

SoundEffects



An Interdisciplinary Journal of Sound and Sound Experience

*Birgitte Stougaard Pedersen
& Iben Have*

Conceptualising the audiobook experience

Iben Have

Associate Professor

Institute of Aesthetics and Communication

Aarhus University

Denmark

ibenhave@imv.au.dk

Birgitte Stougaard Pedersen

Assistant Professor

Institute of Aesthetics and Communication

Aarhus University

Denmark

aekbsp@hum.au.dk

www.soundeffects.dk



Abstract

In this article we wish to introduce and discuss a theoretical framework for a possible conceptualisation of the differences between reading a printed book and listening to an audiobook. We tend to introduce similarities and differences between reading with the eyes and reading with the ears, implying that we should not discuss the audiobook experience as a remediation of the printed book experience only, but as an entirely different experience that could be conceptualised in continuation of mobile listening practises. As a methodological strategy we will emphasise the differences between the literary practices, reading with the eyes and reading with the ears. These different perspectives on reading are used to accentuate the distinct experiences, and future thorough analyses in continuation of this framework would appear much more complex and connected than in the present article.

Introduction: rethinking the audiobook

The audiobook is not a new phenomenon, but during the last ten years the production, distribution, use and reception of audiobooks have gone through radical and crucial changes, which calls for new consideration and investigation of the phenomenon.

In fact, it was the recording of speech, not music, that was Edison's primary goal in inventing the phonograph in 1877. But only a few spoken-word recordings survive from before 1914. In the 1930s novel-length recordings began in Britain and the United States as a service made for blind people, many of them soldiers returning from the First World War with eye injuries (Rubery, 2011, p. 5). After the Second World War the reel-to-reel technique with the quite heavy seven-inch reels gained ground. A reel-to-reel audiobook could easily take up to 20 tapes and each tape weighed 250 grams. The audio cassette was developed during the 1970s and the word 'audiobook' was introduced in that context. In the 1980s the compact disc replaced the audio cassette, and in 2002 audiobooks became available for download from the Internet in digital formats such as MP3 (Rubery, 2011, p. 8). These years, audiobook readers in general move away from physical book consumption towards the convenience of computerised sources; this tendency is also true for the printed book.

The audiobook has historically been associated with children or practices concerning either dyslexia or visual handicaps. Thus, the audiobook has been considered compensatory; it has been treated in terms of its ability to overcome various kinds of insufficiencies or difficulties. With the newer portable formats – from the cassette to the CD and the MP3 format – this has changed over the years. Especially with the development of digital audio media and portable computerised sources, making it possible to stream or download audiobooks from the Internet, the accessibility and

user-friendliness of audiobooks has appealed to new and broader user groups, and during the last ten years the use of audiobooks has increased significantly. As stated by Matthew Rubery in his introduction to *Audiobooks, Literature and Sound Studies*, which the author claims to be the first ever scholarly book on audiobooks:

Whereas an unabridged recording of Tolstoy's *War and Peace* once required 119 records, 45 cassettes, or 50 compact discs, the entire novel could now be stored in digital format on a portable listening device such as an iPod. Improved ease of use is one reason why listening to audiobooks is among the minority of reading practices found to be increasing in popularity as the number of overall readers continue to decline. (Rubery, 2011, p. 9)

An American survey from 2006 made by the Audio Publishers Association (APA) shows that audiobook users today are younger and more economically well-off (APA, 2006)¹ and that 50 per cent of audiobook buyers are men, who generally buy only one out of four printed books. In Denmark, with a population of 5.5 million (estimate 2012), audiobook sales increased more than 100 per cent from 2009 to 2010, and since 2009 between 50,000 and 60,000 new audiobooks have been made available in Danish libraries each year (Nielsen, 2012).

The increasing popularity of the digital audiobook is accompanied by the growth of other digital book formats, for instance the e-book. From several directions, the printed book is thus being challenged as the primary medium for reading literature. Studying the audiobook one might, as one crucial point, bring up the discussion, which will be the key question of this article, of whether the audiobook should be seen as an alternative way of reading a book, or if we instead ought to change our opinion of the audiobook more radically and study it as a popular phenomenon, focusing on the everyday use of audiobooks as part of the digital, mobile audio culture. We consider audiobook use a special instance of mobile listening in line with, for example, music, radio, audio guides or audio self-help therapy.

New media reuse and renew old media, and the representation of a medium (the printed book) in another medium (the audiobook) can, with Bolter and Grusin's (2000) term, be named *remediation* – a concept that has been agenda-setting in conceptualising the use of existing media in digital structuring processes. Studying the audiobook as a remediation of the paper book, there is a tendency to focus on the limitations and to treat the audiobook as a kind of shortcut, a reading practice that is uncritical, unreflective and relaxed, as opposed to a conception of *deep reading* as a contemplative experience. One may sceptically ask: Can we really read with the ears? In the mid-1990s Sarah Kozloff, a film scholar, studied what had been written and published about audiobooks (mostly in American popular media, journals and newspapers), and she concluded:

