to the constructed nature of the text and include self-reference to the book being read, narrators who address the reader directly, multiple views or narratives, varied viewpoints, contrasting images or words and interactive formats (Pantaleo, 2014). The use of metafictional devices in picture books has developed from early seminal works such as *Come away from the water, Shirley* (Birmingham, 1992), *Black and white* (Macaulay, 1990) and *Tuesday* (Wiesner, 1991) to an abundance of more recent works such as the *Chester* series (Watt, 2008), *Interrupting chicken* (Stein, 2010), the *Journey Trilogy* (*Journey, Quest, and Beyond*; Becker, 2013, 2014, 2017) and *Do not open this book* (Lee & McKenzie, 2016).

Picture books are not the only type of multimodal text children read. Classrooms should be sites of engagement with all types of ideas and information, conveyed in a variety of ways. There are many types of texts, ranging from print to moving image and interactive electronic texts, that may also be part of performances or exhibitions. Information texts, whether screen based or paper based, are a significant part of contemporary life. Many contain images and text as well as video and audio resources, where interactivity and text creation also play a role.

Changing times and changing literacy demands

Arguably, the changing nature of technology has been more rapid over the past 15 years compared to previous eras. Yet, teachers understand the importance of engaging with new technologies, while not forsaking practices and texts that experience and research show to be vital for all learners. Fostering reading for pleasure, for example, has always been a foundation for literacy teaching, and as Mackey argues, there is 'no reason why reading for pleasure should not be digital (or non-fictional for that matter)' (Mackey, 2019, p. 118).

Text-rich, supportive classrooms provide the foundation for fostering deep engagement with all types of texts, where we complement and build on important letter/sound knowledge and

comprehension to extend into 'deeper elements of purpose, pleasure, satisfaction and connectedness' (Beavis, 2020, p. 65).

At the same time, simply filling our classrooms with screens without being purposeful about their use will not support deeper engagement or learning (McLean, 2020). Researchers looking at the new literacies that are part of our literacy landscape encourage us to consider what makes things 'new' in technical terms as well as the ethos that underpins new technologies, which is often very collaborative in nature (Knobel & Lankshear, 2014). If we have access to a range of multimodal texts, we can start to work out ways that all of our students can use them across a day or week, as well as what texts best suit some learners for interest, skills and need.

Choosing and teaching with multimodal texts

The following key principles for planning are further developed in Chapter 6.

Who are your learners? In the context of curriculum topics and shared lesson plans, find out who your students are, consider their interests and expertise as well as assessing their needs. Choose the types of texts that will engage, surprise, model, inform and challenge your learners.

Collect and curate resources Use your professional knowledge and expertise to decide what texts and teaching strategies will work for your students – books, posters, screens, live and virtual guests, puppets or performance.

Read, view and enjoy Engage your learners with dramatic reading, a balance of think-alouds (not too many questions on the first reading) as well as time for them to turn and talk about what they think.

Modelled, guided and independent learning Choose clear and engaging examples from a text to revisit, teaching a skill or concept that will enhance the meaning of the text, as well as develop your students' skills and enjoyment.

A range of publications provide explicit guidance about how to teach using various texts. These include *The potential of the visual: Teaching literacy with multimodal texts* (Asha, 2022), *Exploring how texts work* (Derewianka, 2020) and *Teaching poetry for pleasure and purpose* (Murphy, 2021).

Information texts in print and on screen

Information texts can be read for pleasure and enjoyment, as well as forming part of the curriculum learning areas beyond English. Building understanding about the purpose and features of different types of informative texts means teachers need to have some organising principles about how these multimodal ensembles are designed. For example, while narratives mainly rely on a sequenced flow, many informational books and apps present readers and viewers with multiple entry points and focuses (Callow, 2010). More recent ebooks and apps include multiple media and interactive features

including augmented reality (Tan & Chik, 2022). While there are many creative and engaging designs used in print and digital texts, there are also some common features of pages or screens that present factual information, from science and geography to history and health.

