Introduction

Abstract

This chapter aims at clarifying basic concepts related to multimedia: communication, comprehension, and learning. Multimedia communication is considered as the intentional creation, display, and reception of multiple kinds of signs in order to convey messages about some content. It entails two subprocesses: meaning and comprehension. Multimedia meaning is a process in which the producer of a message creates multiple external signs based on his or her prior knowledge in order to direct the recipient's mind so that the recipient understands what the producer means. Multimedia comprehension is the complementary process of reconstructing the previously externalized knowledge in the mind of the recipient. It can be seen as the bottleneck of multimedia communication. Multimedia comprehension and multimedia learning are related but are nevertheless different: While multimedia comprehension results in transient changes in working memory, multimedia learning results in permanent changes in long-term memory. Multimedia learning is a byproduct of multimedia comprehension. Further, an overview of the book is presented.

1.1   What Is Multimedia?

Multimedia is ubiquitous in modern societies nowadays. It plays an increasingly important role in education, business administration, advertising, the economy, finances, news agencies, traveling services, and numerous other fields. The ever-growing Internet is full of multimedia messages about nearly everything, including topics such as how to operate your home trainer, how to change the batteries of your TV remote control, and so forth.

In everyday communication, the concept of multimedia is frequently used in a fuzzy way. Many people understand "multimedia" to be a computer- and web-based combination of digital mass storage devices with

delivery media such as computer screens, loudspeakers, headphones, tablets, or cellphones which deliver spoken or written text with pictures and sound or music. This characterization encompasses three aspects of multimedia: technology, presentation, and perception. The technology aspect refers to the delivery media which include digital networks, computers, screens, and loudspeakers. The presentation aspect refers to the format used to display information such as texts and pictures, which can be in the form of photographs, drawings, maps, graphs, and animations. The perception aspect refers to the organs of perception that receive a multimedia message, usually the eyes and the ears.

The efficacy of multimedia communication depends on all three aspects. The technology aspect is the fundament of multimedia and highly important in terms of practical reliability. Technology enables flexible combinations of different presentation formats. From the viewpoint of cognitive psychology, which focuses on how humans search, perceive, and process information, however, the technology aspect is not very important. Merely reading a text printed on paper, for example, does not fundamentally differ from reading the same text on a computer screen. Generally speaking, the comprehension of multimedia messages is only marginally affected by the technological carrier of the message. Instead, it is heavily influenced by the form in which information is presented and by the way in which a message is perceived by a recipient. Thus, cognitive psychology focuses on the presentation and perception aspects of multimedia communication.

Does multimedia require new technologies? Richard Mayer[1] defines multimedia as the combination of words and pictures. Words can be spoken or written, and pictures can be photographs, drawings, maps, graphs, as well as animations or videos. This straightforward definition of multimedia focuses only on the presentation aspect and ignores the technology aspect. This has important implications: Multimedia does not necessarily require high technology. It also includes the use of books or blackboards instead of computer screens, as well as the human voice instead of loudspeakers. From that point of view, multimedia is not a modern phenomenon. Instead, it has a long tradition which dates back to Comenius,[2] who emphasized the importance of adding pictures to texts in order to improve comprehension in his pioneer work *Orbis Sensualium Pictus* (first published in 1658). Accordingly, one can distinguish between traditional and modern forms of using multimedia.

---

[1] Mayer (2009, 2014); Mayer and Fiorella (2022).   [2] Comenius (1999).

Consider the following examples of multimedia learning. Let us assume that students have to learn about the migration of birds in Europe. To this end, the teacher presents a map of the European continent, indicating where some birds live in summer and where they spend winter. While pointing to the map, she tells the class that many birds breed in middle and northern Europe in summer, but do not stay there during winter. Instead, they fly to warmer areas in the Mediterranean area.

One of the students is assigned the task of learning about a specific migrant bird – the marsh harrier – in order to give a report to her classmates the next day. She walks into the school library and opens a printed encyclopedia of biology, where she finds a text about the marsh harrier and a picture of the bird. Furthermore, she consults the Internet and finds a text and a graph depicting the frequency of marsh harriers during different months in middle Europe. The website also features a sound button which plays the typical call of a marsh harrier near its breeding place. Altogether, the student has practiced three forms of multimedia learning: lecture-based multimedia learning in class, book-based multimedia learning in the library, and web-based multimedia learning at her computer or smart phone. Information was presented in different formats – in the visual modality (written text and pictures) and the auditory modality (oral text and sound).

