Introduction

This chapter starts from the premise that all people have learned, very early in life, how to understand moving-image media, meaning any moving-image material, from short sequences in apps, games, and advertisements to full-length television programmes and films. It may seem strange to insist that viewers have had to learn to understand these media. From its earliest days in the late 19th century, the movie industry has promoted itself as the most ‘lifelike’ medium, and much of the extensive literature on children and moving-image media takes this for granted. The famous – and probably exaggerated (Loiperdinger, 2004) – accounts of audiences recoiling or screaming when in 1896 or soon afterwards they saw the Lumière brothers’ 50-second movie L’arrivée d’un train en gare de La Ciotat formed, as Loiperdinger argues, the ‘founding myth’ of moving-image media: that viewers are supposed to instinctively believe they are watching ‘real life’. But moving-image media are now a distinctive, long-established art form that has been developing for more than a century. They are full of rhetorical devices that do not reproduce daily perceptual experiences: for example, jump-cuts, parallel montage, shot/reverse-shot sequences, non-diegetic sound. These terms may be familiar only to film buffs and movie professionals, but moving-image audiences can effortlessly ‘read’ the features they refer to: if they could not, they would not be enjoying what they watch! Because these audiences do not remember learning to interpret these features, they think that they never had to.

But when does this learning happen? Broadcasters and film-makers know that three- and four-year-olds can be expected to follow and enjoy many feature films and programmes that older age groups like as well. If evidence of the prior learning that has enabled them to do this is sought, it is necessary to look at younger children. This chapter proposes that two-year-olds’ engagement with moving-image media is a socially and culturally important process of learning how to understand the medium itself, alongside – and inherently bound up with – their efforts to follow narratives. This approach challenges two preoccupations that currently dominate academic research about children and moving-image media. First, there is an overriding concern with the risk that these media may present a threat to children’s well-being. Second, most research on children and media in the 21st century has concentrated on digital technologies rather than on the different types of content that these technologies enable viewers to consume and create (Carlsson, 2010; Lankshear & Knobel, 2008; Livingstone & Sefton-Green, 2016).
Research on Young Children and Media

Most Anglophone research in this area takes place within an inherently anticipatory ‘risks and benefits’ paradigm that seeks to identify the psychological, cognitive, and social effects of these media on children’s later development. Within this paradigm, scholars either emphasise the supposed risks of ‘too much’ viewing (Palmer, 2006; Vandewater, Bickham, Cummings, Wartella, & Rideout, 2005; Zimmerman, Christakis, & Meltzoff, 2007), or they argue for its benefits. These may relate to language acquisition or print literacy (Kendeou, Bohn-Gettler, White, & Broek, 2008; Lemish & Rice, 1986; Marsh, 2000; Robinson & Turnbull, 2005), or they may point to the educational or sociocultural value of media content (Davies, 1989; Lauricella, Gola, & Calvert, 2011; Marsh et al., 2005).

Very little research takes an informed approach to the specificities of moving-image media. For example, the stylistic devices media makers employ to convey meanings; their densely multimodal structures; their enormous stylistic and generic variation, both historically and globally; the fact that they are not merely visual but also aural; and that, as in music, movies’ management of sound track duration, rhythm, and pace is an essential dimension of meaning-making. The tendency has been to regard television as a visual (rather than audiovisual) medium, whose defining features reside in the technology – in particular, the screen – rather than in the institutional or aesthetic features that distinguish, for example, different genres and intended audiences (e.g., Anderson & Hanson, 2010; Gola & Calvert, 2011). In addition, critics make little reference to children’s own interests in re-viewing material, and discussions of ‘response’ and ‘attention’ focus largely on gaze, with little consideration of features such as bodily tension, posture, gesture (apart from pointing) and choice of position in relation to the screen (Pempek, Kirkorian, Richards, Anderson, & Lund, 2010).