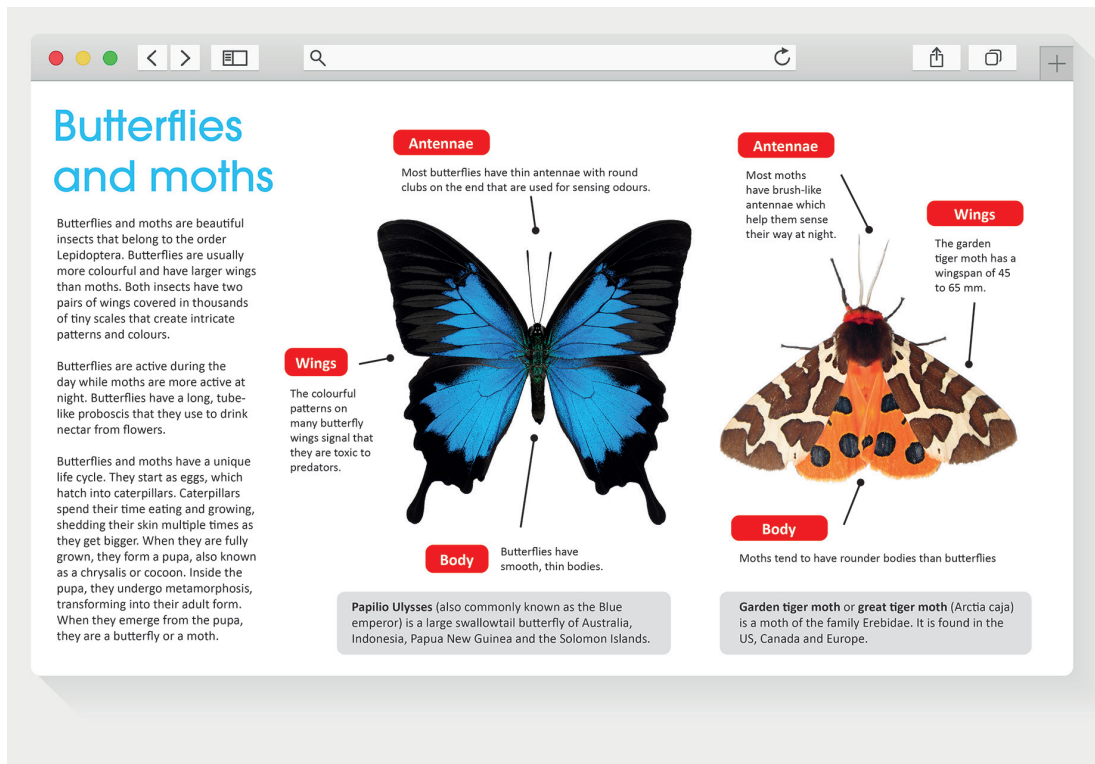
Butterflies and moths

As part of a science unit, a teacher might conduct a modelled reading lesson using a website, chosen with the age and content of the class in mind. As an example, see Figure 2.10. Under the heading of 'Butterflies and moths', the larger written text first classifies butterflies and moths as 'beautiful insects that belong to the order Lepidoptera'. Written as a short descriptive report, the written text describes butterflies and moths by colour, physical characteristics, behaviour and reproduction. Compared to the narrative in picture books, the written text is much more technical and dense.

Visually, the two labelled photographs take up about two thirds of the screen, making them stand out. The labels, when hovered over with the mouse, give extra information. The design for many textbooks and digital texts often means that readers are visually offered many entry points when engaging with the page or screen. Apart from the literacy demands of these choices, there is also research and discussion about the type of pedagogy and learning that these layouts may promote (Coleman et al., 2018).

As the written text and image become more complex, reading and viewing paths become similarly complex and often quite challenging. The nature of the text and associated tasks will play a principal role here. If this is a core text that a class is working with, the teacher should read the introduction and discuss what the scientific classification of animals means, as well as check the visual glossary of how each page and section is designed.

Figure 2.10



Interactive digital texts

Digital-screen-based texts have evolved to include more interactive elements that not only involve video and sound, but also touch-based or haptic elements (McLean, 2020). While every app will be different, teachers need to select and explore the apps for classroom use, not only for content but also to consider what features and challenges may be present for their learners.

Mammals (by Tinybop) is an app with a number of features for teachers to consider before using it in their classroom. In the settings there are over 40 languages that can be selected, which changes all labels and audio description to the chosen language. The main screen shows five mammals with a menu bar on the left.

Figure 2.11 (below)

Figure 2.12 (bottom)



Users select two animals who race, which shows their relative speeds.



A viewfinder using the camera shows users what each animal sees with their type of vision.



Each animal's covering is shown, and touching the screen shows whether it is flexible (like kangaroo fur) or stiff (like elephant hide).

Exploring an elephant

Information is sourced by observing and drawing inferences from what happens when you interact with each animal. For example, when you choose one of the five mammals, the screen shows a menu bar with icons for different body systems such as the muscular, cardiovascular and reproductive systems (Figure 2.13). The elephant's digestive system requires users to choose different foods (meat, frog, vegetables, fruit etc.) and drag them to the elephant's mouth. The elephant will either spit out the food or chew and swallow it. The animated food then progresses through the digestive system. Users discover that the elephant requires quite a lot of vegetables and fruit to fill its stomach, which then leads to the production of elephant dung, complete with sound effects.

Apart from needing to understand what the icons mean and what you need to touch, tap or drag, the important skills here are higher-order inference. What does it mean when the elephant always spits out meat and insects? What happens to the food as it passes through each of the labelled organs? This type of multimodal text requires teachers to model and discuss what is seen, touched and heard, and to then supplement this with other scientific concepts and facts about the animal. The other screens in the app have similar interactive features.

