Handbook of Research on 3-D Virtual Environments and Hypermedia for Ubiquitous Learning

Francisco Milton Mendes Neto
Federal Rural University of the Semiarid Region, Brazil

Rafael de Souza
Federal Rural University of the Semiarid Region, Brazil

Alex Sandro Gomes
Federal University of Pernambuco, Brazil

A volume in the Advances in Game-Based Learning (AGBL) Book Series

Information Science
REFERENCE
An Imprint of IGI Global

Chapter 18
Examining the Effectiveness of Hyperaudio Learning Environments

Joerg Zumbach
Paris Lodron University of Salzburg, Austria

Stephanie Moser
Technical University of Munich, Germany

ABSTRACT

The focus of instructional designers is increasingly shifting towards mobile devices such as smartphones or tablets and their use for learning purposes. These devices not only enable text-based but also audio-based instruction. This chapter presents the concept of hyperaudio, a special type of non-linear auditory learning environment. Contrary to other instructional devices such as hypertext, there is little research on learning with hyperaudio yet. In view of this scientific gap, this chapter aims at examining the effect of non-linear auditory presentation of information in more detail. To begin with, cognitive processes which are crucial for hyperaudio learning are examined. Then, some seminal studies investigating the design of hyperaudio learning environments and their influence on learning processes are presented. Results indicate that non-linear auditory information presentation is not always beneficial in terms of learning outcomes and cognitive load.

INTRODUCTION

The post-modern era has brought about a dramatic change in our requirements from education and information. From a subjective point of view, the need and urgency for retrieving specific information has increased significantly. As a result, learners demand instant access to learning resources and information required for problem-solving as well as a flexible approach to learning. The development of new media offers, amongst other things, various opportunities for such flexible ways of learning. This is supported from a technological perspective by an ever-increasing number of versatile handheld devices. Mobile

technologies such as smartphones are particularly well suited to meet these demands as they facilitate learning outside the bounds of time and place. Such equipment allows a fast retrieval of the right information at any given moment. While the place and time of learning were previously predefined by teachers, new media allow today's learners to decide where and when to learn (Weidenmann, 2001). Mobile learning scenarios, usually referred to as "m-learning", provide enormous potential for autonomous learning. In addition, current standards and software developments (e.g. data reduction) open up a wide range of applications. Consequently, a wide field of research on learning and instruction using mobile digital devices has been emerging in recent years (e.g. Hsi, 2003; Roschelle & Pea, 2002) as developing increasingly sophisticated mobile learning environments requires considerable experience and expertise. Particularly careful planning is required with regard to the uses of mobile devices in order to attract users and learners and to provide them with an integrated learning experience. Mobile technologies such as smartphones or tablets allow the presentation of both auditory and visual or audio-visual information. Given the wide range of technical possibilities available, finding the best way to present information and meet learners' needs still poses a challenge to instructional designers. This paper focusses on auditive information as an arrangement of information that is presented exclusively via the auditory channel. In particular, the dissemination of mobile devices such as MP3-players or smartphones, to name but two examples, has led to an intensified usage of auditive instruction in recent years (O'Bannon, Lubke, Beard & Britt, 2011; Vajoczki, Watt, Marquis & Holshausen, 2010).

This chapter analyses the influence of different formats of audio-based learning environments. Generally, linear presentation of audible information is widely used for learning purposes, e.g. in podcasts, audiobooks, or other instructional material exclusively designed for auditive instruction. In the following paragraphs, however, the authors would like to focus on a different format of information presented via the auditory channel, namely "hyperaudio". To begin with, a closer look will be taken at the concept of hyperaudio as a means of presenting non-linear auditive information. In the context of basic theoretical assumptions, the differences between hyperaudio and existing non-linear information media such as hypertext will be examined. In contrast to Hypertext, for example, there is hardly any research on hyperaudio yet, especially with regard to its use for learning purposes. This paper therefore aims to contribute to this field of research by providing an overview of theoretical and empirical findings regarding the design of hyperaudio learning environments. To this end, a brief definition of hyperaudio and a summary of relevant theoretical concepts such as Cognitive Load Theory (Paas, Renkl & Sweller, 2004; Sweller, 1994) or Cognitive Flexibility Theory (Spiro & Jehng, 1990) will be provided. Furthermore, some empirical findings regarding hyperaudio and related topics such as hypertext or hypervideo will be presented. In this context, different factors influencing self-paced knowledge acquisition by means of non-linear media will be analysed. These include content features such as text type and structure, modality of information presentation (visual or oral), and text arrangement and order (linear versus non-linear). According to research related to linear and non-linear learning environments, the above-mentioned factors are known to have an impact on learning outcomes and cognitive load as they are experienced by learners during the process of knowledge acquisition. The analyses presented in this paper seek to provide an understanding of the influence of different formats of information presentation on knowledge acquisition. Finally, some design principles for supporting knowledge acquisition in hyperaudio learning environments will be proposed.

HYPERAUDIO LEARNING ENVIRONMENTS

One advantage of new media devices such as contemporary handheld devices, e.g. so called "smartphones", can be seen in the opportunity that information can not only be presented visually, but also in an auditory manner. Listening to audio-based instructional material such as audiobooks or podcasts is therefore a common application of mobile devices for learning purposes. This audio material is mainly characterised by a linear sequence of orally presented material. Typical audiobooks are, in most cases, read-out versions of printed books, with the structure of both being similar. By contrast, non-linear information retrieval is less established within audio learning environments. This way of presenting audible instruction is what we refer to as "hyperaudio".

A Brief Definition of Hyperaudio

The arrangement of and access to information within hyperaudio learning environments corresponds to that of hypertext and hypermedia environments, which are based on locally coherent single nodes, i.e. single audio files. These so called "audio nodes" are connected by means of hyperlinks which enable users to navigate within the hyperaudio environment (Zumbach & Kroeber, 2006). Navigating within a hyperaudio learning scenario should contribute to global coherence building and the understanding of relationships between single audio nodes (Zumbach, 2006). The ways in which authors can kink nodes, and thus define the overall navigation structure within a hyperaudio environment, are similar to designing a hypertext or hypermedia environment in that they allow linear, hierarchical, elaborative, or associative structures (Grabinger & Dunlap, 1996).