The period around the third year of life (approximately 18–40 months) is an immensely important time for learning. This period – known in Anglophone cultures as the ‘terrible twos’ but in German and Danish as the ‘age of autonomy’ – is when children learn to speak, become much more mobile and dexterous, and make huge progress in understanding their social and cultural surroundings. However, as Rowe and others indicate, ethnographic studies of two-year-olds are rare (2008), and rarer still in research on children and moving-image media. This gap in the research was noted by Collins as long ago as 1979, but has persisted: Lealand (1998) points out that “among the thousands of research studies and policy statements on children and television, viewers under five years old are usually underrepresented and often ignored” despite general acceptance of the idea that the early years are the most formative (p. 4). Such studies as there are inevitably involve difficult “practical and logistical considerations, including gaining access, involving children as active research participants and negotiating consents” (Plowman & Stevenson, 2013, p. 330). One way of overcoming these difficulties is for researchers to study their own children or grandchildren. A few well-known studies of this nature have resulted in key insights on toddlers’ learning (Britton, 1970; Darwin, 1877; Edmiston, 2008; Halliday, 1975; Piaget, 1928; Weir, 1970).

Research that has depended on information from parents or carers (Certain & Kahn, 2002; Schmidt, Rich, Rifas-Shiman, Oken, & Taveras, 2009; Vandewater et al., 2005), or that has only gathered data through periodic and relatively short visits to homes – the problems of which are discussed by Jordan (2006) – cannot claim to provide confident interpretations of the idiosyncratic ways in which individual children, especially those under three, express their responses to the moving-image media they engage with; or to capture the subtleties of the social contexts in which much viewing may take place. Dafna Lemish’s 1987 paper, “Viewers in diapers” stands out as an important study of this type, given that it followed 16 children for six to eight months, covering an age range from 6.5 to 36 months, aiming “to discover and describe the process
through which babies become television consumers” (1987, p. 54). However, each family was visited only four or five times, and the research depended as heavily on parental viewing logs and diaries as it did on researcher observation. The children’s viewing was also dominated by Sesame Street, the long-running programme from the US-based Children’s Television Workshop. Lemish herself points out that the study “is only one step towards a better understanding of young children’s television viewing” (p. 56). Many changes in moving-image media, technologies, and family life have occurred since 1987, and it is unfortunate that few researchers since then have attempted to undertake similar studies of this age group or to use more intensive observational methods. Two notable exceptions in the UK context are Briggs’ study of his own family’s relationships with the TV series Teletubbies (2006) and Robinson and Turnbull’s (2005) case study of a child’s literacy development – including her relationship with films and TV – from birth to age six; but again, neither addresses the question of how these children learned to understand the medium itself.

The author’s own interest in addressing this question was originally motivated by her professional experiences in developing education about moving-image media, which included persuading teachers that children’s ability to make inferences and predictions about narratives, to recognise generic features, and to appreciate and enjoy stories is enhanced by their film and television viewing. Primary school teachers who overcome their worries about institutional or parental disapproval and start teaching about moving-image media in the classroom are almost invariably amazed by the apparent transformation of their pupils’ knowledge and abilities that results. They tend to infer that the moving-image media have acted as a trigger or accelerator for children’s learning, whereas what has arguably happened is that teacher training and the school system, in general, have failed to recognise or value the prior learning about moving-image media that children bring with them to nursery and reception classes.

This chapter draws on a 20-month observational study of the author’s own grandchildren (Connie and Alfie: dizygotic twins), focussing in particular on the 22–36-month phase (Bazalgette, 2018). Through more frequent (at least weekly) extended and closer contact and familiar surroundings (the author’s home and theirs, separated by only 30 minutes’ travel time), this study collected more material than visiting researchers usually achieve. Research data amounted to over 12 hours of video (taken on an unobtrusive and familiar device: an iPhone), 90 sets of observation notes and over eight hours of parental interviews. A grounded theory approach to analysis of these data (Charmaz, 2006; Glaser, 1967) led to the author drawing upon insights from the broad and evolving field of embodied cognition (Coegnarts & Kravanja, 2015; Damasio, 2000; Daum, Somerville, & Prinz, 2009; Gallese & Sinigaglia, 2011; Trevarthen & Aitken, 2001). In this process, the researcher was in many ways a participant observer who often had to slip between the roles of observer and grandmother. As with other studies of children by family members, the findings may be somewhat subjective due to familiarity with the participants, but they may still provide insights with important implications for future debate and research. The potential ethical issues in the research were discussed with the family, and the project was considered and approved by the relevant university ethics committee.