Teaching and learning are different sides of a specific kind of communication which can be characterized as follows. Teachers, who have greater knowledge about a subject matter than their students, send messages about the subject matter to their students (and sometimes also about their behavior). Students send messages (explicitly or silently) about their understanding, knowledge, and interest to their teachers. Successful teaching and learning take place when the difference between the teacher's and the students' knowledge about the subject matter becomes smaller. However, as multimedia learning is merely a special case of multimedia communication, it follows that multimedia communication also comprises other variants of using multimedia. So, what is multimedia communication?

1.2   Multimedia Communication

To clarify the concepts, we will start with the concept of a medium. A medium is a means for communication that serves to convey messages from a sender to a recipient. These messages are conveyed by external

signs,[3] such as flags, insignia, gestures, spoken or written words, drawings, and so forth. Does that mean the usage of signs is always related to communication? Of course not, because many signs are used outside of the context of communication. We have to remember that the world is full of causality, with causes leading to effects. Thus, effects indicate causes, which means that effects serve as signs for causes. For example, smoke indicates fire; the depth of a footprint indicates the weight of an animal, and so forth. Charles Peirce[4] calls these kinds of signs "indexes." Although we use such signs all the time for our orientation, they do not constitute a form of communication. There is no communication between a burning forest and a firefighter when he or she interprets smoke as a sign of danger and takes action. And there is no communication between a prey and a predator, who is silently following the prey's spoor in order to bring it down. We talk about communication only if signs are produced intentionally (i.e., with a goal in mind) with the aim that the recipient will understand the message and change his or her behavior accordingly. In the teaching context just described, for example, the teacher's communication goal could be to increase her students' knowledge about bird migration. In the context of advertising, the communication goal is usually to convince the addressee to buy a product.

Contrary to animals, whose communication is based on innate, relatively fixed species-specific external sign inventories, human communication is based on a much broader repertoire of powerful signs such as gesture, spoken language, written language, and pictures, which can be flexibly combined and used for all kinds of communication purposes. Many of these signs are human inventions which were created at very different times in history and can nowadays be bound together in multimedia environments.

The distinction between technology, presentation, and perception in the context of media also translates to the analysis of signs. The aspect of technology refers to the carriers of signs: clay, paper, boards, digital devices such as computers or the Internet, and even fleeting carriers such as soundwaves in the case of spoken language. The aspect of presentation refers to how information is displayed by signs such as symbols

---

[3] A sign is an object or event that indicates something else, thus representing it. Signs can exist outside of an individual, such as external objects or events. Internal processes such as pain, fear, imagery, or concepts can also function as (internal) signs. Luria (1973) argued that all higher-order cognitive functions require the use of internal sign systems, because otherwise we would only be informed about the "here and now" and could not reflect about the past or think about the future.
[4] Peirce (1932).

mental representation
(interpretation)

sign   content
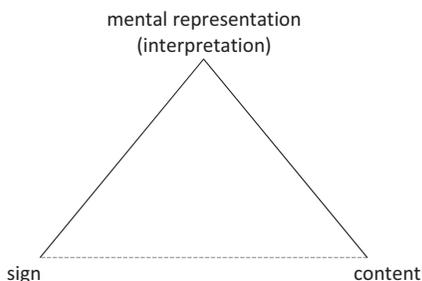
Figure 1.1   Semiotic triangle of Ogden and Richards

(e.g., words) or icons (e.g., pictures). It also refers to the way in which signs are perceived by the recipient. Once again, cognitive psychology focuses on the presentation and the perception aspects. It aims at analyzing how recipients perceive and cognitively process different kinds of signs in order to understand messages. From a psychological point of view, the carrier of signs is not important. If we follow Richard Mayer's parsimonious definition provided in Section 1.1, multimedia communication is the usage of multiple kinds of signs such as texts and pictures to convey messages. Multimedia communication is sign-based communication.

How does sign-based communication work? Signs designate something, which is equivalent to saying that they mean something or they refer to something. The word "bird," for example, refers to all elements in the whole class of birds, whereas the name "marsh harrier" refers only to a portion of this class, and "*this* marsh harrier" refers to a specific animal in the class. To clarify the relations between signs, the meaning of signs, and the content of signs, Ogden and Richards[5] introduced the concept of the semiotic triangle. which is shown in Figure 1.1. The triangle has three constituents: the sign (an external sign in this context), the designated content (also called the referent), and the interpretation of the sign. The interpretation can be understood as a mental representation or as an internal sign of the designated content.[6] The relation between the external sign and the content is not a direct one. Instead, it is mediated by two connected relations: the relation between the external sign and the mental representation and the relation between the mental representation and the designated content.