A specific type of linear hyperaudio environment is commonly known as the audiobook. Audiobooks are narrated versions of printed books that do not have to meet the requirements of local coherence within single audio nodes. The structure of audiobooks is the same as the structure of their printed counterparts. This paper focusses on non-linear auditory information retrieval, which is further described below. Although the concept of hyperaudio is relatively new, the underlying principles are not: non-linear navigation in voicemail boxes or automated helpdesk systems, for instance, are common used in everyday life. A closer look on research related to (non-)linear text-based instruction and audible instruction will be taken below. Subsequently, some conclusions for designing hyperaudio learning environments will be drawn.

Cognitive Aspects of Auditory and Text-Based Instruction

Information presented via mobile technologies may consist in text-based visual information, auditory information or a multimodal presentation combining both. For a long time, audible information was the most common way of providing instruction (Barron, 2004). The cognitive aspects of processing visual and audible information are different and require thorough examination.

In cognitive science, separate Working Memory (WM) instances are thought to be responsible for processing phonological information and visual information. According to Baddeley's (1992; 1998; 2012) WM model, audible information is first processed in a subsystem called phonological loop. Here, auditive information is stored and briefly repeated before being processed within an episodic buffer. By contrast, visual information is represented in a visuo-spatial sketchpad before being processed within the

episodic buffer. While the model provided by Baddeley has gained considerable attention in cognitive research and within the field of multimedia learning (e.g. Mayer, 2014), more recent approaches have criticised it as being codality-specific rather than explaining the differences between the ways in which information presented via different modalities is processed (Rummer, Schweppe, Scheiter & Gerjets, 2008; Schüler, Scheiter, Rummer & Gerjets, 2012). However, empirical evidence suggests that either visual-verbal or auditory-verbal information is processed differently within the human working memory, regardless of whether this difference results from a different modality or codality.

Another theory that addresses the issue of working memory resources involved in processing learning material is Cognitive Load Theory (CLT). In particular, Cognitive Load Theory provides a common approach for examining learning processes across different media (Paas, Renkl & Sweller, 2004; Sweller, 1994). In its origins, CLT assumed that there were three different types of cognitive load: Intrinsic Cognitive Load (ICL), Extraneous Cognitive Load (ECL), and Germane Cognitive Load (GCL). These loads were assumed to be accumulative and result in the total amount of mental effort in the WM. Direct information processing results in ICL and is unavoidable. The more complex the topic, the higher the ICL during mental processing. GCL is necessary for the activation of schematic representations and deeper processing of information. More recent approaches to CLT no longer differentiate between GCL and ICL. Paas and Sweller (2014) define GCL as those working memory resources which are assigned to dealing with ICL. GCL can thus be regarded as free memory resources "(…) devoted to intrinsic cognitive load minus the resources devoted to extraneous cognitive load." (Paas & Sweller, 2014, p. 38). The main reason for merging these two kinds of cognitive load rests on the assumption that the activation of prior knowledge is necessary for both kinds of information processing. Chunking processes involved in determining the amount of ICL as well as activation of prior knowledge, as is assumed to be the case in the activation of GCL, both require the activation of schematic representations in the long-term memory and are therefore difficult to distinguish. Nevertheless, one basic assumption of CLT in its original version still affects the design of learning material and superordinate learning environments: instructional design should be conducted carefully in order to avoid situations which lead to an increased Extraneous Cognitive Load and have a negative impact on learning processes and outcomes (van Merrienboer & Kester, 2014). As the design of the learning material determines ECL, existing guidelines derived from research for reducing ECL should be taken into account in instruction (Mayer & Moreno, 2003). CLT is frequently used when examining different forms of information presentation to determine the best way to facilitate learning. This also includes information presented in a visual or auditory manner.

The first systematic research on audible instruction compared to text-based instruction dates back to the beginning of the 20th century (Barron, 2004). Initial comparative studies seemed to show a general advantage of auditive instruction over text-based instruction (Golas, Gorr & Yao, 1994; Nugent, 1982). However, these early investigations were fraught with constraints with regard to the behaviour and role of the reader as well as the underlying study material (e.g. differences in reading ability, different text formats, etc.). Additionally, shortcomings in the methods of data analysis weakened the results. As Travers (1970, quoted after Barron, 2004, p. 954) stated: "One cannot reasonably ask the general question, whether the eye or the ear is more efficient for the transmission of information, since clearly some information is better transmitted by one sensory channel than by another". In line with this, and following Barron (2004), contemporary research still presents ambiguous research outcomes. Recent publications in the field of information processing assume that text and audio are likely to be represented in the phonological subsystem of the WM only (Rummer, Schweppe, Scheiter & Gerjets, 2008). Verbal

information is processed within the phonological loop, independent of its modality (written or spoken language). If Rummer et al. (2008) are correct in their assumption that the working memory is rather codality-specific, then we should not expect any differences between reading and listening comprehension. And yet, several studies comparing learning with written or auditive verbal information have found differences between the two modalities.

Firstly, there seem to be slight advantages for auditive information. Visual information is more likely to contain irrelevant and distracting stimuli. These stimuli may have little sematic meaning but still have to be processed, e.g. word or number lists (Kürschner & Schnotz, 2008; Pächter, 1996). Using the terms of CLT, this additional information that has to be processed leads to an increased extraneous cognitive load, because it does not really contribute to the genuine information processing process, and thus, to learning.

The proprietary design of each modality might strongly influence the processing of both text and audio. Several design features within a text, such as headings, paragraphs, highlighting, etc. can support reading comprehension processes. These supportive features are not easily transferrable to spoken language (Hartley, 2004). Additionally, during reading, self-pacing facilitates rehearsal strategies when encoding text-based information. When encoding information presented by auditory means, these processes are not available to the listener.

Nevertheless, similar to text-based information, spoken language also has a repertoire of stylistic features including speed control, the use of breaks, emphasis on certain information, etc., which support comprehension processes. In addition, these features are usually not available in written language (Barron, 2004, Kürschner & Schnotz, 2008; of course it is possible to transcribe such utterances from spoken language into written language, but this is not common, except for research purposes, for example). Furthermore, voice carries paralinguistic personality cues such as intonation, which are referred to the speaker. Thus, spoken language contains information in addition to the message itself (Nass & Lee, 2001; Nass & Brave, 2005).

Text-type might also play a major role. There seem to be advantages of written learning material for memorizing details while audible presentation seems to contribute to a better understanding of overarching concepts or ideas such as a plot (e.g. Hildyard & Olson, 1978; Kürschner, Schnotz & Eid, 2006). When processing complex texts, comprehension of the text might be better supported by written information. The basic reason for the advantage of this codality is related to the possibility of self-pacing and the application of specific reading strategies with a focus on memorizing details.