The Social Context of Viewing

Parents who enjoy moving-image media are likely to share this enjoyment with their children, often starting by the time their babies are three months old (Marsh et al., 2005). Many parents soon find that leaving the baby to watch a television programme or DVD enables them to get on with essential household tasks, even though they may feel anxious in case the baby watches for ‘too long’ (Blum-Ross & Livingstone, 2018).
Thus babies quickly become accustomed to their family’s moving-image media viewing practices: the devices family members use, the times and places where they watch, and how they watch. Ofcom’s 2018 report *Children and Parents: Media Use and Attitudes* provided a snapshot, from the middle of the century’s second decade, of how UK children’s viewing habits were changing, which included some data on three- to four-year-olds. They found that, although viewing broadcast programmes on a TV set was in overall decline, 96% of three- to four-year-olds watched TV on a TV set for an average of 14 hours per week, while 30% watched moving-image media on other devices, mainly on a tablet, much of which would have entailed watching YouTube for animated movies, funny videos, or pranks. While most of these figures increased substantially for older age groups, it was still the case that watching movies on a TV set, although declining slowly but steadily overall, remained an important activity for three-year-olds. Judging by anecdotal evidence on social media, and by the digital practices described by Bar Lev, Elias, and Levy (2018), it is possible to infer that two-year-olds and even infants also play frequently with portable devices such as smartphones and tablets and that this is likely to include opportunities to view moving-image media.

Modes of watching will differ between or even within families. Some family members may be doing other things while they watch, while others may sit and watch intently from beginning to end; some may comment frequently to each other, while others may maintain an attentive silence. Some families will have their televisions on continually for much of the day, others will not. Some audiovisual phenomena will not be entirely new to those neonates who have already heard many theme tunes, sound effects, and audience reactions while they were still in the womb (Johansson, Wedenberg, & Westin, 1992).

### How Two-Year-Olds Watch Moving-Image Media

The author’s research indicated that two-year-olds’ engagements with moving-image media follow a similar pattern to their other daily activities. There may be a lot of ‘milling about’ as they explore the spaces they inhabit and search for things that will be interesting to investigate. When they do find something promising, they will give it absolute attention for as long as they can. A movie playing on a flat-screen television may be one of the things in their environment, but they will only pay it attention when they want to, often stimulated by sound-track elements that attract them. Their attention may also be stimulated by a co-viewer, as in “ooh, look at that!” but only if a glance at the screen confirms that it is more interesting than what they are already doing. In other words, like most other things that toddlers do, their attentiveness – and their learning – is largely self-directed, and to interfere with this can result in screams of rage: hence the parental frustrations that have given rise to the unfortunate label ‘terrible twos’.

A close look at how rapt attention in a two-year-old manifests itself reveals an enormous investment of energy. Gazing at a screen is different from, say, investigating the cutlery drawer, in that the hands may not be so active – but they are used to grip on to something if, as often happens, the child is standing up to watch the screen. A toddler’s centre of gravity is higher than an adult’s (Huelke, 1998); if they want to pay attention to a large screen they will want to get as close to it as they can and may need to ‘brace’ themselves against a piece of furniture in order to maintain a steady gaze and follow on-screen movement by moving their heads. Alternatively, if they are not quite so close to the screen, they may simply leave their hands where they were when their attention was first seized, and maintain almost complete stillness as they watch.

### Pleasure, Re-Viewing, and ‘Using It Up’

An initial viewing of a media text may well not reveal pleasure. The child may frown, chew her cheek, purse her lips, and perhaps grip a table or chair tightly as she watches. The toddler’s...
perennial runny nose may ensure that her mouth stays open, so she will periodically lick her lips and perhaps hastily wipe her nose on her sleeve, maintaining her gaze at the same time. If she is drinking from a bottle or cup she will hold it to one side as she does so, in order not to miss anything on the screen. If she sees, hears, or even anticipates anything alarming in the movie, her bodily tension will increase: shown perhaps through clenched fists, raised shoulders, and deeper breathing.