The commentary I have found primarily focuses on comparing the format (often unfavorably) with the experience of reading printed books. To many, listening to

audio books is a debased or lazy way to read, with connotations of illiteracy (only pre-literate children listen to stories); passivity (real reading entails self-construction of the narrative voice); abandonment of control (real reading involves pausing, skimming and savoring); and lack of commitment (real readers sacrifice other activities for their books). These underlying prejudices are reinforced by the format's association with such 'lowbrow' offerings as business advice and self-help titles, as well as radical abridgments of revered novels. Audio books are both dismissed as a negligible fad, and feared as another potent threat to traditional literacy. (Kozloff, 1995, p. 83)

This article focuses on the use of digital audiobooks and discusses the reading experience, as it extends the reading situation to a listening mode. Our study intends to highlight the relationship between reading and listening from a methodological perspective. With the ambition of conceptualising the audiobook experience we will bring in and discuss the methodological usability of Lars Elleström's (2010) model of mediality and multimodality. Our starting point is that we need to rethink the remediation perspective and the possibilities of conceptualising the reading/listening activity and experience. We intend to consider audiobooks as not only restricted to literary culture, but to include audio culture (or media culture in general) as well, as we believe that in the present time audiobooks need to be discussed in terms of their relation to contemporary digital media culture and its listening practices.

From an experience perspective we will take our point of departure in the mediality² of the digital audiobook and try to answer the key question of this article by creating a framework for taking some initial steps in studying how the digital audiobook differs technologically, aesthetically/perceptually and sociologically from the paper book. Thus, this study will adopt a comparative approach, as we believe that the materiality and affordances (technology and user possibilities) shape and frame the experience of reading with the eyes and ears, respectively, including sensorial or modal aspects. The technological affordance of the media frames our use hereof and therefore the epistemological and phenomenological contours must be outlined via the material and technological gaze. The way a certain technology controls our possibilities of experience and use seems to frame our focalisation (focalisation in the way Gerard Genette uses the term designates the offered point of view in a story). This experience, conditional on technology, is what this article intends to investigate and qualify.

Conceptualising the digital audiobook experience

As a fundamental thought and a basis for the study of this interplay between technology and experience, it seems relevant to include a philosophical perspective. Bernhard Stiegler has argued that technologies through history cannot be considered exterior only, as prolongations of the body; they are continuously internalised and control our possible perceptions, imaginations and comprehensions of the world – a

radicalisation of the observations Marshall McLuhan made in *Understanding Media* from 1964. As a historical example of this process Matthew Rubery describes how the phonograph replaces the bodily, speaking voice with a 'disembodied acousmatic voice' (Rubery, 2011, p. 14). In continuation hereof, Stiegler suggests that media as well as memory must be conceptualised as associated milieus, in that it is not possible to sustain the relationship between interior and exterior, which are mutually determined:³ In Stiegler's view, humans in their point of origin are designated by limitations that make us dependent on technical prosthesis. Humans are conditionally constituted by technology, and grammatisation in this connection is characterised, for instance, as a process where something technical and spatial is changing a time-based practise, for instance the transition from speech (flux) to writing (spatial and reproducible).⁴

The history of human memory is that of [...] grammatisation. There is no interiority that precedes exteriorization, but to the contrary exteriorisation constitutes the interior as such. [...] The process of grammatisation is the *technical history* of memory [...] writing, as the discretization of the flux of speech, is a stage of grammatisation. (Stiegler, 2006, p. 2)

As Stiegler proposes, grammatisation in this way, through human history, externalises our memory through writing, for instance, and in this perspective determines our conception of time as well as our conception of the world as such. This philosophical perspective underlines the importance of the main question of this article, namely the status of the audiobook experience as a particular form of mobile listening. If the audiobook experience is not merely a remediated literary experience, it takes part in extending or changing our conceptions of what a literary experience is and thereby questions literature as such. It reinstalls in the reading experience an oral mode that has been thrust into the background for a long time, and in this process silent reading has been conventionally naturalised. The digital audiobook as a phenomenon approaches the literary experience through new technologies and new grammars, and in this process mediality and our possible notions of literature are changed. One of the things that have changed in this grammatisation process is the oral modality of the audiobook, which reoccurs as a secondary orality in the sense of Walter J. Ong (1982), but in an entirely new contextual setting that involves the body as a moving and sensing participant in the reading experience. This new setting is partly the result of the fact that the mode of listening as such is situated and specific. In this sense the digital audiobook experience tends to create a new situation in which an oral modality is matched with a listening activity that takes place in a singular situation and frames the reading experience. This could be called grammatisation, and it occurs in the interplay between bodily and cognitive inputs and a specific listening mode. This coproduction of the interior and exterior

demands a reconfiguration of the epistemological understanding of this specific reading situation – and in this sense Stiegler’s perspective on media as milieus seems productive. We need to understand the experience and the technology that make the experience possible as correlated. We do not consider the relation between the printed book and the audiobook oppositional, but in terms of studying the literary experience we need to understand the technological aspects to produce meaning to the overall definition.