It also seems that the auditory presentation of information puts greater strain on working memory capacity, especially when texts are longer or more complex or contain redundant information. These processes might explain greater differences between written and spoken language, as emphasized by Leahy and Sweller (2011) in their discussion of the transient information effect. The transient information effect occurs when information stored in the WM is replaced by newly acquired information before it has been adequately processed. This might be the case when the aggregation of both information sources exceeds working memory capacity. While readers will find it easy to return to a certain paragraph within a text and reload specific information into WM, audible information remains transient (Singh, Marcus & Ayres, 2012). It is much more difficult for a listener to navigate within auditory instruction. Thus, visual information might contribute to a decreased cognitive load, because information retrieval is simplified. According to Tindall-Ford, Chandler, and Sweller (1997) this might contribute to better overall comprehension.

In summary, research suggests varying results in terms of how information presented either visually or by auditive means is processed. There is currently a lack of insights into how some of these results might interact with other design aspects of learning material, in particular with regard to the dimensions of linearity and non-linearity of learning resources. These dimensions are analyzed within the context of linear and non-linear audio below. Given the above-mentioned research gap, inferences from research within the area of hypertext and hypermedia environments are drawn.

Linear vs. Non-Linear (Audio) Learning Environments

In the following paragraphs, the distinction between linear media, such as "traditional" textbooks or time-dependent media (films, animations), and non-linear media, such as hypertext, hypervideo, or hyperaudio, will be examined. The investigation of linear and non-linear information presentation has become an established field of research in learning sciences and cognitive psychology (Zumbach, 2006).

Information retrieval in hypermedia systems works via interactive cross-references. Those references are called "hyperlinks" and are visible within the content nodes on the display of the digital device. Nodes can be linked in different ways, especially in the case of linear and non-linear content. The difference between linear information presentation and non-linear information presentation lies in the way in which data is accessed by its users. While linear information can only be accessed in a predefined manner (in digital learning environments usually with "next" and "back" buttons), non-linear information can be retrieved freely. Non-linear hyperlinks can connect different information resources in a hierarchical, elaborative, or associate manner, etc. (Chen & Rada, 1996; Grabinger & Dunlap, 1996). Non-linear information retrieval is well-established in hypertext learning environments. Popular examples are Wikipedia articles where users are able to navigate between relevant information via hyperlinks. Compared to hypertext environments, non-linear audio environments are less common. In addition, hyperaudio leaves the linear navigation level. For the purpose of this paper, it is defined as an audio system that is characterised by linked, audio-based nodes which are organised in a non-linear way. Even though hyperaudio environments are comparable to well-established formats such as hypertext, hypermedia, or hypervideo (a special kind of interactive and hyperlinked video format; e.g. Zahn, Schwan & Barquero, 2002), there is hardly any research on navigation strategies within them. Thus, theoretical and empirical considerations concerning learning with hypervideo in comparison to linear audio are required. This is also crucial because cognitive mechanisms underlying information processing are assumed to be different when using non-linear audio or text as opposed to using the visual information processing channel only.

Hyperaudio is "an arrangement of auditorially presented material represented within locally coherent hyperlinked nodes" (Zumbach & Schwartz, 2014, p. 356). This means that several hyperlinks connect different "audio nodes" and enable users to navigate freely within a hyperaudio learning environment (Zumbach, 2006). Users should be able to understand the relationships between those single audio nodes by accessing them via the hyperlinks. This should lead to a better understanding of the single nodes and users should develop a sense of global coherence.

Contrary to linear audio, users in non-linear audio learning environments do not have to visit audio nodes in a predefined manner. Non-linear audio is prevalent for example in telecommunications, e.g. in automated helpdesk systems or voicemail boxes: "If you would like to retrieve your new message, please press 1", etc. Such voice-based menus are typical examples for hierarchically structured hyperaudio systems. However, in associative learning systems, the choice of sequence is up to the learners. Learners can decide with which kind of information they would like to start with, which kind of information they

would like to expand on and which information they would like to skip. One example in this respect are audio-based services for sightseeing purposes, as they offer (sometimes location-based) audio-documents on demand explaining more about exhibited objects or places of interest. Those systems frequently follow a rather simple structure: each audio document has to be navigated separately by specifically selecting it, or is automatically played (e.g. by means of an infrared sensor or a QR-code). More advanced systems can be based on interconnected hyperlinks, for example. This enables learners to define their own focus of interest by simply following their chosen navigation options. If a learner uses an MP3 player to listen to an audio file from the field of molecular biology explaining bacteria, for example, the authors of the document may have linked it to several other audio files about such topics as shapes of bacteria, life conditions, viruses, etc. (see Figure 1).

While listening to the document, the learner can see the links on the display of the player. If any of the documents catches the learner's interest, he or she can simply activate the link by pressing a button on the display. A new audio file opens and the learner can listen to the requested information, e.g. on viruses (see Figure 2). This new file will, in turn, be linked to other audio files, and so on. By this means, each learner creates his or her own learning experience.

One advantage of non-linear learning environments can therefore be seen in the opportunity that the learners can navigate freely within such learning environments. But research on learning with regard to hypertext has shown that this is not always beneficial. In some cases, users might not be able to enjoy the full value of the information provided by the non-linear system.

Deeper insights into these findings can be gained by referring back to the human WM (Baddeley, 1992; 1998; 2012), which is not only essential for visual and verbal information processing but also in terms of linear and non-linear information presentation. As described in CLT (Sweller, 1994; see