When a child has bestowed this much attention on a movie, she is likely to ask for ‘more’ as soon as it finishes. She may excitedly anticipate the bits she remembers, point to the screen and shout out the name of the character or thing that is going to appear, touch the screen to identify an object of interest, grin knowingly as an action that she understands well plays out as expected; and her pleasures here will be enhanced when co-viewers respond appreciatively. By considering a child’s intense – and sometimes maddening – desire to constantly re-view a selected movie as a learning process, one may be better able to respect the child’s choices. And when she no longer wants to re-view it, this is not necessarily because she is ‘bored’ or ‘fed up’ with it but because she has used it up: she has extracted all she can from it and is ready to move on to the next thing.

The twins often gave their greatest attention, and most demands for re-viewing, to moving-image media that not only appealed to them but were also ‘at the edge’ of their ability to understand. For example, when they were 28 months old they shifted their attention from short and relatively simple TV episodes such as the eight-minute Baby Jake (Darrall Maqueen Ltd/JAM Media, CBeebies, 2011–2012), which appeals to toddlers, to Tree Fu Tom (CBeebies/Fremantle Media/Blue-Zoo Production, 2012–2016), a 28-minute narrative programme aimed at ‘up to five year olds’ featuring live-action and CGI animation and a complex mix of arthropod, vegetable, and human characters. This programme was considerably more demanding than anything they had watched before, and yet the toddlers maintained their interest in it for five months. Extended periods of re-viewing or series loyalty, then, related to material whose complexity took a long time for them to ‘use up’.

An Embodied Cognition Perspective on Movie-Watching

Conventional ideas about two-year-olds’ movie-watching fall broadly into three modes of thought. First, there is the mode signalled by using the language of affect, as in “he just loves In the Night Garden” or “Peppa Pig is her absolute favourite”. Second, the use of pop-psychology language reveals, even if only light-heartedly, anxieties about the risks of movie-watching, as in “she’s completely obsessed by Waybuloo” or “he’s really addicted to Paw Patrol”. Third, there is the perspective adopted by many broadcasters and production companies, as in “if children aren’t able to follow the story, they won’t be interested” (Steemers, 2010, pp. 127–30). All three modes focus on supposed deficits in the two-year-old’s brain.

Embodied cognition offers a different approach to two-year-olds’ movie-watching. A broad field that draws upon many academic disciplines including neurosciences, evolution, and philosophy (Shapiro, 2012), this perspective challenges cognitivist approaches to learning and Cartesian dualism’s separation between mind and body. Instead, it proposes that cognition (in animals as well as in humans) is intricately bound up with the body’s motor and perceptual systems and how they interact with the physical world and with other creatures. It is a reminder that the extraordinary brains of homo sapiens evolved in dangerous and demanding environments and that people still carry the instinctive behaviour and modes of thought that were essential to group survival in those environments (Panksepp, 2004). It thus enables a fresh approach to the study of very young children and their “entry into the sociocultural world” (Trevarthen & Aitken, 2001, p. 20), leaving room for the hypothesis that babies and toddlers, confronted by the intense
multimodality of moving-image media, are far from being baffled but will often rise immediately to the challenge of figuring out what the texts they are watching mean.

**Emotion and Cognition**

The power of moving-image media to affect people emotionally has a long history and is an intrinsic element of the wariness with which Anglophone culture has treated the medium. Given the widespread assumption that emotions are something that toddlers have to learn to bring under control, it is understandable how anxieties about the risks of ‘exposure’ to moving-image media have arisen and why they have focussed particularly on the assumed vulnerabilities of young children (AAP, 1999).

According to neuroscience, emotions are vitally important systems that help to motivate both thought and action and are deeply embedded in people’s brains. ‘Primitive’ emotions like fear and rage may have been essential in the environmentally dangerous lives and close social inter-dependence of early humans, but what Panksepp calls the ‘seeking’ emotion was just as important. It generates anticipation and investigation, and by helping individuals perceive causal connections, enables the formation of ideas (Panksepp, 2004, pp. 144–9). An important insight into the role of emotions in social interactions comes from research into mirror neurons and how these enable interpretation of and response to the actions and motivations of others (Gallese, 2001). Gallese explains how the mirror neuron system not only impels people to imitate the actions of others (like smiling back when smiled at) but also functions as an important basis for empathy.

