Julius-Maximilians-Universität Würzburg, Fakultät für Humanwissenschaften

Cognitive Processes of Discourse Comprehension in Children and Adults

- Comparisons between Written, Auditory, and Audiovisual Modes of Presentation -

Inaugural-Dissertation
zur Erlangung der Doktorwürde der
Fakultät für Humanwissenschaften
der
Julius-Maximilians-Universität Würzburg

vorgelegt von
Wienke Charlotte Wannagat

aus Würzburg

Würzburg, 2018

Erstgutachterin:	Prof. Dr. Gerhild Nieding
Zweitgutachter:	Prof. Dr. Peter Ohler

Tag des Kolloquiums: 16. März 2018

Vorwort

Diese Arbeit entstand im Rahmen des DFG-Projekts "Die Entwicklung der kognitiven Filmverarbeitung" (Ni496/9-1).

Mein Dank gilt zunächst meiner Betreuerin, Prof. Dr. Gerhild Nieding für die Möglichkeit ein Thema innerhalb dieses Projekts bearbeiten zu können und für die vielfältige Unterstützung dabei, insbesondere die vielen wertvollen Diskussionen, die mich immer wieder herausgefordert und zum Nachdenken angeregt haben. Vielen herzlichen Dank!

Außerdem bedanke ich mich bei Prof. Dr. Peter Ohler, der sich als Zweitgutachter zur Begutachtung der Arbeit bereit erklärte sowie bei PD Dr. Eva Michel für die Bereitschaft im Kolloquium als Drittprüferin zu fungieren.

Für die hervorragende Arbeit beim Erstellen der Bilder, die einen wesentlichen Teil des Versuchsmaterials darstellen, bedanke ich mich bei Pauline Feichtinger, Sonja Köhler und Anna Laumer. Danken möchte ich auch allen studentischen Hilfskräften sowie Bachelor- und Masterarbeitsstudentinnen, die mich bei der Datenerhebung unterstützt haben. Insbesondere gilt hier mein Dank den Forschungspraktikantinnen (in chronologischer Reihenfolge) Sabrina Guffler, Madita Schild und Annsophie Ernst. Weiterhin bedanke ich mich auch bei allen kooperierenden Schulen und Versuchspersonen sowie den Eltern der teilnehmenden Kinder.

Vielen lieben Dank an alle aktuellen und ehemaligen Kolleginnen und Kollegen vom Lehrstuhl! Ganz besonders danke ich Caro, Catharina, Elisabeth, Frank, Gesine, Juliane und Katharina für die Unterstützung und insgesamt die schöne Zeit vor, während und nach den Mittags- und Tee-/Kaffeepausen! Gesine Waizenegger war außerdem in der Anfangsphase die beste Projektkollegin, die ich mir hätte wünschen können.

Summary

This thesis consists of three studies that investigated if and how different modes of presentation - written, auditory, audiovisual (auditory combined with pictures) - affect comprehension of semantically identical materials. Children aged 7 to 12 and adults were included into the studies.

The theoretical foundation for this thesis and accordingly the definition and operationalization of comprehension refer to comprehensive theories of text comprehension that assume that this process involves at least three levels of processing (Kintsch, 1998): a representation of the text surface that is still close to perception, a representation of the propositional text base that contains the text's semantic content, and finally a situation model that represents the whole situation described in a text. While the mental representations of the text surface and the text base preserve the representational format of the text, the situation model is assumed to also include perceptual simulations of the verbally described situations (e.g., Zwaan, 2014). Furthermore, text comprehension requires the reader to establish coherence on a local level by connecting subsequent sentences, and on a global level by connecting widely separated text passages with each other, or a current phrase with the overall topic (Graesser, Singer, & Trabasso, 1994).

Beneficial effects of pictures added to auditorily presented verbal materials have been found in a vast amount of studies, at least for children up to the age of 12 (e.g., Carney & Levin, 2002). These effects correspond to the *multimedia effect* (e.g., Mayer, 2005a), according to which people supposedly learn better from words and pictures than from words alone. The multimedia effect can be explained by mechanisms of dual coding (Paivio, 1986), thus assuming that pictures offer additional memory code that supports processing of the relevant content (J. Fletcher & Tobias, 2005). Moreover, in contrast to words, pictures are assumed to physically resemble their referents and therefore represent the reality analogically, based on common structural properties (e.g., Schnotz, 1993). Therefore, these specific characteristics of pictures may facilitate situation model construction, assuming that the situation model is, at least in part, a perceptual simulation of the represented information and resembles the representation of a real-life experience (Zwaan, 2014).

If cognitive processing of verbal materials differs depending on the mode of presentation (auditory or written), has been controversially discussed. Unitary process views assume a general ability of receptive language comprehension and therefore assume the same cognitive processes to underlie reading and listening comprehension. In contrast,

dual process views assume cognitive processing and mental representations to differ depending on whether someone reads or listens to a text (e.g., Danks & End, 1987). Studies that compared reading and listening comprehension have produced inconclusive results so far.

Using a sentence recognition task (Schmalhofer & Glavanov, 1986), we examined whether the memory of the text surface, the text base, and the situation model differs between written, auditory, and audiovisual text presentation in a sample of 103 8- and 10-year-olds and adults (Study I), and between auditory and audiovisual text presentation in a sample of 106 7-, 9-, and 11-year-olds (Study II). Furthermore, we examined with 155 9- and 11-year-olds, whether the ability to draw inferences to establish local and global coherence differs between written, auditory, and audiovisual text presentation. These inferences were indicated by reaction times to words associated with a story protagonist's super- (global) or subordinate (local) goal in a word-recognition task.

In terms of the comparison of verbal-only (auditory and written) and audiovisual text, the results of these three studies taken together, indicate that which mode of text presentation promotes text comprehension depends on the specific cognitive process. On the one hand, for children at least up to the age of 12 both memory of the situation model and establishing global coherence, which can be summarized as processing of the text's gist or macro structure, are facilitated by an audiovisual text presentation. On the other hand, it seems that both the memory of the text base and local coherence, which may be more closely linked to the text base than to situation model processing, do not benefit from additional pictures. This effect is discussed within the framework of the integrated model of text and picture comprehension according to which the cognitive processing of pictures does not necessarily include mentally representing a propositional structure (i.e., text base) of the content (e.g., Schnotz & Bannert, 2003). For the adults, in contrast, we did not find an effect of the presentation mode when comparing verbal-only (auditory and written) and audiovisual text.

In terms of the comparison of auditory and written text, the findings of the Studies I and III taken together suggest that the age and thus presumably the decoding ability seems to mediate the effect of presentation mode on the various aspects of text comprehension assessed in the three studies. Eight-year-olds showed better memory of the situation model at auditory than at written text presentation. Older children (ages 10, and 11) and adults in turn, apparently either do not benefit from either form of presentation, as shown in terms of memory of the situation model (Study I), or perform better after reading a text than after listening to it, when it comes to establishing local and global coherence (Study III). Consequently, the result pattern produced by these studies does not speak in favor of either the unitary or the dual process view.

Zusammenfassung

Die vorliegende Dissertation umfasst drei Studien, die untersuchen, ob und wie verschiedene Darbietungsformen - schriftlich, auditiv und audiovisuell (auditiv in Kombintation mit Bildern) - die kognitiven Prozesse beim Verstehen semantisch identischer Texte beeinflussen. Es wurden Kinder zwischen 7 und 12 Jahren sowie Erwachsene untersucht.

Die theoretische Basis dieser Arbeit und somit auch die Definition und Operationalisierung von Textverstehen orientiert sich an der vorherrschenden Sichtweise, dass dieser Prozess mit dem Aufbau von mindestens drei Ebenen mentaler Repräsentationen einhergeht (Kintsch, 1998): einer wahrnehmungsnahen Repräsentation der Textoberfläche, einer Repräsentation der propositionalen Textbasis und schließlich einem Situationsmodell, das die gesamte im Text beschriebene Situation mental repräsentiert. Während bei der Repräsentation der Textoberfläche und der Textbasis das Repräsentationsformat des Textes beibehalten wird, nimmt man für das Situationsmodell an, dass dieses auch perzeptuelle Simulationen der verbal beschriebenen Situation enthalten kann (z. B. Zwaan, 2014). Außerdem erfordert das Verstehen eines Textes, dass Kohärenz hergestellt wird. Dieser Prozess umfasst sowohl eine lokale Ebene, indem aufeinanderfolgende Sätze miteinander verknüpft werden als auch eine globale Ebene, indem weiter auseinanderliegende Textpassagen miteinander und aktuelle Information mit dem übergeordneten Thema verknüpft werden (Graesser et al., 1994).

Dass sich Bilder förderlich auf die Verständnisleistung auswirken können, gilt als gut untersucht (z. B. Carney & Levin, 2002). Dieser Effekt wird auch als *Multimedia-Effekt* bezeichnet (z. B. Mayer, 2005a). Dieser besagt, dass das Lernen anhand einer Kombination aus Text und Bild effektiver sei als das Lernen mit nur einem Text. Der Multimedia-Effekt lässt sich unter anderem auf das Prinzip der dualen Kodierung zurückführen (Paivio, 1986), demzufolge Bilder zusätzliche Gedächtnisspuren bilden, die das Verstehen des relevanten Inhalts unterstützen (J. Fletcher & Tobias, 2005). Weiterhin nimmt man von Bildern an, dass diese, im Gegensatz zu Wörtern, einen Referenten aufgrund gemeinsamer Strukturmerkmale wie äußerlicher Ähnlichkeit repräsentieren (Schnotz, 1993). Diese spezifischen Charakteristika eines Bildes unterstützen dementsprechend möglicherweise die Konstruktion eines Situationsmodells unter Voraussetzung der Annahme, dass auch ein Situationsmodell, zumindest teilweise, eine perzeptuelle Simulation der beschriebenen Situation darstellt, die der Repräsentation einer realen Erfahrung entspricht (Zwaan, 2014).