Figure 1. Example of non-linear audio system (page about bacteria with three other audio files and home button)
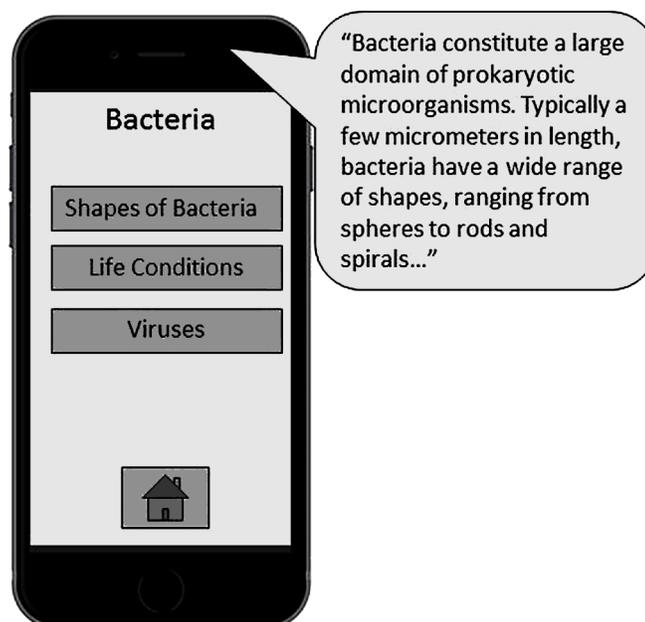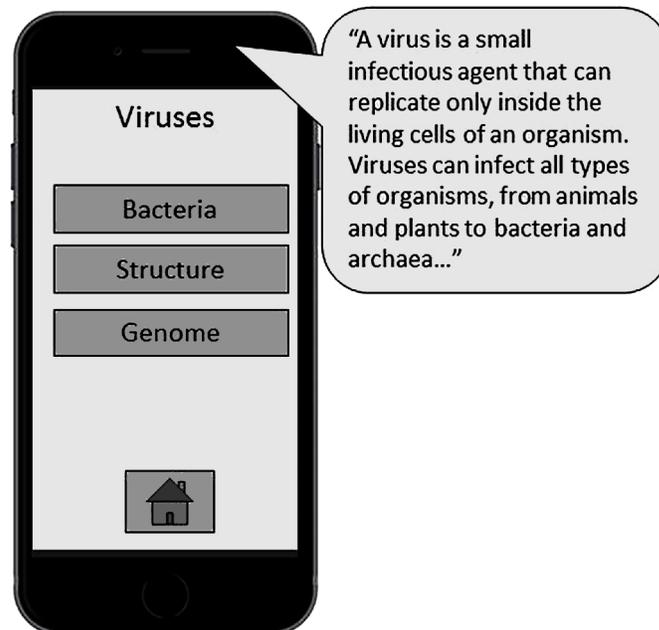
Figure 2. Example of non-linear audio system (page about viruses with three other audio files and home button)

above), the WM has limited resources for processing information. In the field of hypertext learning, much research related to CLT addresses problems like the "Lost-in-Hyperspace" phenomenon or the problem of cognitive overhead (Conklin, 1987). The former describes a feeling of disorientation within non-linear digital environments. Users of such systems may find themselves "lost" within the hypertext environment or might not be able to identify where they are within the learning environment, or be unable to retrieve the information they originally wanted to navigate to. In these cases, the hypersystem might seem like a kind of a maze to these users. The second issue addressed here, namely cognitive overhead, refers to the amount of information that has to be presented and processed within the working memory. Users require additional mental resources and effort to navigate through the hypermedia system. This additional cognitive load, which can be regarded as ECL, is derived from navigation planning. For each hyperlink presented, a learner must decide whether or not following this hyperlink might be useful. This cognitive activity takes place at the same time as information processing itself. Niederhauser, Reynolds, Salmen, and Skolmonski (2000) found that, within hypertext learning environments, linear information retrieval leads to better learning outcomes than non-linear information retrieval. It seems that navigation planning not only causes additional cognitive load within hypertext environments, but also affects learning with hypervideo or hyperaudio, because learners must simultaneously process information and plan their navigation (Bannert, 2004; Gerdes, 1996).

Nevertheless, some research findings postulate that there are also positive effects of non-linear learning environments. Cognitive Flexibility Theory (CFT; Spiro & Jehng, 1990) states that non-linear information retrieval is beneficial in complex (ill-structured) domains. Furthermore, there appears to be an interaction between type of text and text presentation format. The increased mental effort caused by navigation planning in learning environments with non-narrative, encyclopaedia-type texts might

contribute to deeper information processing and might thus likewise increase GCL (Zumbach, 2006; Zumbach and Mohraz, 2008).

When comparing linear and non-linear text-based learning environments with regard to learning outcomes, the meta-analysis provided by Chen and Rada (1996) reveals comparable outcomes between both media or even slight advantages for hypertext. Shapiro & Niederhauser, however, state that "(…) a number of studies have not shown such effect (…)" (2004, p. 609). Thus, non-linear information media might promote but also impede learning for various reasons. This might also affect learning with hyperaudio, because the main differences in the above-mentioned research concern the navigability of these systems (i.e. the differences between linear and non-linear information presentation). In the following paragraphs, research concerned exclusively with learning in hyperaudio environments will be examined more closely.

RESEARCH ON HYPERAUDIO

The research findings outlined above paint an ambiguous picture with regard to information presented either visually or by audible means, especially when leaving the linear sequencing level and using non-linear information presentation. As mentioned before, there is a significant shortage of research on interacting with and learning from non-linear auditory text. This may be due to the fact that the development and distribution of such instructional devices is also less common. Hyperaudio as such is not a new invention: first publications on the use of non-linear audio as an instructional device date back to the 1990s. In a paper presented by Sawhney and Murphy (1996), the authors introduce the *Espace 2* system, a learning environment allowing users to access narrative content presented on several Compact Discs. Users were given the opportunity to navigate through hierarchically structured content, organized within so called "acoustic bubbles" which represented what we would refer to audio nodes. The interface itself was represented on a computer screen. The authors' evaluation was limited to an analysis of the usability of the system, whose outcomes revealed that users experienced the new way of navigating within a multi-auditory information system as difficult. Especially the option of listening to more than one auditive source "…caused some cognitive overload and confusion" (Sawhney & Murphy, 1996, p. 106).

One of the first mobile applications for hyperaudio learning was presented by Petrelli, Not, Sarini, Stock et al. (1999). The authors present an intelligent location-based system for guiding visitors through a museum by providing them with Personal Digital Assistants (PDAs) which enable them to obtain auditory information on a nearby exhibit combined with on-screen hypertext. Upon returning to an object, users are presented with new information that has not been presented to them previously. Navigation through this auditive hyperspace takes place by means of the user's real world movements. In their examination of the system, the authors focus on improving the adaptivity of the system. Their hyperaudio approach relates to an adaptive hypermedia system that is able to incorporate users' characteristics into the information presentation process (see also Petrelli & Not, 2005). However, the authors do not provide any evaluation of their approach, neither in terms of its effectiveness nor with regard to its usability. Especially within the domain of museum and exhibition guides, some additional developments exist, which use non-linear access to auditive information. Zimmermann, Lorenz, and Specht (2003) also present such an adaptive system for supporting visitors of a 3D arts museum. While the adaptivity of these systems was mainly related to cognitive issues, educational characteristics, and interests, more recent developments also tried to incorporate emotion and affect (Yii Lim & Aylett, 2009).

