

7. Audible Comics

This chapter explores the role of sound in comics and provides a critical analysis of how that role has been impacted by the digital mediation of the form. Groensteen asserts that traditional comic books are monosensory, relying solely upon the sense of sight (2013, 69). In contrast to this he identifies digital comics that incorporate elements of audible sound as being ‘plurisensory’ (ibid). During my work as a comics practitioner I have created a number of plurisensory digital comics that incorporate elements of audible sound. This chapter will examine the format of the “audible comic” by drawing examples from both my own practice and the work of contemporary digital comic practitioners. The examination will be aided by a case study of *The Empty Kingdom* (Goodbrey 2014), a new game comic hybrid created as a practice-based inquiry into the incorporation of audible sound with the form of comics.

During the chapter sound will be considered as an element of both navigation and narrative. The differences between ‘imagined’ and ‘perceived’ sound will be outlined (Hague 2014, 65) and the relationship between sound and diegesis in comics will be explored in detail. Parallels and influences from the use of sound in film and videogames will be examined. Included in this examination will be a consideration of the ways in which audible sound can influence the act of reading and the implications of this for digital comic formats and the form of comics. Different approaches to the integration of audible sound will be explored, with an emphasis placed on the importance of reader control over the play and synchronisation of audible elements.

Sound in navigation

Groensteen’s monosensory view of traditional comic books is challenged by Hague (2014), who asserts that the experience of reading a comic book has on some level always been a plurisensory one. Hague notes that the act of reading a comic book is

not silent and that 'the sound of a turning page emphasises the "objectness" of the comic' (72). He also states that although 'they could be classed as incidental, these sounds comprise elements of the comic's character, they tell the reader certain things about the progression of the text and the modification of the comic as an object' (65).

In the earliest of my own experiments with sound in digital comics, my approach was to treat audible sound much as Hague outlines above. In the hypercomic *Doodleflak* (Goodbrey 2002), the reader is presented with a zoomed out view of the entire comic, arranged in the pattern of a snowflake. They can then click on individual branches of the flake, which initiates an animated zoom and rotation that allows them to read their chosen section of the comic. Audible sound is used as part of this navigation process, with one sound playing to accompany the act of zooming in, and another sound playing when the reader chooses to zoom back out. Just as animation has been used as a replacement for the motion of the page turn found in a typical paper-based comic, so these effects serve as a replacement for the sound of the page turn. This approach therefore treats audible sound as an element of the digital comic's character and uses it to give information about the progression and modification of the "virtual" comic object.

Sound in narrative

In Chapter Three, word and image blending was identified as a key characteristic of the form of comics. In this respect Cohn describes comics as being multimodal as they are 'essentially written in two languages: the visual and the written/verbal' (2013, 13). One approach to incorporating audible sound into a digital comic is to treat that sound as an element of the comic's narrative. This approach is enabled by the multimodal nature of the comics form. Within the multimodal narrative of a paper-based comic, sound is represented by images and words on the page. From these visual elements the reader constructs sounds that are 'imagined rather than perceived' (Hague 2014, 65).

Smolderen attributes the origins of this multimodality to a period of experimental 'graphic hybridization' (2014, 47) that took place during the nineteenth century. The result of this experimentation is a hybrid form that operates as 'an audiovisual stage on paper' (ibid). Smolderen asserts that it is this hybridity in comics that provides the opportunity for further hybridisation with other forms (60). In a digital comic, the potential exists to extend the multimodality of the form to also incorporate elements of perceived audible sound. Sounds that become part of the diegesis of the story and 'are consciously integrated into the work to supplement or even facilitate the narrative' (Hague 2014, 73).

To examine the role that audible sound might play in the narrative of a digital comic, it is helpful to first examine the concept of diegetic sound in more detail. In film, diegetic sounds are sounds presented as originating from a source within the story world of the narrative, while non-diegetic sounds are sounds that come from a source outside the world of the story (Bordwell & Thompson 2013, 284). These same definitions can also be applied to comics. In *Doodleflak*, the use of audible sound as an element of navigation is clearly non-diegetic, as the sound has no clear origin within the story world of the comic. In contrast, the visual representation of imagined sound on the page can at times demonstrate a more complex relationship to the diegesis.