Ob die kognitive Verarbeitung verbaler Information sich je nach Darbietungsform (auditiv oder schriftlich) unterscheidet, wurde viel diskutiert. Monistische Positionen gehen von einer generellen Fähigkeit rezeptiver Sprachverarbeitung aus und nehmen dementsprechend an, dass dem Lese- und Hörverstehen die gleichen kognitiven Prozesse zugrunde liegen. Dualistische Positionen hingegen nehmen an, dass sich kognitive Verarbeitungsprozesse und dementsprechend auch die mentalen Repräsentationen unterscheiden, je nachdem, ob ein Text gelesen oder gehört wird (z. B. Danks & End, 1987). Studien, die Hör- und Leseverstehen systematisch verglichen haben, lassen allerdings keine eindeutige Befundlage erkennen.

Mit einer Satz-Rekognitionsaufgabe (Schmalhofer & Glavanov, 1986) wurde in dieser Arbeit untersucht, ob sich das Gedächtnis für die Textoberfläche, die Textbasis und das Situationsmodell bei 8- und 10-Jährigen sowie Erwachsenen (N = 103) zwischen schriftlicher, auditiver und audiovisueller Darbietung (Studie I) und bei 7-, 9- und 11-Jährigen (N = 106) zwischen auditiver und audiovisueller Darbietung unterscheidet (Studie II). Außerdem wurde bei einer Stichprobe von N = 155 9- und 11-Jährigen die Fähigkeit Kohärenz auf lokaler und globaler Ebene herzustellen zwischen einer auditiven, einer schriftlichen und einer audiovisuellen Darbietung verglichen (Studie III). Diese Prozesse wurden im Rahmen einer Wort-Rekognitionsaufgabe anhand der Reaktionszeiten auf Wörter, die mit dem über- (globalen) oder untergeordneten (lokalen) Ziel des Protagonisten einer Geschichte assoziiert sind, untersucht.

Bezüglich des Vergleichs zwischen verbalem (auditiv und schriftlich) und audiovisuellem Text lässt sich aus den Ergebnissen der Studien insgesamt ableiten, dass es vom spezifischen kognitiven Prozess abhängt, ob sich das Textverstehen zwischen den Darbietungsformen unterscheidet. Einerseits zeigte sich, zumindest für die untersuchten Kinder, dass Prozesse wie die globale Kohärenzbildung und der Aufbau eines Situationsmodells, die mit makrostruktureller Verarbeitung assoziiert sind, durch die Darbietung von Bildern gefördert werden. Andererseits zeigte sich dies nicht in Bezug auf das Gedächtnis für die Textbasis sowie für das Herstellen lokaler Kohärenz, einem Prozess, der eher mit Verarbeitung auf Ebene der Textbasis als auf Ebene des Situationsmodells assoziiert ist. Dieser Effekt wird vor dem Hintergrund des "Integrated Model of Text and Picture Comprehension" diskutiert, demzufolge, die Verarbeitung eines Bildes nicht zwangsläufig mit der Konstruktion einer propositionalen Repräsentation einhergeht (z. B. Schnotz & Bannert, 2003). Bei den Erwachsenen fand sich im Bezug auf diesen Vergleich kein Effekt der Darbietungsform.

Für den Vergleich zwischen auditivem und schriftlichem Text zeigt sich in einer Gesamtschau der Studien I und III, dass das Alter und dementsprechend vermutlich die Dekodierfähigkeit den Effekt der Darbietungsform mediiert. Achtjährige zeigten bei auditiver Darbietung ein besseres Gedächtnis für das Situationsmodell als bei schrift-

licher Darbietung. Hingegen zeigte sich für ältere Kinder (10 und 11 Jahre) und Erwachsene entweder kein Effekt wie in Bezug auf das Situationsmodell (Studie I) oder die Leistung war bei schriftlicher Darbietung besser, wie sich hinsichtlich globaler und lokaler Kohärenz zeigte (Studie III). Hinsichtlich der Frage, ob dem Lese- und Hörverstehen die gleichen oder unterschiedliche Prozesse zugrunde liegen, lassen die vorliegenden Befunde dementsprechend keine Aussage zu.

Contents

Introduction

> *"Cartoons don't have any deep meaning. They're just stupid drawings that give you a cheap laugh."*

> *"It's not our fault our generation has short attention spans, Dad. We watch an appalling amount of TV."*

> *"Homer: Marge, I'm bored. - Marge: Why don't you read something? - Homer: Because I'm trying to reduce my boredom."*

These quotes taken from the popular TV show *The Simpsons* illustrate two important aspects about media. Firstly, with television, cartoons, and books - to only name the examples featured in these quotes - both children and adults are confronted with a variety of media in their everyday lives. Secondly, these quotes satirically exaggerate some of the prejudice about the use and effects of these different media: Pictures are superficial and not useful to transport any deeper meaning, TV makes children stupid, and books are not a viable alternative nevertheless, because reading may be perceived as boring.

Also in the academic debate audiovisual media have been suspected to hamper meaningful processing. In an often-cited study, Salomon (1984) argues that information presented via television is superficially processed because children may perceive television as "easy" and therefore do not invest the same amount of mental effort as they do when, for example, reading a book. A similar debate had emerged around books when they became available to a broad audience as a form of entertainment in the 18th century (Wrage, 2010). Also "the internet" in general, or recently smart phones or tablet computers are not only enthusiastically celebrated but also often criticized for causing attention deficits in especially children and adolescents (Riedel, Büsching, & Brand, 2017).

Every newly emerging form of media seems to be accompanied by some general skepticism. To face this skepticism and scientifically examine its truth value to participate in an informed debate on how information should best be provided to a recipient and at what age, it is necessary to initially agree upon a definition of different kind of media. For instance, Mayer (2005b) offers a useful perspective on this matter: You can differentiate on the level of technical devices and compare computers, television, smartphones, books, radio for example. When you try to further define what signifies these

devices, it becomes clear that another descriptive level is more appropriate. Essentially, irrespective of the device, two different types of signs are used to transport meaning in these media: words and pictures. In books, words are written. The exact same words can also be presented auditorily and transmitted via radio or as a podcast available via the internet, to only name a few examples. This distinction introduces another level to be considered. The same signs can be perceived via different sensory channels. Often words are combined with pictures, which can either be static, as in comic books for example, or dynamic (animated).

The importance of different media or representational formats varies with the age of the recipient. From a very early age, children acquire knowledge from auditory information and become quite proficient in listening comprehension before they even start learning to read. Often, the auditory information children consume is further enhanced with pictures, as in illustrated books or film. According to current media usage data, watching films or television is a popular leisure activity among elementary school children (Rideout, 2013). Later, when children learn to read, they acquire one of the most important cultural techniques. Step by step, with their reading skills developing, written text gains increasing importance as a source of information.

Nevertheless, regardless of the underlying sign systems or the perceptual channel, a communicator usually tries to deliver a message in a way that the recipient *understands* its content. From the perspective of cognitive psychology, understanding - or comprehension - means constructing a *coherent mental representation* of the respective subject matter (Trabasso & van den Broek, 1985), irrespective of whether the information is presented in auditory, audiovisual (i.e. auditory presentation combined with pictures), or written form. The cognitive processes that build up to a coherent mental representation and thus describe and explain comprehension are extensively studied in the field of cognitive psychology. This implies that there are many comprehensive theories available that allow to carefully and systematically derive measures of comprehension as the variable of interest when conducting cross-media research. As I will describe later in more detail, cross-media studies (i.e. studies that compare different presentation modes) so far only occasionally considered this theoretical foundation, but rather asked their participants to retell the respective content or asked questions which were assumed would capture comprehension.

This thesis therefore takes well-established theories of text comprehension as the basis for a further exploration of whether and how different ways of presenting information affect comprehension in both children and adults.

Outline of this thesis

This thesis is divided into three chapters.

Chapter 1 deals with the theoretical background of the empirical studies and discusses the cognitive processes that underlie discourse comprehension in general, while also addressing differences between children and adults. This chapter then further elaborates if and how different modes of presentation, namely verbal information and pictures as well as the sensory channel through which verbal information is perceived (auditory and written text) affect these basic cognitive processes that constitute comprehension.

Chapter 2 describes and discusses three empirical studies that compare cognitive processing of texts of different modes of presentation. Study I investigated whether the memory of the text surface, the text base, and the situation model differs between auditory, audiovisual, and written text presentation in samples of 8- and 10-year-olds and adults. Study II, similarly compared memory of these three levels but only between auditory and audiovisual modes of presentation and with different age groups (7-, 9-, and 11-year-olds). Study III further elaborated the idea of text comprehension as the process of building a coherent mental representation and examines in a sample of 9- and 11-year-olds whether the ability to establish local and global coherence differed between auditory, audiovisual and written texts.

Finally, *Chapter 3* integrates the findings of the three studies described and discussed in Chapter 2 in the light of the theoretical background presented in Chapter 1.

1 Theoretical Background

This chapter gives an overview of the theoretical background of the three studies included in this thesis. The first two sections deal with the basic assumptions of cognitive text processing, which can be described as a process of coherence formation that involves at least three different levels of mental representations. These considerations constitute the theoretical foundation of my definition and operationalization of text comprehension, which is employed to examine the overall question of this thesis, whether comprehension differs depending on the mode of text presentation.

Accordingly, the third section then deals with the specifics of the processing of different modes of presentation and explores empirical findings and theoretical accounts on the differences between processing verbal and pictorial information, as well as written and auditory information. All three sections also discuss if and to what extent the introduced theoretical considerations and empirical findings, that all assume adult recipients, can be generalized over different age groups, including children.

1.1 Comprehension as a Process of Constructing Multiple Mental Representations

Irrespective of whether information is externally represented by a text, a picture, or a combination of both, comprehension of this subject matter means that the recipient constructs a coherent internal, i.e. mental representation of this subject matter. However, comprehensive theories that describe this process predominantly refer to the comprehension of verbal information. This is why these theories form the background of the current series of studies, that have the overarching goal to compare comprehension between different kinds of presentation modes and thus also contain pictorial information.