Intermezzo – a scenario

Mediated sound reproduction enables consumers to create intimate, manageable and aestheticised spaces in which they are increasingly able to, and desire to, live. (Bull, 2005, p. 347)

As we will illustrate in the following hypothetical scenario, listening to audiobooks can establish a mental private space, an everyday ritual, just like other (old) everyday media (books, radio, television), but in line with new digital mobile media the digital audiobook affords new kinds of ritualised practices.

Jenny lives ten kilometres outside a big city and routinely listens to audiobooks on her iPod while commuting to and from work in the city. It is a beautiful and easy ride on an asphalted bicycle track, through woods, meadow areas and along a river. No cars or industrial sounds are audible, just other ‘soft’ road users and bird song. About halfway she reaches the heavily trafficked ring road and the traffic situation and soundscape change to urban. She struggles uphill to her office near the city centre.⁵

This Monday morning she listens to the first chapters of a crime story. The female narrator speaks in a deep, soft and calm voice when she describes the mental thoughts and feelings of the characters or the setting in the rough, dismal winter landscape of a small village situated in peripheral Denmark. Jenny is mentally immersed in the story, appreciating the green vigorous landscape she cycles through and knows so well. The story makes her think about a heavy winter, where she had to struggle through the snow to get home, and her bike finally collapsed due to salt and water. Her thoughts sometimes wander away from the story like that, and she accepts the consequent gaps in the narrative and does not bother to take the iPod out of her pocket to jump backwards in the book. She tries to find literature that is not too mentally demanding and compact. The physical exercise makes her feel good, quite comfortable and privileged, with the soft voice in her ears telling a story especially for her. The voice and the narrative frame her experience of the changing surroundings, just as the landscape influences her experience of the audiobook.

When Jenny reaches the city she has to turn up the volume, but is nevertheless unable to drown out the surrounding noise. A larger part of her audiovisual atten-

tion is now used to navigate in the city. Close to her workplace she meets an old colleague whom she stops to talk to. She switches off the audiobook, takes the earplugs out of her ears, so as to not appear impolite and to concentrate on the short conversation (like sunglasses, earplugs can signal a distanced attitude when talking to someone face to face). When Jenny turns on the book again, the conversation with the old colleague still fills her thoughts, and she suddenly realises that she has not listened to the story for the last five minutes. She has to go back to the beginning of the chapter, but she still has trouble paying conscious attention to the story, as the work of the day interferes with her thoughts as she approaches the office. She has to turn off the audiobook and wait until the ride home to continue reading.

It is five o'clock in the afternoon and the day at the office has been quite tough and filled with a lot of problems. Jenny is looking forward to her ride home, in the good company of her audiobook. Actually, the reading is the reason why this bike ride is of value to her. It qualifies the ride as more than just an act of transportation. As soon as she turns on the story she forgets her trouble at work and immerses herself in the narrative, looking forward to get to the more quiet areas by the river. When she reaches her home, she turns off the audiobook in the middle of a chapter, steps out of her private audiobook experience and into her afternoon family routines.

The modalities of the audiobook experience

To conceptualise the audiobook experience, illustrated by this hypothetical example, and discuss the differences between the printed and the audiobook medium and experience, we will introduce a number of methodological distinctions concerning media and modality. In this case two comparable theoretical frameworks, *medium theory* (medium ecology) and *intermedia studies*, both provide an applicable framework with similar agendas: trying to figure out and conceptualise what constitutes a medium and how we can analyse it. In this article, however, we will primarily use and discuss the framework suggested by the Swedish intermediality theorist Lars Elleström.⁶

In the text 'The Modalities of Media: A Model for Understanding Intermedial Relations' (Elleström, 2010), Elleström suggests a way of conceptualising media via four different modalities. These modalities are present in every medium, but can present themselves in distinct ways. Our use of the term *modality* is inspired by applied semiotics and seeks to investigate communication practices especially. Modality studies deal with the ways in which a sender marks her or his communicative act in continuation of the concept of mode. Mode can be defined as ways of being or ways of doing, and it constitutes a resource for creating meaning. Modality studies investigate, using a broad variety of practices, how meaning arises, is

distributed and interpreted through a lot of different communicative modes.⁷ Elleström's text combines this notion of modalities with an intermedial, typological approach by suggesting four modes, which comprise perspectives of a medium's modalities and meaning-creating radius:

When I speak of *modalities* henceforth, I mean these four necessary categories in the area of the medium ranging from the material to the mental, and when I speak of *modes*, I mean the variants of the modalities as describes below. Entities such as 'text', 'music', 'gesture' or 'image' are *not* seen as modalities or modes. (Elleström, 2010, p. 16)