Limited compositional space means that comic creators are selective about which sounds they chose to represent on the page. Usually, only those sounds that directly serve an aspect of the comic's narrative or aesthetic are displayed. Lacassin describes some of the typical approaches taken to the visual representation of sound, observing the existence of:

a high level of phonetic realism. A car... [emits] ...a "roooaaar" or a "vroom vroom". A trotting horse makes a "clippity clop". These graphic onomatopoeias ripple, explode or wind their way along inside the frame, traversing it in every direction (2014, 40).

In addition to these visual effects, a common device used to represent sound within the narrative is the balloon. Balloons are used to convey the spoken words or sounds of the various characters and objects that inhabit a comic's story world. The style in which a balloon is illustrated may be used to convey further information about the qualities and origin of the sound it contains. Lacassin provides some examples of commonly used effects, describing how 'a dotted outline... [indicates] ...whispered asides; saw-toothed edges indicate a retransmitted voice (by telephone or radio) or a recorded voice. A word balloon adorned with stalactites signals the icy tone of hostility' (2014, 39).

Eisner describes the balloon as a 'desperation device' for artists as it 'attempts to capture and make visible an ethereal element: sound' (2003, 26). Smolderen views speech balloons as 'a reified image of human speech' (2014, 53) that situate the speech 'as a sound object within the image' (ibid). In examining this function of balloons, Miodrag notes that while they are 'not visible in the world-of-the-work as they are to the reader, these forms represent diegetic material nonetheless, visualising for the reader what is audible for characters' (2013, 100). The balloon itself is a non-diegetic container or carrier; a signifier of speech that exists outside the story world. But the contents of the balloon are diegetic, as they represent the direct speech of the characters within the story world. Cohn provides a useful examination of the balloon in its role as a carrier of diegetic content, describing how they 'function to encapsulate text (or images) that interface with a "root" through a "tail." With speech balloons [...] the balloon is the carrier, the speaker is the root, and the tail is the tail of the balloon' (2013, 35).

In a digital comic, attempts to integrate audible diegetic sound are problematized by the lack of such elements to contain the sound and connect it with the events being visually depicted. As a result, the role of the rooting object becomes more significant in providing audible sounds with a visual point or origination within the narrative of the comic. An example of this can be found within my own practice in the webcomic series, *The Mr. Nile Experiment* (Goodbrey 2003). *Mr. Nile* is a

metafictional series in which the titular protagonist carries out a series of experiments in an attempt to determine the nature of his fictional reality. The twenty-second instalment of the series depicts Nile standing next to a small radio, which he explains to the reader is 'an avatar, really. Something to give the sound a point of origin within the narrative... [and] ...I'm going to need just a little help on your end to synch things up' (ibid).

The rooting object in this case also functions as a button which the reader is required to click in order to switch on the radio. Once the radio is switched on a piece of audible music begins and Nile can be observed visually reacting to the music that both he and the reader can now "hear." In requiring the reader to click on the rooting object to start the music, it not only integrates the music into the diegesis of the comic, but also aligns the sound to a specific point in the reader's progress through the spatial network of the comic.

A different approach to the integration of audible music into the diegesis can be seen in another webcomic from the same period, *Devil in the Kitchen* by Kean Soo (2003). The narrative of the comic follows a group of friends watching the musician Ashley Maclsaac playing a gig at a local music venue. The layout of the comic is in a similar infinite canvas style to *Mr. Nile*, with the panels arranged on one long webpage which the reader scrolls through to read. However, unlike *Mr. Nile*, a play button and progress bar for an MP3 file sits separately from the comic at the top of the webpage.

To experience the visual narrative of the comic and the audible music track together, the reader must first press play on the MP3 file before then commencing to read and scroll through the comic. As the audio recording of the titular instrumental track *Devil in the Kitchen* plays back, its exact relationship with the diegesis of the comic is uncertain. Rooting objects such as a set of drums, a guitar and a fiddle can be seen in repeated panels in the comic, offering potential points of origin for the music within the diegesis. But mixed in with these are several

panels showing the reaction, excitement and applause of the listening crowd, the audible sounds of which are noticeably absent from the MP3 recording.