Most of the approaches and developments described so far predominantly refer to the technological aspects of designing hyperaudio learning environments, in particular location-based services in museums, and do not adequately address the issue of how learners can really interact with them and benefit from them.

A study by Zumbach and Schwartz (2014) focusses on exactly these issues. The authors compared the presentation of information by means of text and by audible means (modality) in both a linear and a non-linear fashion (presentation format) in a controlled and a randomized experiment. In addition, the authors examined the interaction of the presentation format with different underlying text types. The nature of the learning material (text type) was either an expository structure on molecular biology or a narrative structure with a linear fairytale plot. The two text types had a different process focus: acquiring individual concepts and details versus promoting holistic text processing. This appears to be an important issue: Zumbach and Mohraz (2008), for example, found a strong interaction between text type and information presentation (linear versus non-linear retrieval) in hypertext learning. Both texts in the present study on hyperaudio comprised locally coherent nodes and the content of each node was understandable in itself. The texts were either presented in a linear or in a non-linear fashion. The different versions were either presented as on-screen text or as an audio version using a simulated cell phone. Results of this experiment revealed an overall advantage of linear information presentation with regard to memorization of single facts as operationalized by a specific multiple choice test. The authors did not find the same result with regard to the understanding of the plot as measured by an essay task. Overall, non-linear information presentation led to an increase in cognitive load. This is consistent with prior research results on hypertext (Zumbach & Mohraz, 2008) and indicates that navigation planning in non-linear learning environments might be a secondary task that increases ECL (like in text-hypertext comparisons; Wells and McCrory, 2011). These results also correspond to findings from hypertext and text environments (Chen & Rada, 1996; Shapiro & Niederhauser, 2004). Furthermore, written text was found to provide more advantages than auditory text. At first glance, this contradicts the findings from prior research on listening versus reading (Golas, Orr & Yao, 1994). Zumbach and Schwartz (2014), however, attribute this result to a media-specific appraisal effect: learners may have judged the audible learning environment as more straightforward and may thus have invested less mental effort. Learners in the text-based condition, on the other hand, may have considered the learning environment as more difficult and may therefore have invested more mental effort. This is also consistent with other research (e.g. Salomon, 1984). According to Geary (2004; 2011) learners might experience the learning process with auditory material as straightforward because listening is a primary biological skill which requires little use of the working memory. Reading, by contrast, is a secondary skill which may be considered as more difficult. In addition, the poorer performance of learners in the auditive condition could be explained by means of the transient information effect (Leahy & Sweller, 2011; Singh, Marcus & Ayres, 2012; Leahy, Marcus & Sweller, 2012). This effect is caused by the fact that longer and complex sentences presented verbally might exceed the capacity of the WM. Thus, different revision strategies might result in differences between reading and listening (Kürschner et al., 2006; Rummer et al., 2008). One of the most interesting outcomes of the study by Zumbach and Schwartz (2014) was that the text type did affect learning outcomes: the expository text led to comparable learning outcomes in linear and in non-linear formats, while the narrative text type resulted in lower learning outcomes when presented in hyperaudio or hypertext format. Furthermore, the study showed that hyperaudio led to better learning

outcomes than linear audio if the presented text had a non-linear structure rather than a linear plot-like structure. Presenting the narrative text type in a non-linear fashion led to decreased learning outcomes in both test scenarios as well as an increased cognitive load. As a consequence, the results of this study suggest that different representations (text versus audio and linear versus non-linear) are affected by the nature of the learning material. Zumbach and Mohraz (2008) indicated that a text with a primarily linear structure and a linear plot increases cognitive load when presented in a non-linear fashion. In this case, the load appears to be ECL as navigation planning might represent an additional secondary task. Presenting linear text types either as hypertext or hyperaudio is likely to lead to poorer performance in terms of knowledge acquisition and understanding.

The study mentioned above revealed that, in most cases, hyperaudio leads to decreased learning performance compared to the written codality in a controlled experimental scenario. Under certain conditions, however, the non-linear auditory format led to comparable outcomes to those achieved by means of linear audio or text, especially when the underlying text type was not based on a linear structure. Nevertheless, these findings are sobering as they reveal that hyperaudio as a standalone instructional device is not as effective as other ways of presenting information. In order to support hyperaudio as an instructional device, subsequent research was conducted in order to promote knowledge acquisition through non-linear auditive information presentation. Zumbach and Moser (2013) present two studies examining the influence of learning scaffolds to support learning with hyperaudio. In a first study, summaries were used as tool for triggering coherence formation processes when navigating either linear or non-linear auditive information systems. For this purpose, the content (related to moors as ecosystems) was divided into subsequent chapters which were either navigable in linear or in a non-linear fashion by means of associatively linked audio nodes. In both cases, learners were given a short summary of the content beforehand. Results revealed no significant impact of this instructional device: learners within both conditions (linear audio and hyperaudio), with or without summaries, achieved a comparable performance in a subsequent knowledge test. Learners in the non-linear conditions reported a higher cognitive load than learners in the linear conditions, albeit not significantly higher. In a second experiment, the authors examined the influence of task orientation on learning with either linear text, linear audio, or hyperaudio. Learners from one condition had to navigate one of the three learning environments in order to retrieve information required for solving a pre-defined problem (e.g. "You are a Medical Doctor in an emergency room. Your next patient requires blood. His blood type is B but you only have blood preservation type A, AB and 0. What do you do? Explain!"). In the other conditions, no such task was given. Results revealed a significant impact of the pre-defined problem-solving task on increased learning performance across all three conditions compared to the conditions with no such given task. In addition, linear audio led to a better overall performance than linear text or hyperaudio in this experiment. Furthermore, the hyperaudio group reported the highest cognitive load.