Among researchers, who study discourse processing, it is generally agreed upon that the cognitive processing of a text results in three different mental representations: the text surface, the text base, and the situation model (Graesser, Millis, & Zwaan, 1997). This distinction has been originally introduced by van Dijk and Kintsch (1983), and has been further developed and extended since then (Kintsch, 1988, 1998; Zwaan & Radvansky, 1998; Zwaan & Singer, 2003). While the early work of Kintsch and van Dijk focused on processing at the text base level (Kintsch & van Dijk, 1978; Kintsch, 1988), newer research highlights the nature and features of the situation model, which is considered to represent "deep" understanding.

Furthermore, a text can also be processed on the levels of *communication*, i.e. the pragmatic communicative context, and *text genre* (Graesser et al., 1997). As these two aspects are not central for my research questions, they will not be elaborated further.

In the following, the text surface, the text base, and the situation model will be discussed, focusing on their relevance for discourse processing.

1.1.1 The text surface

The mental representation of the text surface contains the exact wording of a text. It is usually subject to a fast decay (Sachs, 1967; Bransford, Barclay, & Franks, 1972). Sachs (1967) found that readers were equally able to detect alterations made to a sentence that refer to the text surface and alterations that refer to the meaning when the test sentence followed immediately after the original sentence. When more than 80 syllables of filler words/sentences separated original and test sentences, changes concerning the meaning were far more likely to be detected than changes to the text surface. However, some studies found a mental representation of sentences stored in long-term-memory contained some text surface traces (Bates, Masling, & Kintsch, 1978; Tardif & Craik, 1989). A mental representation of the text surface can also serve a specific purpose, for instance, if someone plans to recite a poem and therefore memorizes its surface. However, this surface representation does not imply comprehension, as it is possible to recite a poem that is written in a foreign language without being aware of the meaning of the words. Thus, to gain understanding a recipients has to engage in further semantic processing.

1.1.2 The text base: propositional representations

A text's semantic content, independent of the exact wording, is represented by a mental representation of the text base. The text base is usually depicted as a network of interconnected propositions. Propositions are the smallest possible meaningful units to which you can assign a truth value and are organized in a predicate-argument structure (Hemforth & Konieczny, 2008; Kintsch, 1988, 1998). The sentence "Tom gives Mary a pen." can be expressed as the following proposition: P (GIVE, agent: TOM, object: PEN, recipient: MARY). This proposition symbolizes the relation between the proposition's arguments TOM, MARY, and PEN, and it is further specified by the predicate GIVE (Schnotz, 1994). According to this principle a whole text can be disassembled into its propositions. To also illustrate the hierarchy of and the relations between propositions, the text base can be displayed as a list of propositions (Kintsch, Kozminsky, Streby, McKoon, & Keenan, 1975) or in form of a hierarchical coherence graph (Kintsch, 1974). The text passage *"The Greeks loved beautiful art. When the romans conquered the Greeks,*

they copied them, and thus, learned to create beautiful art." can, according to (Kintsch et al., 1975), be displayed as the following list, in which indentations indicate a hierarchical structure.

1 (LOVE, GREEK, ART)
    2 (BEAUTIFUL, ART)
    3 (CONQUER, ROMAN, GREEK)
    4 (COPY, ROMAN, GREEK)
        5 (WHEN, 3, 4)
        6 (LEARN, ROMAN, 8)
        7 (CONSEQUENCE, 3, 6)
        8 (CREATE, ROMAN, 2)

The following Figure 1 displays the example passage as a coherence graph.

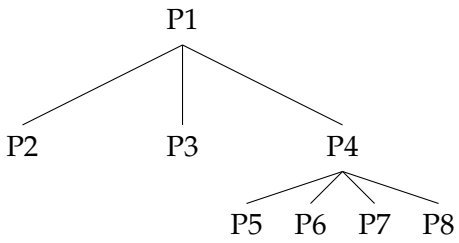


Figure 1. The text base of the example passage displayed as a coherence graph as proposed by Kintsch (1974). P1 - to examplarily explain the notation system - means "proposition 1" and thus refers to the first proposition of the list of propositions (LOVE, GREEK, ART).

Kintsch (1974) further distinguished different types of propositions, such as predicative, modifying, and connective propositions. Predicative propositions indicate actions or states and their predicates are mostly verbs and their arguments can be classified, for instance, as agents, instruments, objects, or recipients. Modifying propositions indicate part-whole-relations by qualitative or quantitative relationships or further specify propositions by negations. Connective propositions have other propositions as arguments and thereby form more complex propositions.

The relevance of propositions as a form of mental representation for text comprehension and memory has been shown in various studies and contexts. For example, Kintsch and Vipond (1979) identified a more complex propositional structure as the reason why Adlai Stevenson's presidential campaign speeches were generally perceived as more difficult to understand than his opponent's Eisenhower speeches, although they did not differ with regard to length of the sentences, or the commonness of the words used

(McKoon & Ratcliff, 2008). Furthermore, Kintsch (1974) explained differences in the reading time of texts of the same length by the amount of propositions. Further evidence of the psychological reality of proposition is provided by Ratcliff and McKoon (1978), who showed that word recognition times of certain words were faster when the presentation of this word was preceded by a word that belonged to the same proposition compared to words that belonged to a different proposition, even when they were closer on the surface-level. For instance, the word *square* as part of the previously presented sentence "The mausoleum that enshrined the tsar overlooked the square." was recognized faster when preceded by the word *mausoleum* than when it was preceded by *tsar*.

A propositional representation can be characterized as some sort of inner speech that internally describes the represented subject matter (Johnson-Laird, 1987). This inner speech is still very close to the natural language and preserves its symbolic character. This definition reveals a major deficit of this type of representation. Consequently, it indicates that a further level of representation is necessary to appropriately describe what is meant by comprehension in its actual sense. If cognitive text processing stopped at this point, various phenomena that occur during discourse comprehension could not be explained. For instance, if the example sentence "Tom gives Mary a pen." was part of a novel, that deals with a series of mysterious incidents in Tom and Mary's house, the reader would probably develop a more or less detailed mental image of the layout of the house. Even if such a description is not explicitly included in the propositional structure. Furthermore, the text base does not acknowledge the important role of prior topic knowledge for text comprehension. Someone with high prior knowledge and low verbal abilities is nevertheless able to capture the overall meaning of a text quite well (Rinck, 2000).

1.1.3 The situation model

In a seminal experiment Bransford et al. (1972) showed that people particularly remember the situation described in a sentence. This finding marks the beginning of the acknowledgement of the situation model and of its relevance for text comprehension. In this experiment, participants heard either one of the following two sentences

(1) Three turtles rested beside a floating log, and a fish swam beneath them.

(2) Three turtles rested on a floating log, and a fish swam beneath them.

These two sentences have an identical deep structure and only differ with regard to the prepositions *on* or *beside*. However, the semantic situation implied differs between these two sentences. Sentence (2) also contains additional information about the spatial arrangement of the turtles, the log, and the fish, as the fish according to this sentence

not only swims beneath the turtles but also beneath the log. This information is not directly contained in the verbatim or propositional information provided, but has to be inferred with reference to general knowledge about spatial relations. Afterwards, the participants were presented with the following recognition sentence

(3) Three turtles rested beside/on a floating log and a fish swam beneath it.

By changing the pronoun *it*, both the sentences (1) and (2) differ from sentence (3) with regard to the surface and propositional structure. However, only sentences (2) and (3) contain the same semantic description of the situation.

Participants who had previously heard sentence (2) were more likely to assume that sentence (3) matched with the previous sentence than participants that had previously been presented with sentence (1). Consequently, this experiment provides evidence that during text processing people construct holistic semantic descriptions of situations.

Accordingly, the situation model is defined as a mental representation of the situation described in a text. In the situation model, the recipient (reader or listener) integrates the textual information with prior knowledge and draws inferences to construct a coherent representation of the situation described. A situation model of the sentence "Tom gives Mary a pen." would probably include further information about the setting, such as the place where Tom gives Mary the pen, and the reader or listener is likely to infer that Mary plans to write something down. In contrast to the first two levels, the situation model goes beyond what is explicitly stated in the text. A coherent situation model is considered the final and most important product of discourse processing, implying deep understanding of the state of affairs represented by the text (Zwaan & Radvansky, 1998).

Kintsch and van Dijk (Kintsch & van Dijk, 1978; Kintsch, 1988) originally assumed the situation model to be a somewhat inferentially enriched text base, and therefore to be largely amodal and abstract. In this sense, the finding by Bransford et al. (1972) may also be explained by the participants having stored the additional spatial information they inferred in form of an additional proposition that specifies the location of the fish (Schnotz, 1994).

This view is opposed or supplemented by researchers who propose that the situation model is not only quantitatively, but also qualitatively different from the mental representation of the text base.

1.1.3.1 Mental models as analogue mental representations

Johnson-Laird (1983) introduced the term mental model in his attempt to explain how information is mentally represented. According to Johnson-Laird (1980, 1983, as cited in Schnotz, 1994) a mental model is an analog mental representation of the represented

situation. This means that specific properties of the model correspond to specific properties of the represented original. Mental model properties are only functionally and not physically analog to the original. Therefore, mental models are not necessarily visual images of what they represent. The structural analogy is based on inherent functional properties. This is why a mental model is defined as an intrinsic representation, in contrast to propositional representations that are made of arbitrary signs. Furthermore, mental models are assumed to be constructed to serve a certain purpose and thus cannot be comprehensive reflections of the represented situations. This specificity to a purpose implies that mental models vary with regard to granularity and perspective. Considering text processing, determining factors can be the recipient's goal or other characteristics. For example, Larkin (1983) found that learners with little experience in terms of a certain topic tend to represent surface features that may not be very useful. Experts, on the contrary, rather construct mental models that contain all necessary structural and functional features.