Ultimately it is the placement of the MP3 player outside the diegesis of the comic that proves most problematic to the integration of the audible music track as, unlike *Mr. Nile*, the comic lacks any fixed points of synchronisation between the soundtrack and the events being visually depicted. This distinction places *Mr. Nile* into the category identified by Hague as 'sounds in comics' (2014, 73) while *Devil* fits better in the alternate category of 'sounds with comics' (77). This latter category is typified by the reader listening to an audible sound recording while reading a separate comic narrative, without any points of direct synchronisation occurring between the two. Although this is not to say that the two share no interaction at all. To take the example provided by *Devil*, it is noticeable that the high tempo of the musical track can influence the reader to adopt a faster pace in their reading and scrolling of the comic. Hague notes that the more audible sound is relied on to set the reading pace, the more the 'visual content of the comic is subordinated, in temporal terms, to the audible, which directs the speed at which the performance should take place' (77).

Parallels to this can be drawn from the relationship between sound and image in film. Chion notes that on 'first contact with an audiovisual message, the eye is more spatially adept, and the ear more temporally adept' (1994, 11). However, for digital comics this is a potentially problematic phenomenon. As examined in Chapter Four, the digital mediation of comics has introduced the potential to include animated, time-based elements to what was previously a spatial, non-time-based form. This has in turn highlighted the importance of the reader's control over the pace of their reading as a key characteristic of the form of comics. In Chapter Three it was asserted that for a digital comic to still operate as a comic, the rate at which information is absorbed must remain under the reader's control.

Responsive soundtracks

The nature of this potential conflict between reader control and audible sound is highlighted by Groensteen, who identifies two different temporalities at work in audible comics, 'the concrete, measurable time of motion and sound, and the indefinite, abstract time of comics narration' (2013, 70). When the reader has to adjust their pace to match the length of a piece of animation or sound, their control over their reading rhythm 'is sacrificed - or else this synchronization may already have been programmed by the author, who therefore also imposes the rhythm' of progression (ibid). Priego asserts that the use of 'synchronous animation with sound' that takes control away from the reader 'belongs to a different realm in which comics stop being comics' (2010, 277).

To overcome this problem, the key lies not in making the reader adapt their reading to the audible soundtrack, but instead making the soundtrack adapt and synchronise to the process of reading. In *Nile*, the clickable rooting object acted as a crude form of reader-controlled synchronisation. While useful in a metafictional story with direct reader-character interaction, such a technique would be less appropriate in a more traditional narrative, where the nontrivial effort of clicking to activate the sound might interrupt the reading process. Hague describes a more elegant approach to the problem, in which the reader navigates through the comic and 'the sound system is set up in such a way that the soundtrack responds to the reader's position in the narrative, replaying sound effects or adjusting the soundtrack to fit the relevant panel' (2014, 76).

The result is an audible comic with a truly responsive soundtrack in which 'the reader is [...] given control over the way in which the soundtrack functions' (76) and there is no interruption of the reading flow. A good example of this approach can be found in the third part of Stevan Živadinović's webcomic, *Hobo Lobo of Hamelin* (2011). The comic's narrative is a take on the *Pied Piper of Hamelin* legend, with each part of the story laid out in a sideways-scrolling infinite canvas format. In terms of the comic's use of audible sound, it is interesting to note that many of the

concepts identified by Chion in his analysis of film soundtracks can also be seen at work in the responsive soundtrack of *Hobo Lobo*. Foremost of these is the principle of 'added value' which Chion defines as:

the expressive and informative value with which a sound enriches a given image so as to create the definite impression [...] that this information or expression "naturally" comes from what is seen, and is already contained in the image itself. Added value is what gives the (eminently incorrect) impression that sound is unnecessary, that sound merely duplicates a meaning which in reality it brings about, either all on its own or by discrepancies between it and the image (1994, 5).

Part three of *Hobo Lobo* opens at night on the edge of the woods. Accompanying the artwork that establishes this scene is a looping diegetic soundtrack that consists of the ambient sounds of the forest at night. Chion describes sounds used in this way as 'territory sounds, because they serve to identify a particular locale through their pervasive and continuous presence' (75). He notes that such ambient sounds can envelop 'a scene and inhabit its space, without raising the question of the identification or visual embodiment of its source' (75). In this instance none of the animals and insects responsible for the sounds in the soundtrack can be seen. Instead it is the overall image of the forest which can be thought of as the rooting visual element within the comic.