---

[5] Ogden and Richards (1923).
[6] With the term "representations," we mean any object, event, or state that stands for something else (cf. Peterson, 1996). The concept can apply both to internal structures in the mind (mental representations of some content) and to external structures referring to the content, such as external texts or pictures.
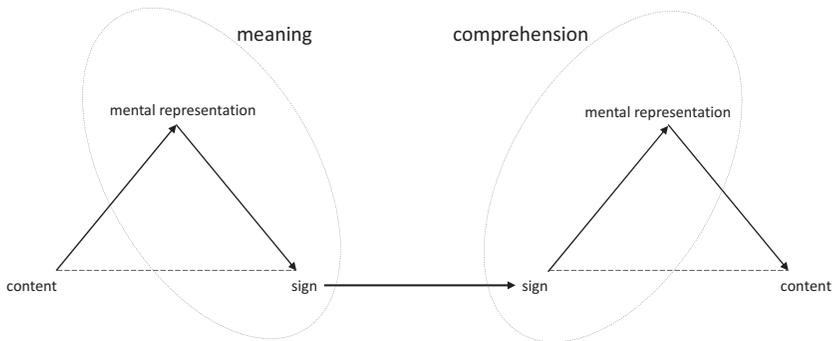
Figure 1.2   Communication considered as a combination of two semiotic triangles

Sign-based communication requires that signs are produced to be understood. In other words, sign production and sign comprehension have to be aligned as well as possible. The alignment of sign production and sign comprehension can be visualized with the help of two semiotic triangles, as shown in Figure 1.2. The sign producer starts with some knowledge about the content (i.e., a mental representation) which he or she has received from another source or from his or her own experience and which he or she wants to communicate. This can be any kind of content, for example, the visual appearance of a bird such as the marsh harrier, its habitat, or migration routes. This knowledge is then externalized by creating external signs. When a producer creates and delivers external signs to someone, we refer to it as "sending a message" to the corresponding recipient. The signs are supposed to mean what the sign producer has in mind. Thus, "meaning" can be considered as a process that creates external signs on the basis of what the sign producer knows or intends. The producer tries to direct the mind of the recipient in such a way that the recipient understands what the producer means.[7]

Based on these external signs, the recipient tries to reconstruct the knowledge that was externalized by the producer. That is, the recipient tries to comprehend the signs. When communication encompasses multiple kinds of signs, the comprehension process is called "multimedia comprehension." If the communication is successful, the recipient's interpretation corresponds to what the sign producer meant; this is what we call "correct comprehension." If the interpretation does not correspond to the intended meaning, we call it "miscomprehension." This implies that the

_____

[7] Clark (1996).

communication was unsuccessful. From a psychological point of view, correct comprehension and miscomprehension involve the same kinds of cognitive processes. However, they both differ from another kind of unsuccessful communication, namely the case where the recipient fails to come up with any interpretation at all. We call this "non-comprehension."

All in all, multimedia communication can be characterized as the intentional production, display, and reception of multiple external signs (corresponding to multiple forms of representation) in order to convey messages about a subject matter.

1.3 What Is Multimedia Comprehension?

We have characterized multimedia comprehension in Section 1.2 as a constituent part of multimedia communication, namely as the reconstruction of knowledge previously externalized by a producer of a multimedia message. This definition raises two questions: First, where does this (re) construction take place? Second, how is multimedia comprehension related to multimedia learning? To answer these questions, we need to understand the human cognitive system.

Most psychologists adopt the view that humans process information in a multiple-store memory system, consisting of sensory registers, a working memory, and a long-term memory.[8] Information from the outside world enters the cognitive system through the sensory organs. Visual information captured by the eyes is stored very briefly (less than 1 second) in a visual register. Auditory information captured by the ears is stored briefly (less than 3 seconds) in an auditory register. Information is stored in the sensory registers only long enough for it to be extracted and passed on for further processing. If attention is directed to information in the sensory registers, this information is transmitted to working memory, where it is further processed in specialized subsystems under the guidance of a central executive.[9] In the case of comprehension, cognitive processing in working memory corresponds to the construction of a mental representation of the content to be understood. The mental construction process draws on external information from the sensory registers and on knowledge of the world (i.e., internal information) retrieved from long-term memory. Due to the limited storage capacity of working memory, which comprises only five elements on average (although this can be effectively increased by chunking), mental representations have to be constructed step by step, in

---

[8] Atkinson and Shiffrin (1971).   [9] Baddeley (1986).

multiple processing cycles. By the same token, complex mental representations cannot be cognitively available as a whole at any time. However, individuals can quickly and flexibly reactivate parts of a mental representation, if needed.[10] If a mental representation that includes new information has been sufficiently interlinked with information from long-term memory and processed repeatedly within working memory, it is likely that the new information is stored in long-term memory, which has a practically unlimited storage capacity. As the term suggests, long-term memory is characterized by a very low decay of information.