Bearing these arguments in mind, when observing two-year-olds viewing and re-viewing a movie, it is clearer how their intense attentiveness has an emotional force that is more than simply ‘enjoying it’. In seeking meaning, very young children are not only learning how to interpret the expressions and actions of characters but also how to assemble causal connections and, therefore, how to follow a narrative. This can be inferred in several different ways.

**Expectations of Significance**

Much of toddlers’ efforts to understand moving-image media are self-driven. Watching TV is not like shared reading of picture books; it is not mediated by an adult or caregiver. The multimodal density of moving-image media offers many ‘ways in’ to potential meanings that the child can latch on to by herself. But co-viewers can still mediate, if in different ways. Through their engagements with any media – books, pictures, computer games, radio, mobile phone calls and texts, music, and of course moving-image media – adults and older siblings unconsciously communicate the importance they ascribe to media content. Lancaster’s (2001) analysis of a two-year-old making drawings with her father describes how a shared activity communicates what she calls “an expectation of significance about the semiotic objects encountered” (p. 136). The two-year-old doesn’t, at this point, learn how to draw a cat that looks like a cat, but she does discover the importance of drawing as a mode of communication and understands that mark-making can be meaningful. Shared viewing of moving-image media contributes to a child’s determination to crack the codes of moving-image media because she knows that moving-image media are important to her co-viewers.

Co-viewers’ contributions to the movie-watching environment may be deliberate and even didactic, as in “Oh, look at the poor cat! He’s sad, isn’t he?”, etc. More often, they are unconscious. When adults watch moving-image media with little children, especially if it’s something they hope the children will like, their spontaneous exclamations such as “oh!” or “wow!” or “uh-oh!” may cue responses from the children. But even the comments adults may make to each
other during the movie, or the bodily tensions or chuckles that may be felt by a child sitting on an adult’s lap, can act as clues to meaning for the less experienced child viewer.

**Diakresis: Selecting the Salient**

Moving-image media use images, movement, colour, voice, sound effects, music, and the duration, transitions, and juxtaposition of shots in constructing and conveying meaning. How does anyone, let alone a two-year-old, take all this in at once? Dehaene (2014) claims that even adults can really only deal with one thought at a time because their consciousness imposes a narrow bottleneck on the multiple perceptions that their daily experiences—including moving-image media—present to them. Drawing on Dehaene, Wojciechowski suggests that in watching a movie, “we perform continual acts of diakresis—a separating out of information that is salient enough to enter into our conscious awareness, and the distinguishing of the salient from everything else” (Wojciechowski, 2015, p. 124). The more extensive the movie-watching experience, the more skilled children are at drawing on their experiences of narratives and genres in the use of diakresis. As two-year-olds repeat-view, they pile up their diakretic selections until they have assembled enough of a coherent whole to satisfy their desire for significance. And because diakresis is a highly individual process, it is difficult for adults to predict just what an individual child may find interesting, or frightening, or baffling.

The beginnings of diakresis can be observed in very young children, starting with the apparently irrational terror that can be sparked off by innocuous material in children’s moving-image media; a phenomenon that is often discussed on social media.1 The diakresis theory would suggest that in these cases, the salience of the disturbing element of the scene overrides what is intended to be gentle and amusing programme content. This corresponds with Kagan’s observation that one-year-olds “are sensitive to events that appear to contrast with those which adults have indicated are proper” like broken toys or clothing that’s damaged or stained (1981, pp. 47–8). But in movie-watching, it is often the narrative that establishes what has ‘gone wrong’, not the family’s social rules: so these odd fears can be a sign that a child actually is beginning to follow stories.

A later example of diakresis at work appeared when the twins were watching *Finding Nemo* (Stanton, 2003) for the first time, aged just three. Connie was unable to ‘correctly’ interpret the death of Nemo’s mother Coral; not just because the actual death is not shown, but also because she could not countenance it emotionally. Instead, she constructed what was, for her, an alternative narrative enigma—Coral has disappeared somewhere, and they have to find her. Connie held on to the narrative until the end of the movie when she was surprised to discover that Coral had simply been replaced by Dory. She had yet to get hold of the generic knowledge and the concomitant awareness of convention that would enable her to accept that Coral was dead and gone, but she was using diakresis to identify a situation she found to be emotionally salient and helpful in constructing her own narrative expectations.