In response to the reader scrolling through the comic, a piece of harmonica music slowly fades up in volume as on the screen a parade of rats can be seen making their way through the forest. Initially the relationship of the music to the diegesis of the story is uncertain but as the reader continues to scroll the rooting image is revealed; the comic's lupine protagonist playing his harmonica and leading the parade of rats. Chion identifies a similar phenomenon in film, where music can 'narrow into' (81) the diegesis once the originating instrument appears onscreen. According to Chion, shifts in music between non-diegetic and diegetic can happen

'at a moment's notice, without in the least throwing into question the integrity of the diegesis' (ibid).

The rat parade then reaches the cliff edge and the protagonist ponders the wealth he will receive when the rats plunge to their doom. Accompanying this sequence a low, ominous droning sound begins to build in volume, creating a sense of foreboding as to the fate that lies ahead for the rats. This is a good example of the way sound 'vectorizes or dramatizes' (13) a sequence, creating 'a feeling of imminence and expectation' (13-14). The sound of a bell tolling is heard, accompanied by the appearance of a bloody scythe on the screen, and then the comic transitions into displaying a series of surreal images of fine dining, wealth and high living. Alongside this visual transition comes a gradual change in the soundtrack from diegetic to non-diegetic, with the harmonica fading out completely to leave only the sound of the ominous, un-ending drone.

Hague asserts that digital comics with responsive soundtracks 'require a relatively modular approach to the sound design' (2014, 76). This modular approach can be heard at work in *Hobo Lobo's* soundtrack, which essentially consists of a series of sound loops and spot effects, setup to play or fade in and out in response to the reader's scrolling progression through the comic. While this approach draws heavily on the language of sound in film, the modular nature of its construction also invites comparisons to the use of sound in videogames. Nitsche identifies the use of 'adaptive audio' systems in games that offer 'a dynamic change of a playing musical piece in relation to the user's interaction' (2008, 135). The resulting combinations of sound effects and music used in games create 'navigable soundscapes' (141). These act as flexible compositions where a 'player's spatial exploration is also a journey through a varying soundscape' (ibid). As *Hobo Lobo* has shown, despite their modular nature these soundtracks are capable of many of the same evocative flourishes found in film and games. Indeed, Nitsche asserts that: 'Elaborate soundscapes can build up a dramatic foreshadowing, provide direct acoustic engagement up to the climax, and mark an end with a cathartic aftermath' (142).

Responsive soundtracks are an example of crossover between videogames and digital comics. The previous chapter examines in detail the potential for direct hybridisation between videogames and comics to create “game comics” that make use of the key characteristics of the form of comics in the mechanics of their gameplay. As part of the practice-based research documented in the previous chapter I created the prototype game comic, *Icarus Creeps* (Goodbrey 2013). Working on this game comic also gave me a new opportunity to experiment with implementing a responsive audible soundtrack in a digital comic.

Icarus Creeps is a hybrid between comics and adventure games. The player navigates the game comic via the use of the arrow keys on the keyboard, which move Icarus panel by panel around the world of the game. The comic features a responsive, non-diegetic musical soundtrack, influenced by the adaptive audio systems commonly found in adventure games. However, my initial intent with *Icarus* had been to create a more complex soundtrack, with elements of both diegetic and non-diegetic sound that would respond to the user’s exploration of the environment and narrative progress. Ultimately, sourcing appropriate sounds and synching these to the player’s actions proved to be a greater challenge than anticipated. To avoid significantly extending development time I opted instead for the simpler musical soundtrack, with changes in the soundscape aligned primarily to transitions between different environment types (such as from inside to outside, or outside to underground).

The Empty Kingdom

To enable further examination of the use of audible sound in digital comics I initiated a new period of practice-based research that resulted in the creation of the game comic prototype, *The Empty Kingdom*. The plot of *Kingdom* follows a videogame player who logs into an empty Massive Multiplayer Online (MMO) game, shortly before the MMO is due to be permanently closed down. Within the game world the player appears as a king, searching his empty island kingdom for

any sign of other players. The game is again controlled by arrow keys and plays similarly to *Icarus*, with the player moving the king from panel to panel through the different parts of the kingdom. While there are puzzles that can be solved in order to reach the narrative's conclusion, the emphasis in *Kingdom* is placed more on the act of exploration itself. My intent was to create an explorable space mediated through the format of an audible digital comic.