1.1.3.2 Embodied theories of language: text comprehension as perceptual simulation

While Johnson-Laird specifically assumed a *functional* analogy of the situation (or mental) model and reality, researchers promoting the idea of text comprehension as a perceptual simulation go even a step further.

Research conducted over the last two decades has increasingly produced evidence for the assumption that comprehension of words, phrases, or sentences means that the represented objects, situations, or events are mentally simulated (Kaup, de la Vega, Strozyk, & Dudschig, 2015). This means that a mental representation of a situation derived from a text can be an analogous, multidimensional, and modality-specific simulation of the described situation, resembling the mental representation of a real-life experience. Returning to the introductory example of Tom and Mary, the embodied account of language comprehension suggests that the situation model is also likely to include a mental representation of what the pen in Mary's hand feels and looks like. In this sense, the situation model is analogous to perception and is no longer explicitly tied to the symbolic code of the text but contains visual, auditory, proprioceptive, motoric, and haptic information (Barsalou, 1999, 2008; Glenberg & Kaschak, 2002; MacWhinney, 1999; Zwaan, 2004).

There is a large amount of evidence of perceptual mental representations in adults. Findings of two experiments by Zwaan, Stanfield, and Yaxley (2002) indicate that people mentally simulate the shape of objects during language comprehension. The researchers created sentence pairs that described animals or objects at certain locations, such as for example

(4) The ranger saw the eagle in the sky.

(5) The ranger saw the eagle in its nest.

They further created pictures of the described objects matching/mismatching each sentence. The eagle in the above mentioned sentence pair can be illustrated by a picture of an eagle with its wings outspread and thus matching sentence (4) and mismatching sentence (5), or an eagle with its wings folded. Participants read a series of sentences and were shown a picture that either matched or mismatched the object/animal as described in the sentence. They were asked to decide whether the picture had been mentioned in the previously read sentence (Experiment 1), or name the object shown on the picture (Experiment 2). Both the reaction times and naming latencies were faster when the object shape implied in the sentence matched the objects as shown in the pictures.

Similarly, further research indicates that also mental simulations of colors are activated when color words are processed, as shown by Richter and Zwaan (2009) who found that lexical decisions on color words were faster when a preceding color matched the color word. These findings are supplemented by research indicating mental simulation also for other aspects of the situation model, such as sounds (e.g., Brunyé, Ditman, Mahoney, Walters, & Taylor, 2010), and perspectives (e.g., Horton & Rapp, 2003).

Evidence of mental simulations is not limited to adults. There is some evidence that children perceptually simulate aspects of a text that are not explicitly stated, indicating that these mechanisms of embodied cognition are already present early in life. For example, Unsöld and Nieding (2009) conducted an experiment, in which children watched short film sequences that were at some point interrupted and a static picture was shown. This picture was either consistent with the state of the object indicated by the course of the narration so far, or not. The participants were then asked to name the object shown on the picture as fast as possible. The results indicated that naming-latencies in children as young as 6-years old, were faster when consistent pictures were shown. Corresponding findings have also been reported with regard to the shape of objects (Engelen, Bouwmeester, de Bruin, & Zwaan, 2011; Hauf, 2016) and taking a protagonist's spatial perspective (Rall & Harris, 2000; Ziegler, Mitchell, & Currie, 2005).

Lately, the debate about perceptual simulation and embodiment in language comprehension has shifted from merely discussing its existence to more closely examining its functional relevance and the interplay with propositional mental representations. Zwaan (2014) as well as Kaup et al. (2015) consider a *dual-format view* of language comprehension that accounts for both amodel propositional representations and modality-specific perceptual simulation. However, these advanced considerations are only of marginal relevance for the current thesis' main questions. The main conclusion to keep in mind from this section, is that perceptual simulation of various aspects of situation model-level text processing indeed occurs in both children and adults.

1.1.4 Empirical evidence of multi-level theories in studies with adults and children

Several studies have confirmed that adults simultaneously construct these three different levels of mental representation during text comprehension (e.g., C. R. Fletcher & Chrysler, 1990; Schmalhofer & Glavanov, 1986). These studies employed a so-called sentence recognition task to separately identify memory traces of the text surface, the text base, and the situation model. The sentence recognition task is based on the assumption that memory traces of the text surface, the text base, and the situation model are the products of cognitive processing on the respective levels (Schmalhofer & Glavanov, 1986).

In the study by Schmalhofer and Glavanov (1986) participants read 10 sentences that were part of a manual for the programming language LISP. They were told that their goal should be to either acquire knowledge or to be able to later summarize the text. For each of these original (O-) sentences (6 in the following example), the experimenters created three different types of distractor sentences. To obtain paraphrased (P-) sentences (7), words in the second clause were replaced with synonyms. Meaning-changed (M-) sentences were constructed by changing the second clause in a way that it was still correct regarding the rules of LISP (8). Finally, correctness-changed (C-) sentences (9) violated the rules of LISP.

(6) PSY100 is a legal atom that concludes with a number.

(7) PSY100 is a legal atom which ends with a numeral.

(8) PSY100 is a legal atom which begins with a letter.

(9) PSY100 is a legal atom that contains a dot.

P-sentences only differed from O-sentences with regard to the text surface, and the ability to discriminate these two sentences consequently indicates memory of the text surface. Both M-sentences and P-sentences differ from the O-sentences with regard to the text surface, but the M-sentences further differ with regard to the text base, while the P-sentences are consistent with the text base. Thus, discrimination of M-sentences from P-sentences indicates memory of the text base. Following this logic, the ability to discriminate C-sentences from M-sentences accordingly indicates memory of the situation model. During the recognition task, participants were presented with original and distractor sentences and asked to decide, whether they had read this sentence before or not. To indicate discrimination ability d' sensitivity scores were calculated based on approval ratings to every sentence type.

Their results indicate that memory of the different levels of processing differed depending on the reading goal. Readers who had been instructed to read with the goal

of text summarization, had better memory of text base information, whereas readers whose goal was knowledge acquisition had better situation model memory. A subsequent experiment with the same materials further showed that accessing situation model-level information seems to be faster than accessing text information. This is plausible because comparing verbatim information or propositional structures requires more cognitive effort, whereas making a decision on a correctness-changed sentence can be based on a judgment whether this sentence could have occurred in the text (Schmalhofer & Glavanov, 1986).

C. R. Fletcher and Chrysler (1990) employed a similar recognition test with texts as test materials that each described a certain order of a set of five objects. For example, a text deals with the value of hobby art collector George's recent purchases and implies the following order by price: carpet < painting < necklace < vase < statue. The distractor sentences all referred to this implied order. This system of text composition allowed a "cleaner" differentiation between changes of the propositional structure and changes of the situation: A change of the situation always implied a violation of the original linear ordering, and other than Schmalhofer and Glavanov (1986), the changes in the distractor sentences were limited to one single word. C. R. Fletcher and Chrysler (1990) found that participants were able to discriminate between original sentences and sentences that had been altered with respect to the text surface but within the scope of the propositional text base, indicating memory of the text surface. They were also able to discriminate between text-surface-changed sentences and sentences that had been changed at the propositional level in addition to the text surface, but within the scope of the situation model. Finally, discrimination was best for sentences that were also inconsistent with the situation described in the text. This result pattern thus supplements Schmalhofer and Glavanov (1986) by showing that memory traces for the three assumed levels of processing can be identified, also for narrative texts and independent of study goals.

Multi-level theories of text processing have been proposed and especially empirically tested in adult samples. Nieding (2006) conducted two studies to test whether these theoretical considerations are equally applicable to children. She developed a child-appropriate version of a sentence-recognition task with test materials that comprised auditorily presented narrative texts about topics children are usually interested in. The findings indicate that already 5-year-old children have memory traces of all three levels (p.90). Mean relative frequencies of "yes" responses significantly differed between original sentence and paraphrases, paraphrases and meaning changes, and between meaning changes and correctness changes. Nieding (2006) classified this result pattern as a "passed model test" (p. 94).

1.1.5 Summary: multiple levels of mental representations

In short, text comprehension is based on the recipient actively engaging with the text, resulting in a coherent mental representation. Considering both children and adults, this process is assumed to involve three different kinds of mental representations: a representation of the text surface that is still close to perception, a representation of the propositional text base that contains the text's semantic content, and finally a situation model that represents the whole situation described in a text. While the mental representations of the text surface and the text base preserve the representational format of the text, the situation model, according to the meanwhile established view of embodied language processing, is assumed to also include perceptual simulations of the verbally described situations.

1.2 Comprehension as a Process of Coherence Formation

The previous section described different types of mental representations that are involved in the process of text comprehension. The previous section also stated that comprehension means that a recipient constructs a *coherent* mental representation of the represented narrative. This section further elaborates the specific cognitive mechanisms necessary to achieve coherence.

A text is not just a random accumulation of sentences, but the sentences that make up a text logically build upon each other and thus form an integrated whole. The recipient is therefore required to not only understand every isolated sentence, but integrate the information of adjacent sentences with each other. This connection of consequent text information is called *local coherence*. Moreover, the recipient can aim further and achieve *global coherence* by connecting the current text information with the overall topic of the text or with information that occurred much earlier in the text and is no longer available in working memory (Graesser et al., 1997). As a text hardly ever describes a situation exhaustively and coherence can unlikely be achieved based on the information explicitly mentioned, the recipient has to fill the gaps that are left by making inferences.