Elleström proposes a model for understanding media that involves four different modalities that are seen as 'essential cornerstones of all media without which medi-ality cannot be comprehended and together they build a media complex integrating materiality, perception and cognition' (Elleström, 2010, p. 15).⁸ In Elleström's model the four modalities are present in every medium, but the way they interact changes from medium to medium. The first modality is the *material*, which is the 'latent corporeal interface of the medium' (for instance, text or sound waves); the second modality is the *sensorial*, that is 'the physical and mental acts of perceiving the present interface [...] through the sense faculties' (for instance, to see, to hear, to smell, to feel) (Elleström, 2010, p. 17). In this sense media must be realised through perception, and Elleström's model incorporates a process that moves from material to mental. The third modality, moving to the cognitive level, is the *spatiotemporal* modality: 'Sense-data cannot be grasped, cannot be conceived as sensation, unless they are given some sort of form, *Gestalt*, in the act of perception. The spatiotemporal modality of media covers the structuring of the sensorial perception of sense-data of the material interface into experiences and conceptions of space and time' (spatiotemporal perception covers, for instance, width, height, depth, time) (Elleström, 2010, p. 18). This modality negotiates cognitive ideas of time and space relations. The last modality, the *semiotic*, concerns meaning production: 'meaning must be understood as the product of perceiving and conceiving subject situated in social circumstances [...] All our sensations are the result of interpreting, meaning-seeking mind' (Elleström, 2010, p. 21).

Building on Elleström's suggestions, we wish in the following to discuss our key question concerning the remediation of new media by producing an outline along his axes of modality, testing whether 'the semiotic modes, together with the spatiotemporal, the sensorial and the material mode, form the specific character of every medium' (Elleström, 2010, p. 23). Discussing the modalities of the printed book and the audiobook as two distinct media experiences, we intend to follow the line of Elleström's model, but also to discuss its limitations and possible lacunas.

The material mode of the printed book can be understood as writing inscribed on the technical medium paper. It is sensorial, perceived partly through sight and

partly by the tactile sense of touch, touching the paper and turning the pages with one's moving hand and arm (sitting in a chair or lying on a bed or similar). The negotiations of space in reading a printed book deal to a great extent with the cognitive level, imagining space and time on the level of the narrated. Time as a mode is present both as imagined time in the narrated space and as actual time in the physical act of reading. Reading a printed book thus involves the liberty to organise; it makes it possible to speed up, slow down as well as jump back and forth in the text. The semiotic modality deals with the way meaning is created in the process of acquiring a fictive literary course of events and a mental fictive universe.

Using this modality perspective, the audiobook experience differs in almost every respect from that of the printed book. As a material modality, listening to a digital audiobook involves sound waves that consist of technologically mediated orality. The sensorial modality involves hearing as well as the possible tactile experience of earphones and an interface, but it also involves a perhaps moving body and sight that act independently of the reading situation. The spatiotemporal negotiation makes the audiobook experience extremely complex. The cognitive process by which one absorbs the story and plot is, in principle, identical to the reading of a printed book, but the imagined fictitious space is constantly, via sensorial inputs, challenged by a real physical space which one's eyes and body move through. The reading situation itself also appears quite different, as the text is performed by a narrator who interprets the text. When a physical, audible voice reads the text aloud, a number of stylistic choices are made concerning the text. The literary intonation is shaped, as a number of decisions are thoroughly made or implemented; accentuations, intensity, tempo, phrasing and voice qualities are all part of the listening process. In a situation where a lot of stylistic choices are made before the input reaches the ear of the listener, a standardisation or a possible hermeneutic closure is completed, both with regard to the stylistic and the semantic part of the literary experience. One is assigned to the intonation, tempo and phrasing of the narrator in a very concrete way. This performativity of the text makes the audiobook experience more vivid and more frozen at the same time or, at least, 'time-forced': vivid in the concrete act of enunciation and frozen in the sense that the reading activity is tied up in a strict sense. A commuter like Jenny can neither make easy jump cuts nor speed up or slow down the reading according to the level of interest, as readers of printed books often do. With regard to the semiotic level it is crucial to understand this performative reading or narration as a parenthetical, interpretative and thereby meaning-creating gesture.⁹ Finally, to read a printed book demands certain reading skills that are typically taught in schools and are very different from the skills used to understand spoken language.

It appears that a number of these modality issues can provide us with interesting theoretical perspectives, not least the interplay between the four modalities.

In continuation hereof one of our theses is that the audiobook experience might not be sufficiently described by the four modalities, as the sensorial, semiotic and spatiotemporal modes seem not only to interact, but also to mutually transform the conditions of the medium itself. At any rate, the semiotic or meaning-creating process must be understood as quite different in the two different manners of reading, as the movement of both sight and body interferes with the semiotic process via mental and physical negotiations of space and time. From our perspective, we may have to discuss in particular the audiobook experience alongside these interplays, as it seems to be the case that confusion of the *concrete* and the *imagined* spaces of the experience plays a specific and significant part, as evident in the Jenny scenario above. In a sense, the relationship between the spatiotemporal mode and the semiotic mode is challenged or confused by the sensorial and, not least, by the intersensorial or multimodal mode of the audiobook experience. Another possible insufficiency of the model concerns bringing together the spatial and the temporal cognitive aspect. Other scholars, such as Morten Kyndrup, would suggest that the organisation of time is clearly separated from the aspect of the spatial status of each mediality (Kyndrup, 2012, p. 8).¹⁰