The overall research findings (for an overview see also Table 1) on the subject of hyperaudio described in this paper suggest that the development and use of this non-linear auditory approach is still rather a niche than a well-investigated domain. Most research from a technological point of view is dedicated to the use of mobile devices as adaptive guides for museums, whereas there is a research gap concerning the use of hyperaudio and its implications for knowledge acquisition as well as its impact on other psychological parameters. With regard to genuine psychological learning research, the above-mentioned studies reveal that hyperaudio is, in most cases, a rather weak instructional device which leads to poorer

Table 1. Overview of relevant cited studies

Authors	Hyperaudio System	Research Focus	Important Findings	Advantages/Disadvantages
Sawhney & Murphey (1996)	- *Espace 2* computing system - Users can navigate within hierarchically linked „acousic bubbles"	- Usability of the hyperaudio system	- Users experienced navigation as difficult	+ Hyperaudio as rather new form of learning/information system + Several shortcomings of hyperaudio sytems and design suggestions are presented - No investigation of learning processes - No detailed study description
Petrelli, Not, Sarini, Stock, Strapparava & Zancanaro (1999)	- Personal Digital Assistants (PDAs) - HyperAudio (context-sensitive adaptive museum guide) - Users can freely move	- Adaptivity of the system	- System supports different forms of adaptivity via content selection and organization, language style, linguistic choices, navigation suggestions	+ Detailed description of different forms of adaptivity - No evaluation of the system (e.g. regarding learning)
Petrelli & Not (2005)	- PDAs - HyperAudio (context-sensitive adaptive museum guide) - Users can freely move	- User requirements - Measured via survey	- Individual differences help improving adaptivity, e.g. differences in usage between frequent and first time visitors - social environment influences usage	+ Investigation of user needs + Development/Improvement of an adaptive environment + Design suggestions regarding adaptivity - No investigation of learning processes
Zumbach & Moser (2013)	- Learning environment (simulated cellular phones) - Users can navigate between audio nodes	- Scaffolding hyperaudio learning - Cognitive Load - Comparison of linear and non-linear text an audio - Task orientation (users had either to solve a pre-defined problem or not)	- Linear audio and hyperaudio as well as conditions with or without summaries led to comparable learning outcomes - Users learning with pre-defined problems showed better learning results - Higher cognitive load in non-linear conditions	+ Investigation of learning processes + Experimental study design: comparison between different forms of (hyper-)media - Rather simple hyperaudio system
Zumbach & Schwartz (2014)	- Learning environment (simulated cellular phones) - Users can navigate between audio nodes	- Comparison between linear and non-linear audio and on-screen text - Interaction between different text types - Cognitive Load	- Advantage of linear format with regard to memorization of single facts - Non-linear presentation led to an increase in cognitive load	+ Investigation of learning processes + Cognitive load perspective + Experimental study design: comparison between different forms of (hyper-)media - Rather simple hyperaudio environment

knowledge acquisition and an increased cognitive load than linear text or linear audio, for instance. However, in some circumstances, for example with a loosely structured text type and a complimentary task or problem to be solved by means of the learning resources provided, hyperaudio can be as effective as other instructional devices. Future research is required in order to identify fields in which the use of non-linear audible information might be a highly appropriate, or indeed the most appropriate, way of delivering and retrieving information.

FUTURE RESEARCH DIRECTIONS

Research and development of hyperaudio environments is still in its infancy. A huge body of research will have to be conducted in order to develop new technologies for teaching and learning, for instance in the fields of hypertext and hypermedia learning. Developments such as the transfer from adaptive hypermedia to adaptive hyperaudio have already been made. Nevertheless, this process should be continued and rigorously evaluated. The authors of this paper support the opinion that such prototypes should also be analyzed with regard to their objective effectiveness and not only with respect to usability. This might also open up a wide range of opportunities for developing truly helpful adaptivity features in order to meet learners' genuine needs. In addition, it may be worth transferring technologically advanced approaches based on smartphones or similar devices as have already been rolled out in museums to a broader context and to enable open content. Based on this idea, authorship of such open (and maybe) adaptive hyperaudio systems is another key topic to be addressed in future research. Similar to hypertext and hypermedia environments, it might be an interesting approach to have users or learners co-construct their own non-linear auditory information resources. This might not only be interesting from a of co-construction perspective, but also from a literature and arts-related point of view, as it would foster true interdisciplinary collaboration beyond the domains of education and technology.

Another major research area would be to examine ideal combinations of content, text type, and the use of hyperaudio. As demonstrated by current research, hyperaudio appears to be rather disadvantageous as an instructional device for teaching and learning purposes compared to written text and linear audio. Two lines of future research could be derived from this, the first of which could be dedicated to the analysis of scenarios where hyperaudio might be rather beneficial for learning, e.g. for complex learning within mobile learning scenarios referring to loosely structured content. This could include the development of supportive tools and scaffolds. The second line of research could focus on analyzing the long-term effects of hyperaudio as a daily, weekly or monthly tool within broader learning environments. All of the above-mentioned studies used ad-hoc interventions which were limited in time and content. By becoming more and more familiar with hyperaudio, disadvantages of this medium known from current research may become obsolete.

CONCLUSION

Learning is increasingly becoming an activity which is independent of place and time. Mobile technologies and location-based services take this trend into account. Hyperaudio appears to be an interesting alternative to other existing instructional devices for formal and informal learning processes. Research from a technological point of view as well as from the perspective of learning with hyperaudio is still in its infancy and might emerge with new applications. All in all, we suggest that, from a current research point of view, additional support is necessary in order to enhance learning from hyperaudio environments. Additional instructions might, for example, be helpful to increase mental effort in information processing and might consequently contribute to learning performance. Research regarding non-linear auditory instruction is still emerging and must also incorporate new technologies and instructional developments.

REFERENCES

Baddeley, A. D. (1992). Working Memory. *Science*, *255*(5044), 556–559. doi:10.1126/science.1736359 PMID:1736359

Baddeley, A. D. (1998). *Human Memory; Theory and Practice*. Boston: Allyn and Bacon.

Baddeley, A. D. (2012). *Working Memory: Theories, Models, and* Controversies. *Annual Review of Psychology*, *63*(1), 1–29. doi:10.1146/annurev-psych-120710-100422 PMID:21961947

Bannert, M. (2004). Designing metacognitive support for hypermedia learning. In H. Niegemann, D. Leutner, & R. Brünken (Eds.), *Instructional design for multimedia-learning* (pp. 19–30). Münster: Waxmann.

Barron, A. E. (2004). Auditory Instruction. In D. H. Jonassen (Ed.), *Handbook of research on educational communications and technology* (pp. 949–978). Mahwah, NJ: Lawrence Erlbaum.