Against the backdrop of these assumptions regarding the cognitive system, we can characterize multimedia comprehension based on multiple external signs such as text and pictures as the construction of mental representations of what the multimedia message is about in a recipient's working memory.

1.4   Differences to Multimedia Learning

Whereas multimedia comprehension is a transient change in working memory, multimedia learning is a process that takes place in long-term memory. If no change has occurred in long-term memory, nothing has been learned. There is no way of changing an individual's long-term memory directly, as changes must be triggered by cognitive processing in working memory, such as comprehension. The cognitive processes during comprehension introduce memory traces into long-term memory which allow an individual to remember what he or she previously understood. The individual can reconstruct previously constructed representations in working memory provided that memory traces are still accessible.[11] Learning strategies frequently suggest reiterating comprehension processes systematically and at an increasingly deeper level in order to develop such memory traces in long-term memory, because this makes mental representations easier to reconstruct. Thus, multimedia learning can be considered as a by-product of multimedia comprehension. Conversely, due to the limited capacity of working memory, multimedia comprehension can be considered as the bottleneck of multimedia communication and multimedia learning.

---

[10] The corresponding access structures have been described as "long-term working memory" (Ericsson and Kintsch, 1995).
[11] Cermak and Craik (1979).

This view suggests that multimedia comprehension and multimedia learning are closely related. Good comprehension is usually associated with good learning. Nevertheless, the two processes are different because multimedia comprehension results in transient changes in working memory, whereas multimedia learning results in permanent changes in long-term memory. The difference between comprehension and learning was demonstrated by the infamously striking and bizarre case of neurological surgery performed on Mr. Henry Gustav Molaison (also known as patient H. M.). After a major brain operation in 1953, patient H. M. was still able to comprehend information, but could no longer learn new declarative knowledge. Here is a short description of his deficits:

> The patient suffered from severe epilepsy and headaches. The epilepsy was localized to the left and right medial temporal lobes, which were surgically removed with most of the amygdala and the entorhinal cortex at the age of 27. The operation had no effect on his speech behavior, test intelligence, social behavior, and emotional responses. His working memory was also intact: He could easily perform tasks that required a short-term storage of information. He was also able to remember events before his operation without any problems. However, he could no longer acquire new declarative knowledge. So, he could not remember a new address after a move. He could read the same newspaper repeatedly without realizing that he had already read it. He could play with the same puzzles repeatedly, without remembering that he had already solved them. Like other subjects, he improved more and more, but said he had never done the job and had no idea how to solve it. He thus acquired automated procedural knowledge without the accompanying conscious declarative knowledge about this procedure. Thus, the patient was able to acquire new procedural knowledge in long-term memory after the operation, but he could not learn new declarative knowledge.[12]

In other words, Mr. Molaison's comprehension was still intact, but his learning was severely impaired. His deficits seemed to be caused by a blocked directed connection from working memory to long-term memory.

1.5   Intraindividual Communication with Multimedia

Comprehension is usually the starting point for further cognitive processes, further reflections, thinking, and problem-solving. When individuals think about a subject matter, they frequently externalize their ideas by talking (silently or aloud), writing, or drawing a diagram or a graph. In other

---

[12] Corkin et al. (1997).

words, they express their ideas with the help of different external representation formats. After having produced this externalization and having activated other parts of their prior knowledge, they can reinvestigate their own representations and understand the subject matter from a new perspective. This frequently leads them to other insights and new ideas. When individuals externalize their ideas by talking, writing, or drawing, and then reconsider and reinterpret their externalized ideas, they practice a kind of "communication with themselves." They switch between the roles of the sign producer and the sign recipient until they come up with a mental representation of the subject matter which allows them to answer the question they were concerned about. Thus, besides the interindividual communication described in Section 1.2, intraindividual communication can also exist. It involves individuals creating their own external representations and operating on them in order to come up with new insights, elaborate their thinking about a subject matter, and solve problems.