Summing up the modal conditions of the printed book and the audiobook, the preliminary answer to our key question – Is this a remediation of a literary experience, an alternative way of reading a book, or is it a distinct example of mobile listening (according to the distinctions suggested by Elleström)? – must be that the audiobook experience differs from the paper book experience in nearly all respects and thus seems to promote a discussion of the two as different experiences of what seems to be the same content (a novel, for instance). So following Elleström's model seems to entail an evident need for rethinking and reconceptualising the literary experience in relation to the media. We also need to make a clearer distinction inside the model, sophisticating the sensibility of the differences between the temporal and the spatial aspects of the reading experience. The semiotic content of Tolstoy's *War and Peace* might be treated as identical in audio and paper format, respectively, but the question is whether this content is affected by the different modal aspects and, if so, to what extent; we cannot conceptualise the experience along the same axes. As we have already suggested, the delicate interplay between the modalities may even entail that *War and Peace* does not necessarily afford identical literary experiences, due to the different sensorial, spatiotemporal and social settings.

Elleström makes a distinction between *technical* media, *basic* media and *qualified* media (Elleström, 2010). All media need technical media to be realised, and in the case of the audiobook the technical medium is, for example, the iPod, the smartphone and the MP3 file. Basic media are identified by their modal appearances, like auditory texts or moving images, and qualified media refer to different kinds of art forms and cultural media types like music, literature, dance, film, radio etc. In

continuation of the discussion above we might need another term than literature to capture the qualified media aspect of the audiobook – a word like *listenrature* could be a tentative suggestion.

To read /to listen – in context

Below we discuss more specifically the phenomenological differences between the literary experience of the printed book and the digital audiobook, differences that will underline that we should no longer discuss the experience as literary; instead, we need to identify and understand it as a new kind of mediated auditory experience. One important discussion, following the line of modality issues, is that the digital audiobook via its technical conditions (that is, the small portable format) establishes new affordances. The advantage of using J. J. Gibson's concept (1979) in this phenomenological context is that, in opposition to more deterministic approaches to the media/user dichotomy, affordance defines a quality that neither belongs to the object nor the subject, but appears in interaction between the two. Technological artefacts are social as well as technical and do not merely determine agency, but constitute agency: perception, cognition, feelings and habits (deNora, 2000, p. 40). By using Elleström's model analytically we have qualified some of the different affordances of the printed visual book and the digital audiobook, respectively. The affordances of the audiobook offer the possibility of specific mobile listening practices, and in continuation hereof we need to understand the material, technical conditions in close relation to the distinctive character of the sensory possibilities of the reading/listening situation. Before moving closer to the multimodal character of the experience, we wish to frame the reading situation from a contextualised (social) perspective.

Reading in a printed book, or an e-reader for that matter, can take place in various settings: in our homes, at work, in a train or bus. Common to these examples is that they rely on some kind of bodily tranquillity, due to the fact that one's sight is tied to the pages. They also require some kind of relatively low or, at least, stable level of auditory distraction in order to concentrate on the reading experience.

The audiobook experience can take place in approximately the same settings as the printed book experience, but it involves one central difference: body and sight are free to move around during the process of reading. In the last decades, the audiobook reading experience has often taken place in spaces connected to everyday life and everyday obligation – a listening situation that is similar to that of radio (Scanell, 1996; Larsen, 1999). One can listen to an audiobook while cooking, gardening or doing the laundry as well as when moving between work/school and one's home, like Jenny in the scenario above. Commuting as a concept in this setting becomes an interesting frame for the reading experience. One can read printed and audio-

books when using public transport, while the audiobook alone is suitable for the car, bicycle or walking. Ben Highmore in *Everyday Life and Cultural Theory* suggests that much everyday experience in industrialised Western modernity is potentially characterised by monotony: 'The repetition-of-the-same characterizes an everyday temporality experienced a debilitating boredom' (Highmore, 2002, p. 8). Commuting is a daily routine that extends one's working hours and is thus associated with a feeling of monotony too. On the other hand, the commuting space can be regarded as a pause, as a momentary oasis in between the obligations at work and the obligations at home. In this sense, the character of the commuting space is ambivalent, in that reading or listening activities, for instance, make it possible to fill up or qualify time in tendentiously meaningless everyday monotony. Highmore states that 'The everyday offers itself up as a problem, a contradiction, a paradox: both ordinary and extraordinary, self-evident and opaque' (Highmore, 2002, p. 16).

Listening to audiobooks is to a great extent comparable to the semi-distracted listening mode of talk-radio. Many users listen to audiobooks not only to get a literary experience, but also to be entertained or maybe above all to be in good company and feel good. The presence of a calm human voice has since (and even before) birth had a calming effect on us. As Stig Hjarvard writes, inspired by Simmel's concept of sociability, media are social technologies as well as information technologies, providing entertaining parasocial contact with other people (real or fictive), without having to contribute much yourself (Hjarvard, 2005, p. 22). In media studies the theory of parasocial relationships deals mostly with television and the audiovisual experience (Horton & Wohl, 1997) in which the viewer feels a kind of relationship with, for example, television hosts. But having a constant human voice in one's ear telling a story may result in a similar feeling of parasocial company and comfort.