1.2.1 Different types of inferences

As the situation model can be a detailed representation of the situation, the possible inferences that can be made in the process of constructing this situation model are plentiful. Plentiful are also the possibilities to organize different inferences based on their function. Inferences can be classified by distinguishing text-connecting and extratextual inferences, that is inferences that can be made based on the text itself and inferences that require further (or prior) knowledge. A recipient can generate inferences

on-line (during) reading, or off-line (afterwards). A further possibility is to classify inferences based on whether they are necessary or elaborative. As the notation implies, necessary inferences are pivotal to maintain referential and causal coherence (Keefe & McDaniel, 1993). One example would be backwards inferences that are generated to explain current text information with information that occurred earlier in the text (Magliano, Baggett, Johnson, & Graesser, 1993). Another example are anaphoric inferences that are decisive to connect subsequent clauses with one another. When reading the two sentences "Tim is hungry. He orders Pizza.", the reader, most certainly infers that the pronoun "He" in the second sentence refers to Tim. Without making this connection these two sentences would not make sense as a whole. Elaborative inferences are not needed to obtain a coherent mental representation and rather refine and extend a mental representation (Estevez & Calvo, 2000). These are for example predictive inferences that anticipate how a plot proceeds. A predictive inference, you may draw when reading the sentence "After standing throughout the long debate, the tired speaker walked over to her chair." is that the speaker will sit down (Estevez & Calvo, 2000). Whether an inference is necessary or elaborative to some degree depends on your assumptions on what inferences a recipient routinely generates and thus incorporates in the situation model. Available theoretical positions on this matter will be discussed in the next section.

1.2.2 Theoretical positions on inference generation and represenation of coherence relations

Over time, different theories have emerged, that make different assumptions about which of these inferences a reader routinely generates on-line. These theories thus differ in their assumptions about the scope and richness of detail of the situation model and ultimately in their idea of what signifies comprehension of a discourse.

All theories agree that at least some inferences are generated. The most parsimonious assumption that no inferences are drawn and a text is mentally represented only based on its explicit propositions has been shown to be highly unlikely (see also section 1.1.3, p.7).

The *minimalist hypothesis* (McKoon & Ratcliff, 1992) assumes that readers only routinely generate the inferences they need to establish local coherence based on information that is available in short term memory. This includes mainly referential or anaphoric inferences, or inferences about causal antecedents. Similarly, the *current-state selection strategy (CSS)* assumes that a reader first and foremost aims for local coherence that is achieved by making causal connections between adjacent propositions (C. R. Fletcher & Bloom, 1988; C. R. Fletcher & Chrysler, 1990). Compared to the minimalist hypothesis, this approach thus emphasizes causal rather than referential connec-

tions based on argument overlap (Graesser et al., 1994). According to the CSS, a reader achieves coherence when he or she connects a current clause with a causal antecedent under the precondition that this antecedent can be retrieved from short term memory. A causal antecedent remains available until the reader identifies a consequence. Therefore, according to the CSS, inferences about a character's superordinate goals are only drawn when establishing local coherence fails, in which case an earlier mentioned goal is retrieved.

According to the *constructionist theory* of text comprehension, coherence formation can be described as a process of "search after meaning" (Graesser et al., 1994). Graesser et al. (1994) postulate three fundamental mechanisms to underlie text comprehension. First, they assume, that comprehension is driven by the reader's goal. This implies that the mental representation of the same text can differ between different readers with different reading goals, while at the same time every reader usually aims for a semantic representation and does not settle for surface-level representations (reader goal assumption). Secondly, the constructionist theory, other than the minimalist assumption or the current state selection strategy assumes, that a reader attempts to achieve both local and global coherence (coherence assumption). Thirdly, comprehension as a process of search-after-meaning implies that the reader actively engages with the text, trying to make sense of every new incoming piece of information by answering the question why this piece of information is mentioned and how it relates to the rest of the text processed so far (explanation assumption). These three assumptions further imply that readers generate inferences about characters' superordinate goal, causal antecedents of actions, states and events, and about the global theme.

Based on their constructionist theory of text comprehension, Graesser et al. (1994) propose a useful and further refined overview of different types of inferences, in which they differentiate between 13 different types of inferences that they assume to possibly occur during discourse processing. Referential inferences, inferences that assign a case structure role or specify a causal antecedent are critical for establishing local coherence. Inferences about a protagonist's superordinate goal, the overarching, main topic or moral of a text as well as inferences about a character's emotional reaction contribute to establishing global coherence. Further inferences, such as for example predictive inferences about causal consequences, are considered elaborative and not needed to achieve a coherent mental representation. Inferences may also refer to the pragmatic communicative context, such as inferences about the author's intention behind a text. These can also be considered elaborative inferences. Graesser et al. (1997) even state that the pragmatic communicative context may be considered a further level of processing, in addition to the level of the situation model (see also section 1.1).

Graesser et al. (1994) emphasize that there is no guarantee that everyone succeeds in his or her attempt to construct a coherent referential situation model, tailored to the specific reading goals that explains why actions, events, and states are mentioned. For example, establishing global coherence may fail because a reader either lacks specific prior knowledge, or because the text composition makes it difficult or impossible to identify coherence relations (Graesser, Wiemer-Hastings, & Wiemer-Hastings, 2001).

The constructionist perspective on text comprehension specifically emphasizes the importance of causal relations when it comes to establishing and maintaining coherence. This view is supplemented by the *event-indexing model* (Zwaan, Magliano, & Graesser, 1995) that postulates that a reader constructs a multi-threaded situation model, requiring him or her to monitor five different dimensions. Other than in real-life, an omniscient narrator can move freely between different times and locations, and causes and effects are not necessarily presented in the right order. Also, a protagonist's goals and motives may be revealed step by step. In order to construct a comprehensive situation model, a reader has to keep track not only of causality, but also of the protagonist, temporality, spatiality, and intentionality. The latter aspect refers to a character's goals that provide explanations for his or her actions and maintaining coherence on this dimension thus requires keeping track of described goals and connecting them with actions.

Nevertheless, the findings of various studies indicate that readers predominantly rely on causal relations in order to establish and maintain a coherent representation of a narrative (e.g., Bohn-Gettler, Rapp, van den Broek, Kendeou, & White, 2011; Rapp, van den Broek, McMaster, Kendeou, & Espin, 2007). Such findings corroborate the importance of causal relations for situation model construction.

1.2.3 Overview of empirical studies on inference generation and represenation of coherence relations

Plenty of studies have addressed the question whether readers usually attempt to construct a situation model that is coherent both on a local and a global level, as the constructionist theory would predict, or whether, in accordance with the minimalist hypothesis or the current-state selection strategy, only local coherence is established. Overall, research suggests that adult recipients generally attempt to construct a situation model that is both locally and globally coherent (Albrecht & O'Brien, 1993; Dopkins, Klin, & Myers, 1993; Graesser et al., 1994; Hess, Foss, & Carroll, 1995; Huitema, Dopkins, Klin, & Myers, 1993; Myers, O'Brien, Albrecht, & Mason, 1994; Suh & Trabasso, 1993; Rizzella & O'Brien, 1996; Poynor & Morris, 2003). In the following, I will describe some of these studies exemplarily in further detail, also in order to familiarize

the reader with the methods usually used to assess the coherence relations that a reader (inferentially) establishes.

In a series of experiments, Rizzella and O'Brien (1996) examined whether readers accessed distant causal antecedents in locally coherent texts when more recent causal antecedents are available. They further examined the effects of elaborating an early causal antecedent. Participants in their experiments read texts such as the narrative of Billy who loses his key, that is displayed in Table 1.

Table 1. *An example story used in the experiments by Rizzella and O'Brien (1996).*

Neither causal antecedent elaborated	Early causal antecedent elaborated
Billy was walking home from school after playing a game of basketball. Billy looked for his keys to unlock the front door of his house. He searched everywhere but couldn't find the keys. Billy needed to find another way to unlock the door. In order to unlock the door, Billy broke a small window. The window fell in pieces on the ground. He knew that once his father came home he would be in trouble.	Billy was walking home from school after playing a game of basketball. Billy looked for his keys to unlock the front door of his house. He searched everywhere but couldn't find the keys. *He realized there was a big hole in his pocket. Now, he had no idea where to look. Billy shuddered when he recalled the warning his father gave him about being more responsible. His father told him that if he was not more responsible, he would ground Billy for an entire month.* Billy needed to find another way to unlock the door. In order to unlock the door, Billy broke a small window. The window fell in pieces on the ground. He knew that once his father came home he would be in trouble.

This story provides an early and a late causal antecedent for the consequence that Billy would be in trouble. The early antecedent (Billy lost his keys) is represented by the underlined word "key". The late antecedent (Billy broke the window) is represented by the underlined word "window". Rizzella and O'Brien (1996) created a second version of each of their experimental stories, in which the early causal antecedent was further elaborated (see the right column of Table 1. The additional text that elaborates the early antecedent is printed in italics.). Naming latencies of the probe words representing the early or the late causal antecedent were collected (Experiment 1A). The rationale behind this procedure is that the time needed to name a probe word depends on how readily available the concept behind this probe word is in working memory. Availability in working memory in turn depends on whether the antecedent is kept in working memory to be connected with a consequence. A reader who aims for only local coherence would not have the early antecedent available as a reason for why Billy would get into trouble, especially because a consequence, following the early antecedent (He needs to find another way to open the door) is also provided. The results indicate that when nei-

ther causal antecedent was elaborated the mean naming time was significantly faster for probe words associated with late antecedent. When the early causal antecedent was elaborated, the mean naming time did not differ, indicating that the early antecedent was available to the same extent as the late causal antecedent. This finding suggests that readers try to form coherence both locally and globally. The findings of a subsequent experiment (1B) suggest that it is the consequence implied by the last sentence ("He knew that once his father came home he would be in trouble." in the example text presented in Table 1) that (re-)activates the distant cause: Naming times of probe words associated with the early and the late causal antecedent did not differ when this last sentence was removed from the text.

Other studies found that a superordinate goal is inferred even when not elaborated. Huitema et al. (1993) recorded reading times of sentences that either were consistent or inconsistent with an earlier introduced goal. For example, participants read a story in which the protagonist wants to take a vacation somewhere where he can swim and sunbathe. This introduction was then followed by some backgrounding sentences about how the protagonist buys a travel guide and looks at newspaper ads. The consistent target sentence in this example would be "He went to his local travel agent and asked for a plane ticket to Florida.", whereas by replacing Florida with Alaska, this sentence became inconsistent with the goal.