Figure 2.13

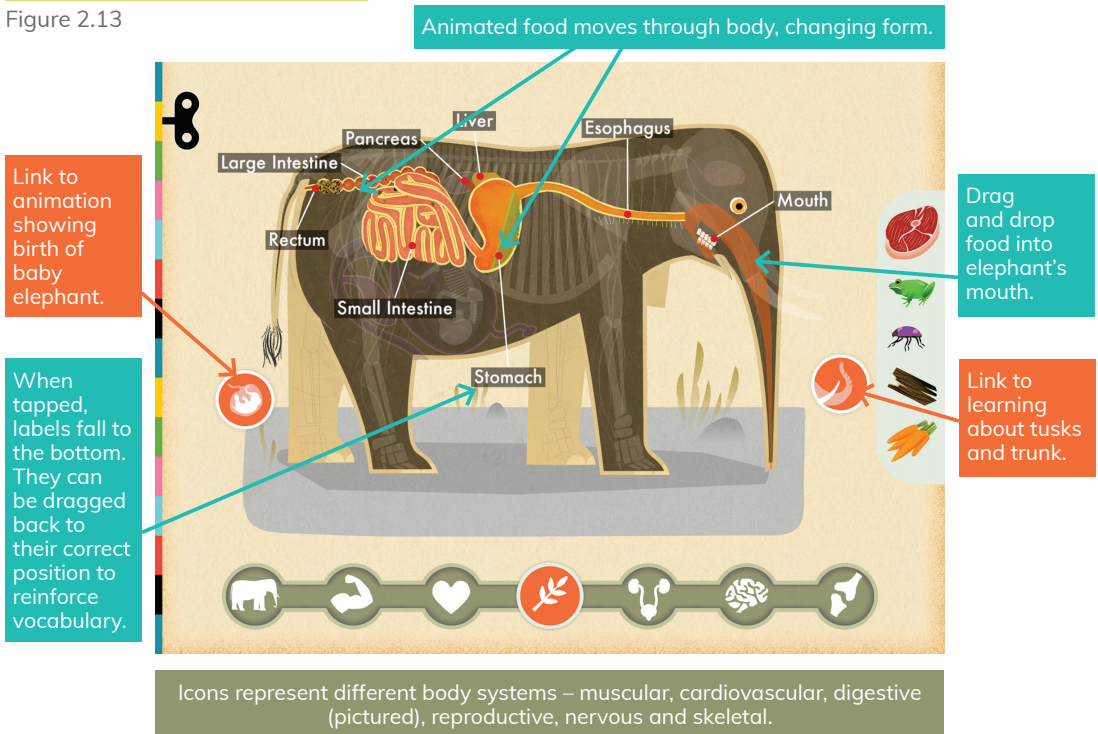


Table 2.6 Analysis of written and visual elements of Figure 2.13

Mammals by Tinybop		
An interactive tablet-based text written for children aged five to nine. It presents information using interactions with body systems as well as animations. This screen allows users to feed the elephant and watch the digestive process, where you need to infer that elephants only eat fruit and vegetables, and that they must eat a great deal before their stomach is full.		
VISUAL	MULTIMODAL MEANINGS	WRITTEN
<p>Happenings</p> <p>The elephant illustrations are conceptual images, which present analytical views of their (external and internal) physical features, using simple labels. Food is pictured using small images. The animations show the processes of eating and digestion using chewing and swallowing food, then the food moving as smaller and smaller pieces.</p>	<p>How do image and text work together to create meaning?</p> <p>This interactive screen is quite complex. Users need to read the labels, following the line to each feature. As they use touch to feed the elephant, they need to deduce what foods the elephant will eat, then consider how much food is needed and what the process of digestion entails. Ideally, thinking out the concepts aloud would bring both modes together, e.g. 'When the elephant eats fruit, the fruit goes down the oesophagus to the stomach ...'</p>	<p>Happenings</p> <p>The written text is a set of noun groups, which can be dragged and dropped to the bottom of the screen by tapping. They label parts of the digestive system but not the actions involved such as chew, grind, pass through, expel etc.</p>
<p>Interacting and relating</p> <p>The elephant is shown in sideview in long shot, creating a more detached, objective feel. The images are muted in colour tone and are more illustrative or diagram-like, rather than naturalistic or photographic.</p>	<p>How do image and text shape our feelings and interactions?</p> <p>The visuals are appealing and simple, with the interactive features of sound, touch and animation adding to the emotive engagement. The written text is scientific and, teamed with sound effects and animation, creates an exploratory, hands-on learning experience.</p>	<p>Interacting and relating</p> <p>The labelled features use correct scientific terms.</p>
<p>Design and layout</p> <p>The elephant is central and the most salient element. The menu bars on the right and bottom frame the main image, while the labels add another layer to the main image. The haptic or touch features allow food to be dragged to the mouth. The animation guides the eye around the screen.</p>	<p>How does the layout of the image and text guide the reader/viewer?</p> <p>The visual and haptic features require some initial exploration to discover what to do. The combined actions and animations create a sequence that guides the users' eyes and hands to start the digestive processes.</p>	<p>Text design</p> <p>The noun groups are not part of sentences or clauses.</p>

Conclusion

This chapter has introduced a social semiotic framework, presenting an overview of the main concepts and metalanguage associated with the model. The sample texts from a picture book and interactive screens show key features that allow teachers to think more deeply about how multimodal ensembles work. The next

chapters of this book will look at the three meaning-making functions, or metafunctions, (happenings, interacting and relating, and design and layout). Each chapter ends with two annotated examples of different types of multimodal texts in order to revisit the key concepts.