Chen, C., & Rada, R. (1996). Interacting with hypertext: A meta-analysis of experimental studies. *Human-Computer Interaction*, *11*(2), 125–156. doi:10.1207/s15327051hci1102_2

Conklin, J. (1987). Hypertext: An introduction and survey. *IEEE Computer*, *20*(9), 17–41. doi:10.1109/MC.1987.1663693

Geary, D. C. (2004). Evolution and cognitive development. In R. Burgess & K. MacDonald (Eds.), *Evolutionary perspectives on human development* (pp. 99–133). Thousand Oaks, CA: Sage Publications.

Geary, D. C. (2011). Evolutionary Educational Psychology. In K. Harris, S. Graham, & T. Urdan (Eds.), *APA educational psychological handbook* (Vol. 1). Washington, DC: American Psychological Association.

Gerdes, H. (1996). *Lernen mit Text und Hypertext.* Lengerich: Pabst.

Golas, K. C., Orr, K. L., & Yao, K. (1994). Storyboard development for interactive multimedia learning. *Journal of Interactive Instruction Development*, *6*, 18–31.

Grabinger, S., & Dunlap, J. C. (1996). Nodes and Organization. In P. A. M. Kommers, S. Grabinger, & J. C. Dunlap (Eds.), Hypermedia Learning Environments (pp. 79–114). Mahwah, NJ: Lawrence Erlbaum.

Hartley, J. (2004). Designing instructional and informational text. In D. H. Jonassen (Ed.), *Handbook of research on educational communications and technology* (pp. 917–947). Mahwah, NJ: Lawrence Erlbaum.

Hildyard, A., & Olson, D. R. (1978). Memory and inference in the comprehension of oral and written discourse. *Discourse Processes*, *1*(2), 91–117. doi:10.1080/01638537809544431

Hsi, S. (2003). A study of user experiences mediated by nomadic web content in a museum. *Journal of Computer Assisted Learning*, *19*(3), 308–319. doi:10.1046/j.0266-4909.2003.jca_023.x

Jacobson, M. J., & Spiro, R. J. (1995). Hypertext learning environments, cognitive flexibility, and the transfer of complex knowledge: An empirical investigation. *Journal of Educational Computing Research*, *12*(4), 301–333. doi:10.2190/4T1B-HBP0-3F7E-J4PN

Kürschner, C., & Schnotz, W. (2008). Das Verhältnis gesprochener und geschriebener Sprache bei der Konstruktion mentaler Repräsentationen. *Psychologische Rundschau*, *59*(3), 139–149. doi:10.1026/0033-3042.59.3.139

Kürschner, C., Schnotz, W., & Eid, M. (2006). Konstruktion mentaler Repräsentationen beim Hör- und Leseverstehen. *Zeitschrift für Medienpsychologie*, *18*(2), 48–59. doi:10.1026/1617-6383.18.2.48

Leahy, W., & Sweller, J. (2011). Cognitive load theory, modality of presentation and the transient information effect. *Applied Cognitive Psychology*, *25*(6), 943–951. doi:10.1002/acp.1787

Mayer, R. E. (Ed.). (2014). *The Cambridge handbook of multimedia learning*. New York: Cambridge University Press. doi:10.1017/CBO9781139547369

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, *38*(1), 43–52. doi:10.1207/S15326985EP3801_6

Nass, C., & Brave, S. (2005). *Wired for speech: How voice activates and advances the human-computer relationship*. Cambridge, MA: The MIT Press.

Nass, C., & Lee, K. (2001). Does computer-synthesized speech manifest personality? Experimental tests of recognition, similarity-attraction, and consistency-attraction. *Journal of Experimental Psychology. Applied*, *7*(3), 171–181. doi:10.1037/1076-898X.7.3.171 PMID:11676096

Niederhauser, D. S., Reynolds, R. E., Salmen, D. J., & Skolmoski, P. (2000). The influence of cognitive load on learning from hypertext. *Journal of Educational Computing Research*, *23*(3), 237–256. doi:10.2190/81BG-RPDJ-9FA0-Q7PA

Nugent, G. C. (1982). Picture, audio, and print: Symbolic representations and effect on learning. *Educational Communications and Technology*, *30*, 163–174.

O'Bannon, W., Lubke, J., Beard, J., & Britt, V. (2011). Using podcasts to replace lecture: Effects on student achievement. *Computers & Education*, *57*(3), 1885–1892. doi:10.1016/j.compedu.2011.04.001

Paas, F., Renkl, A., & Sweller, J. (2004). Cognitive Load Theory: Instructional Implications of the Interaction between Information Structures and Cognitive Architecture. *Instructional Science*, *32*(1/2), 1–8. doi:10.1023/B:TRUC.0000021806.17516.d0

Paas, F., & Sweller, J. (2014). Implications of Cognitive Load Theory for Multimedia Learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (2nd ed.; pp. 27–42). Cambridge, UK: Cambridge University Press. doi:10.1017/CBO9781139547369.004

Pächter, M. (1996). *Auditive und visuelle Texte in Lernsoftware*. Münster: Waxmann.

Petrelli, D., & Not, E. (2005). User-centred design of flexible hypermedia for a mobile guide: Reflections on the hyperaudio experience. *User Modeling and User-Adapted Interaction*, *15*(3-4), 303–338. doi:10.1007/s11257-005-8816-1

Petrelli, D., Not, E., Sarini, M., Stock, O., Strapparava, C., & Zancanaro, M. (1999, May). HyperAudio: location-awareness+ adaptivity. In *CHI'99 Extended Abstracts on Human Factors in Computing Systems* (pp. 21–22). New York: ACM. doi:10.1145/632716.632732

Roschelle, J., & Pea, R. (2002). A walk on the WILD side: How wireless handhelds may change CSCL. *International Journal of Cognitive Technology*, *1*, 145–168. doi:10.1075/ijct.1.1.09ros

Rummer, R., Schweppe, J., Scheiter, K., & Gerjets, P. (2008). Lernen mit Multimedia: Die kognitiven Grundlagen des Modalitätseffekts. *Psychologische Rundschau*, *59*, 98–107.

Salomon, G. (1984). Television is ''easy'' and print is ''tough'': The differential investment of mental effort in learning as a function of perceptions and attribution. *Journal of Educational Psychology*, *76*(4), 647–658. doi:10.1037/0022-0663.76.4.647

Sawhney, N., & Murphy, A. (1996). ESPACE 2: An Experimental Hyperaudio Environment. In M. J. Tauber (Ed.), *CHI Conference Companion on Human Factors in Computing Systems* (pp. 105-106), New York: ACM.