McCloud asserts that 'in comics at its best, words and pictures are like partners in a dance' (1993, 156), where each takes turns in leading the narrative. For audible sound to successfully join the multimodality of comics, it too must be given opportunities to lead. In *Icarus* I had begun by creating the comic and game systems before later beginning work on the soundtrack. In *Kingdom*, I began the creative process with the sounds themselves. Using the creative commons sound archive *Freesound.org*, I assembled a library of ambient territory sounds that evoked a range of different environments. From these I teased out an imagined geography in which the sounds could interrelate as part of an adaptive soundscape. Next I drew thumbnails of the landscape in comics form and from this guide created the final artwork, constructing and integrating the modular soundtrack as I progressed.

In a responsive soundtrack, synchronisation between sound and image relies on accurately tracking the progression of the reader through the comic. The "fidelity" of the soundtrack's response is linked directly to the accuracy with which this progression can be measured. An audible comic that groups panels on digital "pages" has a low fidelity of response, as it can only track the points at which the reader navigates from one page to the next. Changes in the soundtrack are therefore limited to these digital page transitions, meaning any loops of sound must match appropriately with all the panels contained on each page. To achieve a higher fidelity of response, the reader can be limited to viewing a single frame of the comic at a time, or be forced to click regularly to build up or change the composition of panels on the screen. Changes in the soundtrack can then be synchronised to the appearance of each new panel, allowing for spot sound effects to be used alongside loops of sound tailored more closely to specific image

sequences. However, as explored in Chapters Three and Four, this approach to panel display can erode key characteristics of the form of comics such as the spatial network and the simultaneous juxtaposition of images.

In *Kingdom*, the comic is divided into fixed, page-like compositions of simultaneously juxtaposed panels. Within these compositions the reader controls the position of the king, moving him from panel to panel using the arrow keys and triggering the transition to a new composition by moving him off the edge of the screen. In this way the king serves as an avatar for the reader within the environment being simulated in the comic and allows for a high degree of precision in the tracking of reader progression. This approach therefore allows for the retention of some traditional concepts of page layout while at the same time allowing for a very high level of fidelity in the responsiveness of the comic's soundtrack. From the practitioner's standpoint, it also necessitates a process of careful experimentation in order to successfully determine the variations in volume needed for each sound loop in each panel of the comic.

Groensteen asserts that in comics, text and image 'enter into an intimate, almost fusional relationship' (2013, 71). However, he also cautions that once other elements such as sound are added to into the multimodality of comics, 'it becomes much harder to achieve this perfect degree of integration: often, they remain disparate elements, aggregated but not fused, unsystematic' (71). In attempting to understand the fusion of comics and audible sound, it is helpful to consider the concept of 'synchresis', which Chion defines as 'the spontaneous and irresistible weld produced between a particular auditory phenomenon and visual phenomenon when they occur at the same time' (1994, 63). As Groensteen asserts, synchresis is problematized in an audible comic due to the conflict between the definite, measurable time of sound and the indefinite, abstract time of comics narration. The use of modular, looping and ambient sound elements in responsive soundtracks are one approach towards successfully achieving synchresis. Loops of audible sound, lacking definite beginnings and endings, can be more easily matched with the

indefinite sequences of fictional time that are created by the reader through the process of closure.

Kingdom primarily makes use of ambient loops in its responsive soundtrack, but it also contains some spot sounds that play at specific points in the reader's progression through the comic. To understand the approach taken in integrating these sounds, it is useful to consider Chion's concept of the 'synch point', which he defines as the 'salient moment of an audiovisual sequence during which a sound event and a visual event meet in synchrony' (58). Spot sounds are by their nature relatively short sounds of a definite length, designed to accompany a specific event or action within the narrative. Placing the synch point of a spot sound in an audible comic represents the hardest challenge to achieving synchresis, as it has the most potential to draw attention to the conflict between definite and indefinite time.