Intersensorial/multimodal perspectives in reading and listening practices

The phenomenological experiences of reading with the eyes and reading with the ears are fundamentally different, and this leads to a new epistemological situation for the literary experience, concerning the act of listening. Historically the literary experience seems attached to the imagination, an inward space, where sight is tied to the writing on the page. Reading a printed book is a silent act, where voices are called forth by the reader's activity. The experience of reading a book thus represents a possible metaphorical audibility; the sound takes place in a silent zone of attention and the literary experience happens inside the reader's imagination and body. The experience of an audiobook involves an amount of different sensorial inputs from the surroundings, consisting of different and alternate visual and tactile inputs; the listening can be disturbed or enriched by audible inputs, for

instance the sounds of the city, as a result of moving around while listening, exercising or commuting.

As we have shown via the use of Elleström's model, an interplay between the four modalities takes place in both reading with the eyes and reading with the ears, but as we have suggested the *interplay* between modalities may be the most interesting parameter in investigating the phenomenology of the situations.

Multimodal research investigates types of meaning that view language only as part of a multimodal ensemble. Multimodality thus steps away from the notion that language always plays the central role in communicative interaction (Jewitt, 2009, p. 14; Kress, 2010, p. 80). All modes are shaped through their cultural, historical and social uses.¹¹ Multimodal research sheds light on the interplay between modalities especially, and this interplay can be both aligning and contradictory.¹² Listening to audiobooks as a multimodal experience can be compared to watching a film, which is also a multimodal experience. Even though the semiotic input of the film's soundtrack and the visuals seem to be contradictory, it is very natural for the human cognitive processing of sense stimuli to synthesise it into comprehensible wholes, just like Jenny did, when listening to the descriptions of the winter landscape while sensing the green vigorous landscape around her.

The experience of reading a printed book is multimodal on the material level: the technical medium, paper and ink are combined with the modality of text. The sensorial modality both involves vision, tactility and the movement of the hand and arm. The spatiotemporal modality is primarily linked to the level of the narrated and the way in which it distributes imaginations of time and space. Even though the reading situation of course takes place in concrete physical time and space, it is often stressed that being absorbed in a reading experience makes one forget time and space.

The audiobook experience is multimodal at the material level, regarding sound waves and orality transported through the technical medium of, for example, an MP3 player or an iPod. The sensorial modality involves, as stated, both the ears, the tactile feeling of the earplugs and the possible movement of body and sight. The spatiotemporality of the narrated is challenged by sensory inputs from the surrounding space, as in Jenny's experience in the scenario above. The semiotic modality is being pushed or fed by the delicate interplay with the body that senses and moves in continuation of the narrated time and space. A fifth modality may be said to occur in continuation of this situation – as a kind of atmosphere (Böhme, 2001) or associated milieus (Stiegler, 2006) that are created by the interplay. This fifth modality designates the situation as a whole, and in this sense it might be preferable to discuss it as a metamodality emerging from the specific situation.

We will suggest that the audiobook experience not only creates a physical and cognitive bubble, as suggested by, for instance, Michael Bull in his study of audio

culture and mobile listening practices. Bull argues that listening to music is an auditory privatising of, for instance, the workplace (Bull, 2007, p. 113), creating a kind of cocoon or what he describes as the *aural solipsism* of the iPod (Bull, 2007, p. 26). Instead, we believe that both a synchronic and a diachronic flow emerge: synchronic through different layers of perception in a specific moment, and diachronic while the body moves on through time, all layers contributing to the overall situation and experience. This overall experience or flux both affects the listening process – the story you are listening to will be influenced by the nature of your body movements and your surroundings – and the listening experience, which will become part of the listener's experience and construction of her or his environment.

The listening process might thus take part in creating a momentary and changing sense of place. Reading with the ears comprises a kind of filling up of both time and space; it is a possible answer to the monotony of everyday commuting. As Bull states in *Sound Moves: iPod Culture and Urban Experience*, 'Audiobook listening changes the way we perceive our social environment' (2007, pp. 40–41). Each listening situation is singular with regard to the exchange of meaning in or from the situation. In continuation of this it is suggested by Wittkower (2011) that listening to an audiobook while moving around makes the story turn into the *figure*, whereas the walking or other movements of the body become the (back)ground: 'The audiobook *forms* the context for physical and social experience rather than being experienced *within* a physical and social context' (Wittkower, 2011, p. 228).

In keeping with the notion that the interplay between modalities creates a singular and specific situation, Henri Lefebvre states that everyday places are created by the intervention of the body (2004). Other thinkers define places as a matter of interplay between humans and their surroundings (Norberg-Schulz, 1976) or between a temperament and a locality (Ringgaard, 2010). All of these definitions of place, in one way or another, seem to concern the body as a mediating parameter. The body, as in the case of audiobook listening, is sensitive to auditory clues, visual clues as well as to movements; and this body also includes the soaring thoughts that might interrupt the narrative flow of the audiobook, as in Jenny's case. We believe that the interplay between the listening process and the reader's sense of place is actualised in a highly interesting way due to the digital audiobook experience. This interplay may create a sense of place or locality, but through mobile practice, producing a more dynamic experience of an atmosphere or a milieu, we can consider either a fifth modality – a *metamodality* – or in other words a specific characteristic of the medium as such. Jenny's gaze on the changing surroundings, from rural to urban, is influenced by the plot and the voice of the audiobook, which thus establish a sense of place. This may be connected to the way she perceives the listening situation as a whole: a mood or atmosphere that is created by the interplay between the story (including the narrator's voice), the changing surroundings, the body which