Schüler, A., Scheiter, K., Rummer, R., & Gerjets, P. (2012). Explaining the modality effect in multimedia learning: Is it due to lack of temporal contiguity with written text and pictures? *Learning and Instruction*, *22*(2), 92–102. doi:10.1016/j.learninstruc.2011.08.001

Shapiro, A., & Niederhauser, D. (2004). Learning from hypertext: Research issues and findings. In D. H. Jonassen (Ed.), *Handbook of research on educational communications and technology* (2nd ed.; pp. 605–620). Mahwah, NJ: Lawrence Erlbaum Associates.

Singh, A.-M., Marcus, N., & Ayres, P. (2012). The transient information effect: Investigating the impact of Segmentation on spoken and written text. *Applied Cognitive Psychology*, *26*(6), 848–853. doi:10.1002/acp.2885

Spiro, R. J., & Jehng, J. C. (1990). Cognitive flexibility and hypertext: Theory and technology for the nonlinear and multidimensional traversal of complex subject matter. In D. Nix & R. Spiro (Eds.), *Cognition, education and multimedia: Exploring ideas in high technology*. Hillsdale, NJ: Lawrence Erlbaum.

Sweller, J. (1994). Cognitive Load Theory, Learning Difficulty and Instructional Design. *Learning and Instruction*, *4*(4), 295–312. doi:10.1016/0959-4752(94)90003-5

Tindall-Ford, S., Chandler, P., & Sweller, J. (1997). When two sensory modes are better than one. *Journal of Experimental Psychology. Applied*, *3*(4), 257–287. doi:10.1037/1076-898X.3.4.257

Vajoczki, S., Watt, S., Marquis, N., & Holshausen, K. (2010). Podcasts: Are they an effective tool to enhance student learning? A Case Study. *Journal of Educational Multimedia and Hypermedia*, *19*, 349–362.

Van Merrienboer, J., & Kester, L. (2014). The Four-Component Instructional Design Model: Multimedia Principles in Environments for Complex Learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (2nd ed.; pp. 104–148). Cambridge, UK: Cambridge University Press. doi:10.1017/CBO9781139547369.007

Weidenmann, B. (2001). Lernen mit Medien. In A. Krapp & B. Weidenmann (Eds.), *Pädagogische Psychologie* (pp. 415–466). Weinheim: PVU.

Wells, A. T., & McCrory, R. (2011). Hypermedia and learning: Contrasting interfaces to hypermedia systems. *Computers in Human Behavior*, *27*(1), 195–202. doi:10.1016/j.chb.2010.07.036

Wong, A., Leahy, W., Marcus, N., & Sweller, J. (2012). Cognitive load theory, the transient information effect and e-learning. *Learning and Instruction*, *22*(6), 449–457. doi:10.1016/j.learninstruc.2012.05.004

Yii Lim, M., & Aylett, R. (2009). An Emergent Emotion Model for an Affective Mobile Guide with Attitude. *Applied Artificial Intelligence*, *23*(9), 835–854. doi:10.1080/08839510903246518

Zahn, C., Schwan, S., & Barquero, B. (2002). Authoring hypervideos: Design for learning and learning by design. In R. Bromme & E. Stahl (Eds.), *Writing Hypertext and Learning* (pp. 153–176). Amsterdam: Pergamon.

Zimmermann, A., Lorenz, A., & Specht, M. (2003). User modeling in adaptive audio-augmented museum environments. In P. Brusilowsky, A. Corbett, & F. de Rosis (Eds.), *Proceedings of the 9th International Conference on User Modeling* (pp. 403-407). Berlin: Springer Verlag. doi:10.1007/3-540-44963-9_57

Zumbach, J. (2006). Instructional Design for Hyperaudio Learning with Cellular Phones. In V. Uskov (Ed.), *Proceedings of the Ninth IASTED International Conference on Computers and Advanced Technology in Education* (pp. 388-393). Anaheim, CA: ACTA Press.

Zumbach, J., & Kroeber, C. (2006). Learning with Hyperaudio: Cognitive Load and Knowledge Acquisition in Non-Linear Auditory Instruction. In G. Clarebout & J. Elen (Eds.), *Avoiding simplicity confronting complexity. Advances in studying and designing (computer-based) powerful learning environments* (pp. 359–170). Rotterdam: Sense.

Zumbach, J., & Mohraz, M. (2008). Cognitive load in hypermedia reading comprehension: Influence of text type and linearity. *Computers in Human Behavior*, *24*(3), 875–887. doi:10.1016/j.chb.2007.02.015

Zumbach, J., & Moser, S. (2013, August). *Supporting knowledge acquisition in Hyperaudio learning environments.* Paper presented at the 15th Biennial Conference of the European Association for Learning and Instruction (EARLI), Munich, Germany.

Zumbach, J., & Schwartz, N. (2014). Hyperaudio Learning for Non-Linear Auditory Knowledge Acquisition: A Comparison of Text Type, Presentation Format, and Modality. *Computers in Human Behavior*, *41*, 365–373. doi:10.1016/j.chb.2014.10.024

KEY TERMS AND DEFINITIONS

Cognitive Flexibility Theory: Cognitive Flexibility Theory is a prescriptive instructional design theory which suggests that non-linear learning environments such as hypertext or hypermedia support complex learning and avoidance of oversimplification in learning.

Cognitive Load: Cognitive Load refers to the amount of information to be kept and worked with within the Human Working Memory.

Cognitive Load Theory: Cognitive Load Theory assumes three kinds of cognitive load which occupy resources within the working memory. Intrinsic cognitive load refers to the number of elements and their interactivity, extraneous cognitive load is a cognitive load which does not contribute to learning, and germane cognitive load is the space within the working memory devoted to genuine learning processes.

Scaffolding: Scaffolding describes mechanisms which support learners in self-directed learning without using direct instruction.

Hyperaudio: The non-linear arrangement of audio files within a digital learning environment which allows users to navigate freely between these auditory documents.

Hypermedia: The non-linear arrangement of visually presented information such as text, images, or movies, allowing users to navigate freely by using hyperlinks.

Working Memory: Working Memory refers to a unit of the human information processing system. The duration of information retention is limited.