Whether this intraindividual communication is successful or not depends heavily on the representation formats and how these formats fit with the questions at hand. When individuals have a large scope of representation formats at their disposal, allowing them to select and combine different representations flexibly according to specific requirements, their capacity to think and reflect about a subject matter and find better solutions to problems might be enhanced. Thus, multimedia can also serve as a tool for thinking and problem-solving when self-made multiple representations are used in the context of intraindividual communication.

1.6   Overview

Multimedia technology has developed at high speed in recent years. Despite all technical innovations, however, multimedia comprehension is still constrained by the characteristics of the human cognitive system. Given the central role of working memory in the process of comprehension and its severely limited processing capacity, multimedia comprehension can be considered the bottleneck of multimedia communication and multimedia learning. The use of technologies will only be successful if the psychological laws governing comprehension processes are taken into account.

The present book is about multimedia comprehension. It deals with the construction of mental representations in a recipient's working memory based on multiple external signs such as text and pictures. The book aims at explaining general issues related to multimedia comprehension from a psychological point of view, focusing on the presentation and perception

aspects of multimedia without dealing with questions of technology. Its central question is how recipients perceive and cognitively process different kinds of signs in order to understand messages.

The organization of the book is as follows. After this introduction, Chapter 2 provides an overview of the history of human sign systems used in multimedia communication, including the use of gestures and oral language. These also include early forms of writing, which were based on concepts, and later forms which were (and still are) based on phonemes (but still use hidden signs for concepts). The chapter further deals with the use of different kinds of pictures, including realistic pictures, maps, and graphs. Together, these various sign systems can be combined as tools for creating multimedia messages, which serve specific communication purposes.

Chapter 3 analyzes the principles of representation used by different sign systems more closely. Its main question is how the different sign systems and their representations are related to one another. The chapter argues that the various kinds of representation can be classified into two basic categories: descriptive representations and depictive representations. The two kinds of representation differ in terms of their representational power and inferential power. Both kinds of representation can take the form of external, physical representations and the form of internal, mental representations.

Chapter 4 deals with the comprehension of descriptive representations in the form of written or spoken texts. It analyzes the nature of the meaning of text and the creation of multiple kinds of mental representation. Special attention is given to processes of coherence formation and ways of directing the reader's or listener's flow of consciousness during text comprehension.

Chapter 5 relates to the comprehension of static or animated depictive representations which include realistic pictures, maps, and graphs. Picture comprehension is described as the creation of multiple representations through sub-semantic processing, semantic perceptual processing, and conceptual picture processing.

While Chapters 4 and 5 focus on the comprehension of texts and the comprehension of pictures separately, Chapter 6 analyzes the integrated comprehension of text–picture combinations which is at the core of multimedia comprehension. The chapter is substantiated in the Integrated Model of Text–Picture Comprehension[13] (ITPC model) which includes distinctions between descriptive and depictive representations,

---

[13] Schnotz (2005, 2014, 2022).

external and internal representations, and perceptual surface-structure processing and semantic deep-structure processing. Integrated processing is considered as being embedded in the human cognitive architecture, which is assumed to consist of modality-specific sensory registers, limited-capacity working memory, and long-term memory. The model covers listening comprehension, reading comprehension, visual picture comprehension, and sound comprehension.

Chapter 7 provides further and deeper analyses of the integrated comprehension of texts and pictures with a special focus on inter-representational coherence formation and mental model construction. The chapter also points out the interplay between the ambiguity of representations and their disambiguation by providing complementary representations. It further elaborates on the different, but complementary functions of texts and pictures during different phases of task-oriented multimedia comprehension.

The focus of Chapter 8 goes beyond comprehension. When multimedia comprehension has been successful, an individual can use his or her mental representation in order to infer new information or to solve problems by means of productive thinking. The chapter discusses different views of productive thinking and problem-solving. It points out that productive thinking and problem-solving make specific use of descriptive and depictive representations. In addition, it shows that the use of descriptive and depictive representations has implications, for example, for mathematics education and science education. Finally, referring to historic examples of high practical relevance, the chapter describes the use of depictive representations for data collection and statistical problem-solving based on content-related hypotheses.

Finally, Chapter 9 analyzes the conclusions that can be drawn from the previous theoretical analyses and sets out the practical implications. It provides suggestions on how developers should design multimedia messages, as well as on how recipients should process them.