senses and moves and the technical media. Together these modes form a specific situation: a fifth meta-modality that takes into account the specific situation as a dynamic interplay between the four modalities – a phenomenology of the audiobook that understands the medium experience as a particular sensory atmosphere or a milieu. This fifth modality might, instead of a general system, be conceptualised as a meaning-creating process, including a clear focus on the experience of a process that would be the next analytical step of Elleström's model, which seems more analytical in the philosophical sense, engaged in creating a general typological framework.

Conclusion

To read a book in silence is historically connected to a contemplative, absorbed experience, in which sight is bound to the page. The sensorial experience is connected to vision and space, but also to sound (metaphorically) and time. These features are all linked to the process of reading. We can describe this as an exclusive, high-absorbency practice, which monopolises much of one's attention.¹³

The practice of listening to an audiobook seems concentrated, yet also distracted. Both reading and listening develop through time, and the experience is to a higher degree suffused by other sense stimuli. Visual or physical inputs are not consistent with auditive inputs. This entails a risk of a contradictory experience, but it could also be seen as a possibility for other sensorial inputs to contribute to the literary experience – of *listenrature*. We may discuss it as a non-exclusive, low-absorbency occupation – listening, which one can do while engaged in other practices; like radio, a secondary medium (Larsen, 1999, p. 259). It is interesting to notice that the majority of people listening to audiobooks are not substituting it for the printed format, but use it as a supplement to visual, text-based reading in relation to other kinds of social practices (Mediatore, 2003, p. 318).

As we have demonstrated in this article, studies of the audiobook will be limited if research approaches choose to view and emphasise the audiobook as a remediation of the paper book. A new sensory experience – a new kind of use, which has more in common with mobile audio media than with the printed book – produces the notion of and the need for approaching the audiobook phenomenon as a new medium experience that calls for a new theoretical framework. In this article we have aimed to outline a preliminary sketch for analysing the digital audiobook inspired by Elleström's theoretical framework on mediality and modality. Thus, we have presented a number of possible ways of conceptualising the audiobook experience and discussing the interplay between modalities, creating a certain atmosphere, a kind of dynamic *meta-modality*, which should be taken into account. So, can

you really read with the ears? We believe you can, but it requires that you reconceptualise the concept of reading something not necessarily attached to sight.

References

- Audio Publishers Association (APA). Accessed on 2 July 2012: <http://www.audiopub.org/LinkedFiles/2006ConsumerSurveyCOMPLETEFINAL.pdf>.
- Bolter, J.D., & Grusin, R. (2000). *Remediation. Understanding New Media*. Cambridge, Massachusetts: The MIT Press.
- Böhme, G. (2001). *Asthetik*. Munich: Wilhelm Fink.
- Brügger, N. (2002). Theoretical Reflections on Media and Media History. In: Brügger, N., & Kolstrup, S. (eds.), *Media History: Theories, Methods, Analysis*. Aarhus: Aarhus University Press.
- Bull, M. (2005). No Dead Air! The iPod and the Culture of Mobile Listening. *Leisure Studies*. Vol. 24, no. 4, 343-355.
- Bull, M. (2007). *Sound Moves: iPod Culture and Urban Experience*. New York: Routledge.
- Danish Association of the Blind. Accessed on 2 July 2012: <http://www.dkblind.dk/indsats/kultur-fritid/lydbøger/lydbogens-historie>.
- Cook, Nicolas (1998). *Analyzing Musical Multimedia*. Oxford: Oxford University Press.
- Danske Lydbøger (2010). *Mænd lærer at læse igen*. Accessed on: <http://danske-lydbøger.dk/lydbøger/maend-laerer-at-laese-igen.html>.
- deNora, T. (2000). *Music in Everyday Life*. Cambridge: Cambridge University Press.
- Elleström, L. (2010). The Modalities of Media: A Model for Understanding Intermedial Relations. In: *Media Borders, Multimodality and Intermediality*. New York: Palgrave Macmillan.
- Gibson, J.J. (1979). *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.
- Highmore, B. (2002). *Everyday Life and Cultural Theory*. London: Routledge.
- Hjarvard, S. (2005). *Det selskabelige samfund*. Copenhagen: Samfundslitteratur.
- Horton, D., & Wohl, R. (1997). Massekommunikation og parasocial interaktion: Et indlæg om intimitet på afstand. *MedieKultur*, 13(27), 27-39.
- Jewitt, C. (2009). *The Routledge Handbook of Multimodal Analysis*. London: Routledge.
- Kozloff, S. (1995). Audiobooks in a visual culture. *Journal of American Culture*, 18(4), 83-95.
- Kress, G. (2010). *Multimodality – a social semiotic approach to contemporary communication*. New York: Routledge.
- Kress, Gunther & van Leeuwen, Theo (2001). *Multimodal Discourse. The Modes and Media of Contemporary Communication*. New York: Bloomsbury Academic.
- Kyndrup, M. (2012). Hvorfor medialitet? In: Pedersen, B.S., & Sørensen, M.M.Z. (eds.), *Medialitet, inter-medialitet og analyse*. Aarhus: Aarhus University.
- Lefebvre, H. (2004). *Rhythmanalysis – Space, Time and Everyday Life*. London: Continuum.
- Larsen, B.S. (1999). Radio as Ritual: An Approach to Everyday Use of Radio. *Nordicom Review*, 21(2), 259-275.
- McLuhan, M. (2006 [1964]). *Understanding Media*. London, New York: Routledge.
- Mediatore, K. (2003). Reading with Your Ears. In: Chelton, M. K. (ed.), *Readers' Advisory*, vol. 42, no. 4, pp. 318-323.
- Meyrowitz, J. (1994). Medium Theory. In: Crowler & Mitchell (eds.), *Communication Theory Today*. Stanford: Stanford University Press.
- Mitchell, W.J.T., & Hansen, M.B.N. (2010). Introduction. In: Mitchell, & Hansen (eds.), *Critical Terms for Media Studies*. Chicago and London: The University of Chicago Press.
- Nielsen, M. (2012). Lydbøger hitter – også i førerhuset. Accessed on 2 July 2012: <http://www.dr.dk/Nyheder/Kultur/2012/04/132418.htm>.
- Norberg-Schulz, C. (1976). *Fenomenet Plats*. KAIROS, 5.