**Modality Judgements: ‘Real’ and ‘Pretend’**

Hodge and Tripp’s concept of ‘modality judgements’ (Hodge & Tripp, 1986) is taken directly from linguistics and does not equate with the later connotations of ‘modality’ within multimodality theory. It simply refers to what the assumed truth or reality status of a movie is supposed to be, thus making the important assertion that the appearances of ‘truth’ and ‘reality’ can be highly questionable or at least uncertain. In the context of this chapter’s argument, modality judgements can also be usefully linked to the ways in which children operate their rules about ‘real’ and ‘pretend’ in play; these can be arbitrary but hold good for the duration of the game. Anxieties about the risks of movie-watching often include the mantra that “children cannot distinguish between
fantasy and reality” but, as Woolley points out, children are not fundamentally different from adults in their ability to distinguish fantasy from reality, and that everyone operates “a continuum of ontological commitment to what we think the world is really like” (Woolley, 1997, p. 991): the ‘fake news’ phenomenon provides ample demonstration of this.

Two-year-olds are as interested as anyone in what is or is not meant to be real but, being small and relatively vulnerable, their threshold for fearful reactions is set sensibly lower than that of larger, stronger, and more experienced people, and they are more likely to be frightened by characters or events that appear to claim a high modality status; in other words, that seem to be meant to be real. For example, Alfie (at 34 months) remained uncertain for some time about whether there really was a ‘big bad mouse’ in The Gruffalo’s Child (Welland & Heidschtter, 2011), exorcising this fear through many games of running through the woodland in the local park with Connie, waving a stick and screaming “Monsters!” Thinking about and negotiating the modality status of moving-image media, especially where this involves potentially frightening things, is thus an important part of learning how to understand moving-image media and, by extension, narratives in general. It is important to recognise that two-year-olds can, to some extent, do this for themselves.

**Conclusion**

The argument of this chapter is founded on the principles, first, that moving-image media are not ‘transparent’ and instantly accessible to the novice viewer, but instead employ complex, multimodal devices to convey meanings, and second, that these multimodal devices have to be learned, usually very early in life. These principles distinguish the research reported here from themes that have long dominated the study of children and moving-image media, such as the over-arching concern with the risks and benefits of children’s ‘exposure’ to audiovisual material, and the growing tendency to focus on technology – digital media for example – rather than on the specificities of different types of content, and how and why children engage with them.

The research findings on which this chapter is based are drawn from the author’s 20-month study of her twin grandchildren’s viewing practices between the ages of 22 and 42 months, focussing particularly on when they were two: an age range relatively neglected in academic research, given the difficulties of access and methodology that it presents. Two aspects of their viewing behaviour were described: social contexts such as family viewing, in which co-viewers’ utterances and physical behaviour formed part of the children’s experience; and the phenomenon of focussed attention, which provided evidence of the enormous investment of energy that children can make when a film or television programme engages their interest. By recognising two-year-olds’ “expectations of significance” (Lancaster, 2001) it is possible to see their attentiveness and demands for re-viewing as evidence of learning processes at work, rather than as idiosyncratic desires. Their changing preferences about what they watch may often be grounded in the desire to move on to more challenging material rather than merely ‘becoming bored’.

Insights drawn from the developing field of embodied cognition help to illuminate these findings further. Children’s instinctive emotional responses to what they watch may be the starting point for how they interpret and reflect upon a scene: identifying aspects that they judge to be salient, for example, and trying to assess how ‘true’ or ‘real’ a character or setting is intended to be.

Moving-image media have been an important part of human culture for more than a century. They are now made and shared in ways that, only 20 years ago, most commentators could not imagine. An appreciation that learning to understand these media is a significant achievement for most two-year-olds may enable a deeper understanding of children’s personal, cultural, and social development.
Cary Bazalgette

Note


References


