Computers in Human Behavior 41 (2014) 365–373

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

Hyperaudio learning for non-linear auditory knowledge acquisition

Joerg Zumbach [a,*], Neil Schwartz [b]

[a] University of Salzburg, Austria
[b] California State University, Chico, CA, USA

ARTICLE INFO

Article history:
Available online 7 November 2014

Keywords:
Hypermedia
Mobile learning
Hyperaudio

ABSTRACT

In this research, we present the concept of Hyperaudio as non-linear presentation of auditory information in the context of underlying theoretical assumptions of how Hyperaudio differs from existing non-linear information media. We present a study comparing text and auditory represented information either in a linear or non-linear manner and the interaction of these presentation formats with different underlying text types. Learners had to learn from two different text sorts either from text only in linear or non-linear manner from a computer screen or the same information presented as audio files also presented either in linear or non-linear manner. Results show overall advantages of linear information presentation compared with non-linear information presentation, and the advantages of written text versus auditory text on learning performance assessed with an essay task and a multiple-choice test. Interaction effects indicate that non-linearity increases cognitive load assessed with a self-report measure in auditory instruction compared to linear information presentation while cognitive load in processing written text is not affected by linearity. Further, effects reveal that the text type (ex-pository vs. linear text type) interacts with presentation format showing that expository text leads to comparable learning outcomes in linear and non-linear formats, while presenting linear text type as hypertext or Hyperaudio is here rather unbeneficial.

1. Introduction

In this study, we examined the influence of linear vs. non-linear Hyperaudio on learning outcomes, measured by two different metrics of knowledge acquisition, under the lens of cognitive load theory. The basic research questions were aimed at understanding how auditory and text-based learning material interact with information access (linear vs. non-linear) and different types of text.

The use of auditory instruction has increased during recent years (O'Bannon, Lubke, Beard & Britt, 2011; Vajoczki, Watt, Marquis, & Holshausen, 2010). A major reason for this increase is the increasing ubiquity of mobile audio devices like MP3-players or cell and smart phones. A common use of these mobile devices for learning purposes is to listen to instructional material designed exclusively for auditory instruction, like audiobooks or podcasts. This material is mainly characterized by a linear sequence of orally presented information.

In this paper, we suggest a different format of auditory instruction—specifically, nonlinearly-presented audio information—termed "Hyperaudio". Hyperaudio is comparable to non-linear visual learning material like Hypertext or Hypermedia and audio-visual non-linear learning environments like Hypervideo (e.g., Zahn, Schwan, & Barquero, 2002). The difference between linear text and Hypertext is characterized by the way information is presented and accessed. While linear text (usually presented on a digital display) can only be accessed in a linear manner with "next" and "back" buttons, Hypertext nearly always allows non-linear information retrieval by using interactive hyperlinks that connect different information resources in an associative manner (cf. Chen & Rada, 1996; e.g., like the hyperlinks available in Wikipedia articles). Interestingly, there is a large body of research related to Hypermedia learning and some about Hypervideo, but a dearth of research about Hyperaudio.

We define Hyperaudio as an arrangement of auditorially presented material represented within locally coherent hyperlinked nodes. These audionodes are connected via hyperlinks that enable users to navigate within a Hyperaudio environment. By navigating this Hyperaudio environment, users should be able to understand the relationships between single audionodes as well as develop understanding across nodes for a sense of global coherence. The possibilities of linking nodes and creating an overall navigation structure of a Hyperaudio environment are the same as in common hypertexts or hypermedia environments. Nodes can be linked in a

* Corresponding author at: School of Education, University of Salzburg, Hellbrunner Str. 34, A-5020 Salzburg, Austria. Tel.: +43 (0)662 8044 5801.

E-mail address: joerg.zumbach@sbg.ac.at (J. Zumbach).

http://dx.doi.org/10.1016/j.chb.2014.10.024
0747-5632/

linear, hierarchical, elaborative, or an associate manner (cf. Grabinger & Dunlap, 1996).

However, contrary to linear audiobooks, audio nodes are not visited in a predefined sequence. For example, when a learner listens on a cell phone to an audio document explaining bacteria, the document might be linked by the authors to several other audio files, like cell cores, viruses, etc. These links are displayed on the cell phone while the learner is listening to the document. If the learner is interested in any of the documents, the learner might simply activate the link (e.g., by pressing a corresponding button on the phone) and a new audio file would open. The new file would be linked to other audio files, and so on. In short, current standards and developments in Software (data reduction, e.g., MP3) and portable hardware (e.g., portable MP3-Players, or integrated audio-players in cellular phones) open a wide range of applications and can contribute to research on ubiquitous learning with handheld computer devices (e.g., Hsi, 2003; Roschelle & Pea, 2002).