- Ong, W.J. (1982). *Orality and Literacy. The Technologizing of the Word*. London, New York: Methuen.
- Pedersen, B.S. (2012). Intermedialitet – materialitet, typologisering, oplevelse. In: Pedersen, B.S., & Sørensen, M.M.Z. (eds.), *Medialitet, intermedialitet og analyse*. Aarhus: Aarhus University Press.
- Ringgaard, D. (2010). *Stedssans*. Aarhus: Aarhus University Press.
- Rubery, M. (ed.). (2011). *Audiobooks, Literature, and Sound Studies*. New York and London: Routledge.
- Scannell, P. (1996). *Radio, Television and Modern Life*. Oxford: Blackwell.
- Sexton, Jamie (2007). *Music, Sound and Multimedia: From the Live to the Virtual*. Edinburgh: Edinburgh University Press.
- Stiegler, B. (2006). Anamnesis and Hypomnesis. *Ars Industrialis* (Association internationale pour une politique industrielle des technologies de l'esprits). Accessed on 1 July 2012: <http://arsindustrialis.org/anamnesis-and-hypomnesis>.
- Wittkower, D.E. (2011). A Preliminary Phenomenology of the Audiobook. In: Rubery, M. (ed.), *Audiobooks, Literature and Sound Studies*. London: Routledge.

Notes

- 1 The APA found that the average age of audio listeners was roughly 45 years; their average yearly income was above 47,000 dollars, and their average educational attainment quite high.
- 2 *Mediality* or *mediacy* can in a broad sense be understood as a medium's material manner as well as its specific potentials for use and experiences. See for instance Mitchell & Hansen (2010), Brügger (2002) or Pedersen (2012).
- 3 The milieu metaphor can again be compared to the one used by media sociologist Joshua Meyrowitz, who defines it with the analytical question, 'What are the relatively fixed features of each means of communicating and how do these features make the medium physically, psychologically, and socially different from other media and from F2F interaction?' (Meyrowitz, 1994, p. 50).
- 4 For further definition, see <http://www.arsindustrialis.org/grammatisation>.
- 5 Jenny is a fictive person, but the scenario is inspired by one of the author's own experiences using and listening to audiobooks.
- 6 This choice of theoretical sources implies that we use the term *mediality* (Elleström, 2010) instead of *mediacy* (Brügger, 2002), which is a more common term in media studies.
- 7 See, for instance, Kress (2010) and Jewitt (2009).
- 8 Brügger calls this 'variables' in the media constitution.
- 9 In continuation hereof it seems relevant to discuss the role of the literary voice in its meeting with the voice of the narrator. We intend to deal with this issue in future publications.
- 10 For further studies on the relation between time and space in the user's experience, it might have been relevant to consult Eskelinen's *Cypertext Poetics* from 2012 (London: Continuum).
- 11 The field of multimodal research is in itself differentiated and complex. In Gunther Kress and Theo Van Leeuwen a mode is understood as a semiotic source in a very broad sense. This can make it difficult to compare single events, understood as modes, and the field in general might need more precise differentiations. This will, however, not be discussed further in the present article.
- 12 Similar interests and concepts are developed in media studies as multimedia analysis or theory of audiovisual perception, as developed by Nicolas Cook (1998), Kress & Leeuwen (2001), Jamie Sexton (2007) and others. The concepts of aligned and contradictory relationship between music and the moving image have for example a long history in film music studies.
- 13 We wish to thank Steven Connor for this way of describing the practice of reading.