The rationale behind this task is that a reader who has integrated the character information into his or her mental representation of the story, takes more time reading a sentence that contradicts this character description because he or she tries to resolve this inconsistency. Indeed, Huitema et al. (1993) found that the reading times of the sentences inconsistent with an earlier introduced goal were higher than those of the corresponding sentences that were consistent with the goal (experiment 1). This finding could be confirmed in subsequent experiments in which the backgrounding sentences did not describe actions related to the goal, or in which these sentences introduced another goal (experiments 2 and 3, Huitema et al., 1993). Poynor and Morris (2003) replicated these findings using the same materials but with the goal only stated implicitly and therefore to be inferred by the recipient.

Adult readers' ability to establish local and global coherence however seems to be affected by working memory span (reading span, Daneman & Carpenter, 1980), as found by Whitney, Ritchie, and Clark (1991). They conclude from think-aloud protocols of adults that had read a difficult text passage that low-span readers either struggled with regard to global coherence because they allocated available capacities to the connection of subsequent phrases, or they focused especially on global coherence as a top-down processing strategy that helps compensating for low working memory capacity.

1.2.4 Developmental aspects: (How) do children infer and represent coherence relations?

As with the earlier discussed aspects of discourse comprehension (section 1.1), also available theories about inference generation and coherence formation are predominantly based on research with adult readers. Therefore, the question remains if and to what extent these theories also apply to children of different age groups.

Overall, children are able to make the necessary inferences to understand causal connections when following a narrative already at an early age, whereas at the same time, this ability improves and expands with age, as implied by a number of studies that will be described in further detail in the following.

In their study, Thompson and Myers (1985) presented 4- and 7-year-old children three different versions of child-appropriate short narratives. In one version, the connection between a cause and an event was physical, meaning that a mechanical cause lead to a target event. In another version, the connection between a cause and an event was psychological. Here, for example a character's emotional state lead to a target event. In yet another version, the connection was enabling, that is the cause of a target event was a necessary but not sufficient precondition. Afterwards, the children were asked questions that required them to infer (among others) causes of story events. The results indicate that 4-year-olds perform at the level of 7-year-olds with regard to inferences that link a physical cause to an event. It seems however, that the ability to infer psychological causes still develops between the ages 4 to 7. This finding aligns well with the findings of a study by Whiteman (1967) that imply that children's understanding of causality in general develops with age. He interviewed two groups of children aged 5-6 and aged 7-8 about the motivations of protagonists in story situations. This research was motivated by Piaget's (1929, 1930) assumptions about the shift from the preoperational to a concrete operational stage that comes with the ability to decenter the focus of attention away from salient features, and in turn the ability to also consider underlying thoughts and motivations. While Whiteman (1967) found that children's understanding of both psychological and physical causality develops with age, there seems to be even more room for improvement with regard to the understanding of psychological causality.

Studies with older children indicate further development. Findings of studies that included participants aged 6 and 7, 9 and 10 as well as 18 years indicate that the ability to infer plausible causes for an unexpected story outcome increases with age (Ackerman, 1986; Ackerman & McGraw, 1991). In these studies, the participants read short narratives that end with an outcome inconsistent with an earlier introduced protagonist's goal and thus invite the recipient to infer a reason for this unexpected outcome. For example, a story deals with a woman who plans to wash clothes and dry them using the

dryer. Later, she is described hanging the clothes outside on the clothesline, instead of putting them into the dryer. Afterwards, the participants answered forced-choice questions (yes/no) that described a referential inference, an inference presenting a plausible explanation for the story outcome, and an inference describing an implausible outcome. The authors varied the amount of clues that served as possible explanations for the outcome inconsistent with the goal. The results indicate that the ability to make inferences increases with age. In general the 18-year-old college students made more inferences to establish causal and referential coherence than both the 7- and the 10-year-olds. This effect of age especially showed when there were no clues available to explain a goal-inconsistent outcome, as Ackerman (1986) only found an effect of clue support among the youngest age group.

Especially with regard to the studies by Thompson and Myers (1985) and Whiteman (1967) it has to be pointed out that when asking questions that require oral answers, children's verbal abilities may constitute a confounding factor.

Two experiments by van der Schoot, Reijntjes, and van Lieshout (2012) not only employed alternative methods that are less dependent on verbal abilities but also examined reading comprehension skills as a possible factor that may affect inference generation, or maintaining coherence relations respectively. Van der Schoot et al. (2012) used the same inconsistency-detection task that Huitema et al. (1993) and Poynor and Morris (2003) used in studies with adult participants. They examined if 10- to 12-year-old students who differed with regard to their reading abilities could uphold over longer text distances information about a story character, which would enable a recipient to detect and/or resolve an occurring inconsistency within a text. As indicators of the ability to resolve inconsistencies, van der Schoot et al. (2012) recorded self-paced reading times (Experiment 1) and sentence fixation durations and regressions to the character description via eye-tracking (Experiment 2). It was further varied, whether the information about the story character was directly followed by an inconsistent sentence (local condition), or separated by a filler paragraph (global condition). The results of both experiments indicate a difference between good and poor comprehenders in the global condition, but not the local condition. When inconsistency detection required the readers to uphold the relevant information over a longer time period (as in the global condition that contained a filler paragraph), poor comprehenders did not take more time to read the inconsistent target sentence, nor did they regress more to the character description, presumably because the information no longer resided within the working memory.

Helder, van Leijenhorst, and van den Broek (2016) report that 8- and 9-year-old children also differed with regard to their abilities to detect inconsistencies in a narrative depending on their reading comprehension ability. However, this study did not differ-

entiate between a local and a global condition. Both studies (van der Schoot et al., 2012; Helder et al., 2016) controlled for decoding skills, thus explicitly defining comprehension skills as linguistic comprehension skills, which is the ability to assign meanings to words and sentences (for a more detailed discussion of the difference between decoding and linguistic comprehension skills see the section 1.3.2, p. 33)

Further research expands the scope to other dimensions that are monitored during comprehension and thus situation model construction as assumed by the event-indexing-model (see section 1.2.2, p. 16; Zwaan et al., 1995). While adults seem to routinely monitor all dimensions, 12-year-olds still only seem to monitor causality, as indicated by reading times of sentences that indicate a shift within the respective dimensions (Bohn-Gettler et al., 2011). Wassenburg, Beker, van den Broek, and van der Schoot (2015) used a task based on the same inconsistency paradigm as van der Schoot et al. (2012) with the variation that the inconsistencies not only referred to character information but also to other situational dimensions, such as emotion, causation, time, and space. This study indicates that children aged 8 to 13 seem to monitor causality and emotions, but not spatial and temporal information even at the local condition. At the global condition only emotionally charged information was monitored.

In addition to this finding, some studies indicate that goal structures play a crucial role in children's comprehension of narratives as they constitute a first organizing principle and are a prominent concept when acquiring a theory of mind (Sodian & Thoermer, 2006). Goals have been found to provide coherence for children who retell narrative texts (Lynch & van den Broek, 2007; Lynch et al., 2008; Perfetti, Landi, & Oakhill, 2005). For instance, Lynch and van den Broek (2007) found that goal-based inferences that children aged 6 and 8 made according to think-aloud-protocols predicted their recall performance. Also, when understanding films, macro-structural information about protagonists' superordinate goals help 7-year-olds resolve temporal ellipses based on global coherence structures (Nieding, Ohler, & Thußbas, 1996). Moreover, Nieding (2006) found that 5-year-olds reconstructed protagonists' movements in auditorily presented texts more easily when these protagonists pursued superordinate compared to subordinate goals.

In conclusion, it seems that the ability to establish and maintain both local and global coherence develops with age. Furthermore, this process is associated with text comprehension skills and while an understanding of causality seems to be available already at an early age, monitoring causal relations within the situation model is apparently easier than monitoring other dimensions. This seems to hold true even in children at least up to the age of 12. In younger children, also protagonists goals have been found to provide coherence during comprehension of narratives.

1.2.5 Summary: mental representation of coherence relations

Research on text comprehension implies that an adult recipient usually attempts to achieve a mental representation that is coherent on a local as well as a global level. This is a demanding process and success depends on text characteristics as well as on an individual's characteristics, for example working memory span. The ability to represent coherence relations of a text also develops with age, while an understanding of causality and thus the ability to monitor causality within the situation model develops early in life. Research with children also indicates that the ability to maintain coherence relations depends on children's text comprehension skills.

1.3 Effects of Presentation Mode: Theoretical Approaches and Empirical Evidence

In their everyday lives, both children and adults are confronted with a variety of media that contain not only written text, but also auditory text, which may also be combined with pictures.[1] The basic underlying cognitive processes, as described in the two preceding sections, are supposedly very similar, irrespective of whether a text is presented auditorily, audiovisually (i.e. audio combined with pictures), or in written form. At least for adults there is evidence of a general comprehension skill which extends beyond non-verbal materials such as picture stories (Gernsbacher, 1990). Nevertheless, these different forms of text differ in terms of the signs they contain (verbal or pictorial) and the perceptual system (auditory or visual) through which they are perceived. This section elaborates possible differences in cognitive processing and their implications for comprehension between verbal information and pictures, and between auditory and written texts.

1.3.1 Cognitive processing of verbal information and pictures

Before examining how the presentation mode can affect the cognitive processing of information, it is decisive to take a closer look on the different signs that an audiovisual text is made up of.