With regard to hypertext, research on learning has shown that hypertext learning is not always beneficial. Although a meta-analysis provided by Chen and Rada (1996) revealed comparable outcomes or slight advantages of hypertext compared with linear text on learning, Shapiro and Niederhauser stated that "(...) a number of studies have not shown such an effect (...)" (2004, p. 609). Foltz (1996, see also Wells & McCrory, 2011) also argued that learning with non-linear texts increases complexity and makes it harder to learn. Factors like navigation planning, the structure of hypertexts, and learner prior knowledge also influence information retrieval (cf. Salmerón, Baccino, Cañas, Madrid, & Fajardo, 2009). Finally, Zumbach and Moharz (2008) added that while learner characteristics influence performance in hypertext learning compared with linear text, issues of instructional design are also important. They found that basic text design, text type, and complexity of learning material play an important role when comparing learning from text and hypertext. In particular, learners must compensate for poor readability of text when instructional design is poor. This puts a heavier load on working memory resources, which are limited in hypertext because of the dual task of information retrieval and navigation planning. In addition, some text types are traditionally designed in a linear sequence and follow a linear (time) plot. Breaking this linear plot in hypertext can also increase working memory load because learners have to re-construct the plot, which is not necessary in traditional text presentation. Thus, Zumbach and Moharz (2008) argue that complexity of content plays an important role in hypertext learning. When the complexity of a learning domain is high, learners have to make use of multiple information sources. Therefore, designers of a hypertext system might offer links between different aspects of a topic, which is impossible to do with traditional linear text. Taken together, research reveals an ambiguous picture showing that non-linear information media may foster, but also harm, learning depending on learner characteristics, the nature of content, and issues of instructional design.

2. Auditory vs. text-based instruction

Early comparative studies seemed to show an advantage for auditory instruction over text-based instruction (cf. Golas, Orr & Yao, 1994; Nugent, 1982). However, there were many constraints of these early investigations because of the role and behavior of the reader, in addition to shortcomings in the methods of data analysis. More recent research presents ambiguous outcomes (for an overview see Barron, 2004). In fact, Travers (1970, quoted after Barron, 2004, p. 954) stated that "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". The processing of either text or audio is also strongly influenced by the proprietary design of each modality. That is, within text, several design features like headings, paragraphs, highlighting, etc. can support reading comprehension processes that are not easily transferable to spoken language (cf. Hartley, 2004). In addition, reading and rehearsal strategies are made easier by self-pacing during the encoding of information presented in text, but it is difficult to engage in the same processes when information is presented auditorially. In spoken language, there are also stylistic features that can contribute support to comprehension processes. These features consist of the control of speed, use of breaks, emphasis on certain information, etc. that are not available in written text (cf. Barron, 2004; Kürschner & Schnotz, 2008). Another advantage of auditory instruction is that spoken language contains additional information to the message itself. For example, voice carries paralinguistic personality cues that are referred to the speaker (e.g., Nass & Lee, 2001; Nass & Brave, 2005). Common to both modalities is the representation learners create from the text base; that is, learners construct mental representations based on the propositional text model, and this representation is modality *unspecific* (cf. van Dijk & Kintsch, 1983; Johnson-Laird, 1983).

3. Working memory and processing of text and audio

The complexity between visual and verbal material can be informed by an analysis of the way learners process information in working memory (WM). According to Baddeley's (1992, 1998) WM model, auditory information is processed in a phonological loop where it is stored and shortly repeated before subsequently being processed within an episodic buffer. Visual text-based information is first represented in a visuo-spatial sketchpad before being processed within the episodic buffer. Rummer, Schweppe, Scheiter, and Gerjets (2008) assume that text and audio are likely to be represented only in the phonological loop, arguing that Baddeley's model is not modality-specific but rather codality-specific. That is, verbal information processing takes place in the phonological loop independent from its modality; visual-verbal information is imported in the phonological loop by rehearsal. In short, WM load should be independent of modality. If Rummer et al. (2008) are correct, then differences between reading and listening comprehension should not be expected.

And yet, several studies comparing learning with written or auditory verbal information have shown differences between both modalities. For example, there seems to be slight advantages for auditory information compared to visual information, because visual information more often contains irrelevant and distracting stimuli that may have less semantic meaning but still has to be processed (e.g., word or number lists; cf. Kürschner & Schnotz, 2008; Pächter, 1996). For text comprehension, there seems to be advantages of written learning material regarding memory for details, while auditory presentations seem to contribute to better understanding of more overarching concepts or ideas—like a plot, for example (e.g., Hildyard & Olson, 1978; Kürschner, Schnotz, & Eid, 2006).

It is possible that a major advantage of complex text comprehension is due to self-pacing and the application of specific reading strategies—both of which might explain advantages of written text compared to auditory text—at least with regard to the memorization of details. Indeed, working memory capacity is stressed with auditory presentation formats, especially when texts are longer or more complex and contain information that is redundant. Leahy and Sweller (2011) underscore this difference in advantage in their discussion of the transient information effect—an effect characterized by a learner's tendency to drop information kept in