In *Kingdom*, my approach was to treat these synch points as occurring in the gutters between panels, rather than in the panels themselves. One example of this is provided by a section of the comic that includes a bird sitting on top of a rock, which flies away as the king approaches it. In this sequence the reader first sees the bird perched on the rock. Then, as the reader moves the king towards the rock, they hear but do not see the bird take flight. In the next panel the bird is then shown already in full flight away from the rock. Rather than conflict with the still images that make up this sequence, the spot sound is in essence synched with the imagined motion that the reader creates in their mind through the process of closure. This approach plays to the strength of audible sound to suggest unseen movements (Chion 1994, 12), without negating the role of the reader in mentally constructing the 'continuous, unified reality' (McCloud 1993, 67) represented in the panels of the comic.

Conclusion

This chapter has examined the use of sound in audible comics, both as an element

of navigation and as an integrated part of a comic's narrative. The integration of audible elements into the multimodality of comics may include either diegetic or non-diegetic sounds, with the former benefiting from the use of rooting objects to situate them within the diegesis. In audible comics sound may influence the pace of reading and potentially lessen reader control over their progression through the narrative. These issues can be addressed through the use of responsive soundtracks that link control and modification of audible sound directly to reader progression.

In addition to providing a theoretical framework for the study of sound in digital comics, the chapter has provided a practitioner's perspective on the challenges of creating audible comics. The modular, looping nature of a responsive soundtrack is sympathetic to the indefinite time of comics' narration and allows creators to draw tropes from film and videogames in the construction of their soundscapes. The fidelity of response in these soundscapes is determined by the precision with which reader progression can be tracked. A higher fidelity of response can complicate soundtrack creation, but also provides more opportunity for the use of spot sounds alongside looping elements of audio. In placing the synch point for such spot sounds in the gutter between panels, audible sound can support (rather than conflict with) the reader's role in constructing time and narrative within the form of comics.

References

Bordwell, D. and Thompson, K. (2013) *Film Art: An Introduction*. Tenth edition. New York: McGraw-Hill.

Chion, M. (1994) *Audio-vision* (trans. Gorbman, C). New York: Columbia University Press

Cohn, N. (2013) *The Visual Language of Comics*. London: Bloomsbury.

Eisner, W. (2003) *Comics And Sequential Art*. 24th Edition. Tamarac: Poor House Press.

Goodbrey, D. (2002) *Doodleflak*. Available at: <http://e-merl.com/flak.htm> [Accessed 5 August 2015]

— (2003) *The Mr. Nile Experiment 22: Turn On (Tune In)*. Available at: <http://e-merl.com/mrnile/day22.htm> [Accessed 20 July 2015]

— (2013) *Icarus Needs*. Available at: <http://www.kongregate.com/games/Stillmerlin/icarus-needs> [Accessed 17 August 2016]

— (2014) *The Empty Kingdom*. Available at: <http://www.kongregate.com/games/stillmerlin/the-empty-kingdom> [Accessed 17 August 2016]

Groensteen, T. (2013) *Comics And Narration* (trans. Miller, A.). Jackson: University Press of Mississippi.

Hague, I. (2014) *Comics and the Senses: A Multisensory Approach to Comics and Graphic Novels*. New York: Routledge.

Lacassin, F. (2014) 'Dictionary Definition.' In Miller, A. and Beaty, B. (eds) *The French Comics Theory Reader*. Leuven: Leuven University Press.

McCloud, S. (1993) *Understanding Comics*. New York: HarperPerennial.

Miodrag, H. (2013) *Comics and Language*. Jackson: University Press of Mississippi.

Nitsche, M. (2008) *Videogame Spaces: Image, Play, and Structure in 3D World*. Massachusetts: MIT Press.

Priego, E. (2010) *The Comic Book in the Age of Digital Reproduction*. Doctoral thesis, University College London.

Smolderen, T. (2014) 'Graphic Hybridization, the Crucible of Comics.' In Miller, A. and Beaty, B. (eds) *The French Comics Theory Reader*. Leuven: Leuven University Press.

Soo, K. (2003) *Devil in the Kitchen*. Available at:

<http://keaner.net/kitchen/kitchen.html> [Accessed 13 April 2015]

Živadinović, S. (2011) *Hobo Lobo of Hamelin – Page 3*. Available at:

<http://hobolobo.net/tale/3> [Accessed 13 April 2015]