Words and pictures consist of two fundamentally different sign systems (Schnotz, 2005). Words are symbolic signs that do not resemble their referent (C. Peirce, 1906; Schnotz, 2002). Neither the letter sequence "c-a-r" nor the sound of this word ([kaːr])

[1] This section is based on the introductory sections of the following papers: Wannagat, W., Waizenegger, G., & Nieding, G. (2017). Multi-level mental representations of written, auditory, and audiovisual text in children and adults. *Cognitive Processing*, *18*(4), 491-504. doi: 10.1007/s10339-017-0820-y, Wannagat, W., Waizenegger, G., & Nieding, G. (2018). Mental Representations of the Text Surface, the Text Base, and the Situation Model in Auditory and Audiovisual Texts in 7-, 9-, and 11-Year-Olds. *Discourse Processes*, *55*(3), 290–304. doi: 10.1080/0163853X.2016.1237246.

resembles an actual car. The relation between a symbol and its referent is arbitrary and based on conventions (De Saussure, 1960). This implies that the recipient has to be familiar with these conventions in order to understand the meaning of a symbol. The organization of verbal symbols is determined by (grammatical) rules and also the processing sequence is given, such as for example, from left to right and from top to bottom (Eysenck & Keane, 2000). Pictures, by contrast, consist of iconic signs. Iconic signs represent a referent via common structural properties, such as superficial/physical analogy. There is neither a predefined repertoire of signs nor are there syntactic rules comparable to the rules applicable in language.

With regard to pictures, a further disctinction has been introduced by Hegarty, Carpenter, and Just (1996), who differentiate between three different types of pictures. *Iconic drawings* are realistic pictures and illustrate a situation as it occurs, and thus there is a high correspondence between the picture itself and the subject matter it represents. Typical examples are drawings, naturalistic paintings, photographs, or cartoons. In terms of *schematic diagrams* and *charts & graphs*, on the contrary, the analogy relation is not based on physical but on structural properties and these types of pictures are therefore especially useful to illustrate abstract concepts, in which connections and other structural properties are not directly accessible for perception. *Schematic diagrams* depict structural and qualitative relations, such as in organigrams or circuit diagrams. *Charts & graphs*, for example pie or bar graphs, depict quantitative relations. The present thesis focuses on illustrated, auditorily presented text and thus on what Hegarty et al. (1996) call *iconic drawings*.

To explain supposedly beneficial effects of pictures on the cognitive processing of verbal information, there are several theories available. In the following, I will describe three different theoretical accounts shortly with a focus on the components and processes that allow predictions about the processing of text and pictures.

1.3.1.1 Paivio's dual coding theory

Paivio's dual-coding theory (Paivio, 1986, 1969, 1975, 1991) first of all acknowledges the previously described obvious differences between words and pictures. Paivio distinguishes between two independent cognitive subsystems for the processing of verbal (linguistic) and non-verbal information, each with its own representational format: logogens within the verbal, and imagens within the nonverbal system (Eysenck & Keane, 2005).

The verbal system is specialized on the sequential processing of linguistic information. This linguistic information is represented by logogens, which are assumed to be modality-specific. The word "car" can be represented by the sound of the word or its visual form (the letter sequence "c-a-r").

The non-verbal system engages in the processing of spatial and synchronous information, among others, such as in scene or picture perception and analysis. This system is further assumed to contain subdivisions specific to the visual, the auditory, and the haptic sensory modality, because the imagens (like the logogens) are specific to the original sensorimotor modality.

In conclusion, the dual-coding theory essentially assumes that words and pictures are processed independently by means of distinct cognitive mechanisms. Thus, pictures added to words provide additional memory code (J. Fletcher & Tobias, 2005).

1.3.1.2 Baddeley's conception of a modular working memory

In his widely accepted conception of a modular working memory, Baddeley (1986, as cited in Eysenck & Keane, 2005) differentiates between two independent subsystems for the processing of phonological and visuo-spatial information, which are coordinated by a central executive. [2]

Visuo-spatial information is processed by the *visuo-spatial sketchpad*, which is further subdivided in subcomponents for different object features, such as color and shape, and for spatial information.

The other main component is the *phonological loop*, which processes acoustic and articulatory information. This includes both spoken and written langugage. Visually presented verbal information (= written text) is translated into articulatory code and then stored within the phonological loop. Subvocal rehearsal, a kind of inner speech holds this information available within the phonological loop. In case this process is suppressed, phonological memory traces decay within 1-2 seconds. This means that every incoming information that is not articulated within 1-2 seconds decays and is no longer available for further processing. Both components are assumed to have only limited capacity. At the same time, these two components can operate independently from one another.

These assumptions have an important implication. When simultaneously performing two tasks, performance in both tasks is more impaired when these tasks take up the capacity of the same component compared to when these two tasks use different components. A vast amount of dual-task studies provide evidence for the assumption of two separate working memory components. For example, performance in a verbal task, such as holding up a sequence of consonants is less disturbed by a visual-spatial task, such as pressing keys on a keypad following a simple sequence, than by a concurrent task that requires resources of the same component.

[2] Later, Baddeley (2000) extended his model by proposing the episodic buffer as a further component of the working memory. As this component is not central for the purpose of this thesis, it is left out of further discussions of Baddeley's model.

According to this model, the processing of words takes place within the phonological loop, whereas the processing of pictures takes place within the visual-spatial sketchpad. Simultaneous presentation of the same information in form of words and in form of pictures means that the same information can be processed twice at the same time and thus efficiently makes use of the given cognitive prerequisites.

Graesser et al. (2001) make some points about the differences between text and film with regard to inference generation that are related to the idea of a working memory of limited capacity: A viewer, in contrast to someone reading a book of the same content, directly observes the spatial layout of an environment, visual features about objects and people, the actions that are performed by people, and events that are visually prominent. However, a viewer still has to generate inferences that *explain* visible actions, events, and states; these inferences include the superordinate goals, causal antecedents, emotions, and perhaps traits of people. From a capacity point-of-view, one might therefore assume that when visualizations of the situation are available that can replace some sorts of inferences to a certain degree, cognitive resources may be allocated to the inferences about things that cannot be directly observed. This consideration would then allow the assumption that pictures may support inferences and ultimately text comprehension.

1.3.1.3 The instructional psychologists' perceptive: theories of multimedia learning

Against the background of a debate about the promises and risks of "multimedia" for comprehension and learning, instructional psychologists have proposed theoretical models that combine the previously described basic assumptions about the nature of human cognitive processing and the cognitive architecture in which it occurs into comprehensive frameworks about the cognitive processes that underlie learning with different representational formats, such as words and pictures.

It is important to keep in mind that this group of theories specifically deals with *learning* and thus goes a step further than the actual topic of this thesis which deals with *comprehension*. Learning means that the mental representations resulting from cognitive processing that led up to comprehension are integrated and aligned with the knowledge stored in long-term memory in form of cognitive representations. As a result of this process, cognitive representations are permanently modified to account for this newly processed information.

The essence of two influential theories, the Cognitive Model of Multimedia Learning (CTML, Mayer, 2005a) and the Intergrated Model of Text and Picture Comprehension (ITPC, Schnotz & Bannert, 2003; Schnotz, 2005, 2014) will be described in the following.

The cognitive theory of multimedia learning. The CTML assumes a cognitive architecture based on Atkinson and Shiffrin (1968) that consists of a modality-specific sensory memory, a working memory of limited capacity, and a long-term memory as a knowledge storage. Furthermore, the CTML is based on three assumptions. The *dual channels* assumptions, based on the previously described theories of Paivio and Baddeley postulates two channels, one for auditory/verbal and one for visual/pictorial information. Both channels are of *limited capacity*, but operate independently from one another. Learning is assumed to be a result of *active processing*.

First, information perceived via the eyes or ears is briefly upheld in sensory-specific memory stores. During this period of time, an individual decides[3] which pieces of information are relevant and thus further processed. Relevant auditory information, such as spoken language is processed in an auditory/verbal channel within the working memory, whereas visual information is processed in a visual channel. Written text that is initially perceived visually and that is therefore also processed in the visual channel is assumed to switch to the auditory/verbal channel some time during further processing. The cognitive processing within the working memory that ultimately results in meaningful learning involves selecting relevant items (words or pictures) and organizing these into a verbal and a pictorial mental model. These two models are then integrated into a mental model that contains the information of the verbal and the pictorial model and prior knowledge stored in long-term memory.
Although integrating information from two sources can be demanding, the independent operation of the two channels enhances the overall capacity of the working memory (J. Fletcher & Tobias, 2005).

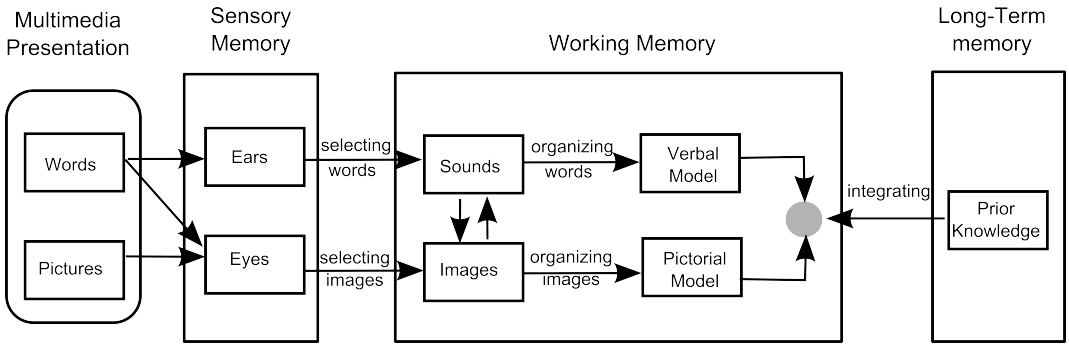


Figure 2. The cognitive theory of multimedia learning. Adapted from Mayer (2005). The grey circle represents the (situation/mental) model in which the information of the verbal and the pictorial model have been integrated with each other and with prior knowledge.

[3]This process (as are the processes that will be described in the following) is actually not as conscious and deliberate as the use of the verb "decide" implies in this context. However, to convey the general idea that information processing and comprehension is an active and constructive (though not always conscious) process, I occasionally resort to maybe inappropriately "active" vocabulary.

The integrated model of text and picture comprehension. Very similar assumptions are made by the Integrated Model of Text and Picture Comprehension (ITPC, Schnotz & Bannert, 2003; Schnotz, 2005, 2014). The ITPC assumes the same Atkinson-and-Shiffrin-inspired cognitive architecture as the CTML that includes modality-specific sensory registers, a working memory, and a long-term memory. Although sensory registers can be assumed for every sensory modality, this model is limited to an auditory (ear) and a visual (eye) register. Information is perceived and further transferred to the working memory via the channel specific to its original sensory modality. Further processing in the working memory takes place in two different channels. One channel processes verbal information (written and spoken text), the other channel processes pictorial information (pictures or non-verbal sounds).

As in the CTML, the processing of both text and pictures is seen as an active process of coherence formation that includes multiple levels of mental representation for the processing of both words and pictures, and thus also refers to Kintsch and van Dijk's assumption in terms of text processing in general (see sections 1.1 and 1.2).

The ITPC distinguishes between a perceptual and cognitive level of processing. The perceptual level refers to the transfer of information between the environment and the working memory and thus to the processes within the sensory registers. The cognitive level refers to processing within the verbal and pictorial channel of the working memory and to the processes involved with integrating information stored in the long-term memory with information that is currently processed within the working memory.

Processing on the perceptual level is specific to the original sensory modality of the processed information. This implies that both a text and a picture printed on paper are first processed by the visual sensory register and transferred to the visual working memory although a text and a picture are assumed to be different representational formats.

Only on the cognitive level, the representational format determines how the information is further processed. Both written and auditory verbal information as symbolic signs are processed within the verbal channel. Pictures as iconic signs are processed within the pictorial channel. Figure 1.3.1.3 provides an overview of the features of the ITPC. This figure allows reconstructing the processes involved in comprehension of words and pictures. For example, auditory text as auditory verbal information is perceived by the ear and is, via the auditory channel, transferred to the auditory working memory where it is represented as a text surface. A verbal filter (represented by a black triangle in Figure 1.3.1.3) selects verbal information to be further processed within the verbal channel to be then represented as a propositional representation. This propositional representation is afterwards used to build a mental model.

A picture as visual pictorial information is perceived by the eyes and is, via the visual channel, transferred to the visual working memory where it is stored as a visual-perceptual representation. A visual-pictorial filter selects pictorial information, which is then be further processed within the pictorial channel to be finally represented as a mental model. This mental model can be used to read off new information that is encoded within the propositional working memory. Mental model construction within the pictorial channel is assumed to consist of processes of analog structure mapping (Gentner, 1983). This means that the recipient constructs a mental image of the perceived elements of a picture by mentally reproducing or *modeling* its structural features.

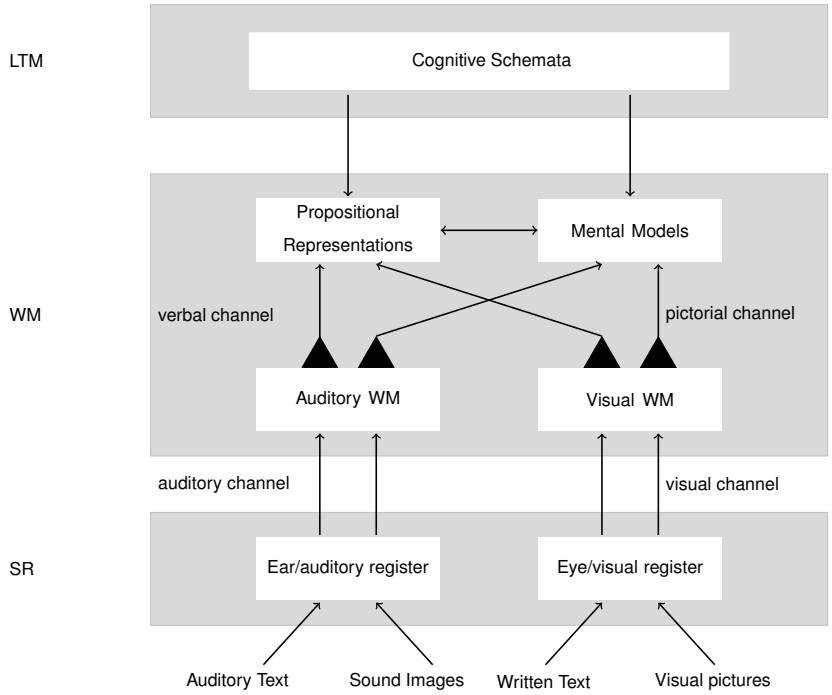


Figure 3. The Integrated Model of Text and Picture Comprehension, adapted from Schnotz (2005). "LTM" refers to long term memory, "WM" to working memory, and "SR" stands for "sensory registers". Schnotz (2014) presents a revision of this model that contains more details. As the key components remained the same, I display the 2005-version for reasons of clarity and comprehensibility.

Comparison of the CTML and the ITPC. The CTML and the ITPC are very similar in terms of the general assumptions about the integrated processing of verbal information and pictures. However, the two models differ in terms of their assumption on how this integration takes place. The CTML assumes specific verbal and pictorial (mental) models that are integrated into an overall model in a subsequent step. In contrast, the ITPC assumes only one mental model that integrates the propositional information origina-

ting from the processing of verbal information and the visual-perceptual representation of a picture.

These models may be criticized for oversimplifying the cognitive processes involved in the processing of words and, especially in case of the CTML, the processing of pictures. Although one may say that the ITPC makes more detailed assumptions about the specifics of the processing of words and pictures than the CTML and it is more closely oriented on the prevailing theories of text comprehension.

Nevertheless, both models contribute to our understanding on how an *integrated* processing of verbal information and pictures works out and the authors of both models derived several assumptions about the usefulness of pictures added to words and the conditions under which pictures may not be useful after all. Two of these assumptions, the multimedia effect and the redundancy effect will be described in the following

The multimedia effect. As an overall message, you may infer from the above describes theories what Mayer (2005a) calls the *multimedia effect*, that people learn better from words and pictures than from words alone. Distributing a certain amount of information across two channels, both on the level of sensory-specific and cognitive processing, ensures that the limited capacity of the working memory is used efficiently.

This positive effect is however tied to certain prerequisites. The information from both sources (words and pictures) has to be present in the working memory at the same time. As the working memory only has limited capacity, words and pictures have to be semantically interconnected (coherent) and need to be presented in a temporal and spatial contiguity, depending on whether verbal information is presented auditorily or in written form (Schnotz, 2014).

The redundancy effect. Other than predicted by Paivio's dual-coding theory, and other than a general positive multimedia effect would imply, the two different memory codes provided by words and pictures, do not necessarily have a positive effect on comprehension and learning. According to the ITPC (Schnotz, 2014), adding pictures to verbal information can also have detrimental effects. For experts, i.e. people with high prior knowledge concerning the topic in question, verbal information may be sufficient for constructing a situation model. Additional pictures that depict something that the recipient has already understood based on verbal information nevertheless draw attention. In this case, pictures impose an unnecessary cognitive load that interferes with meaningful processing and thus comprehension and learning. This effect has also been described as the expertise reversal effect within the framework of cognitive load theory (Kalyuga, Ayres, Chandler, & Sweller, 2003).

1.3.1.4 Some assumptions derived from embodied cognition accounts

As described earlier, embodied accounts of text comprehension assume text comprehension to be a process of mental simulation that results in a mental representation similar to that of a real-life experience. Mentally simulating the situation described in a text implies that the resulting mental representation contains perceptual features and thus represents a situation based on analogy. Comparably, as pointed out above, pictures also represent a situation in a visuospatial structure as an analogy based on common structural properties (Chun & Plass, 1997; Clement & Gentner, 1991; Gentner, 1983; Johnson-Laird, 1983; Schnotz, 1993; Vosniadou & Ortony, 1989).

Consequently, pictures as external representations and situation models as mental, i.e. internal, representation have something in common, and Schnotz (2005) points out that both a picture and a situation (or mental) model are *depictive* representations. This implies that a picture by itself or when supplementing a verbal narrative, by virtue of its analogous nature has the potential to initiate and support perceptual simulation and situation modeling of a narrative, or as Barsalou (1999, p. 69) puts it: "The visual information available in television makes linguistic indexing, and therefore comprehension, much easier".

1.3.1.5 Overview of empirical studies

Several studies and meta-analyses have explored whether audiovisual or auditory presentation is better for memory, comprehension, and learning of narrative text content in children aged 3 to 11 (Beagles-Roos & Gat, 1983; Carney & Levin, 2002; Gibbons, Anderson, Smith, Field, & Fischer, 1986; Greenhoot & Semb, 2008; Guttmann, Levin, & Pressley, 1977; Hayes, Kelly, & Mandel, 1986; Levie & Lentz, 1982; Pezdek, Lehrer, & Simon, 1984; Ricci & Beal, 2002). Overall, in line with the multimedia effect, at least for children, audiovisual presentation seems beneficial on a broad range of outcome measures. The following summary of studies, is restricted to studies that include an audiovisual condition that ensured temporal contiguity of the verbal and pictorial information.

When asked to retell a story, children from age 5 remembered more explicitly mentioned story details (Beagles-Roos & Gat, 1983; Gibbons et al., 1986; Greenhoot & Semb, 2008; Ricci & Beal, 2002) and story characters' activities (Hayes et al., 1986) when the story had been presented audiovisually as compared to a purely auditory version. The same was found for detail questions about story content (Beagles-Roos & Gat, 1983; Guttmann et al., 1977; Pezdek et al., 1984; Ricci & Beal, 2002) and questions requiring inferences to be drawn (Beagles-Roos & Gat, 1983; Pezdek et al., 1984; Ricci & Beal, 2002), as well as for the number of inferences incorporated when retelling a story (Gibbons et

al., 1986). Auditory presentation, in contrast, was only found to foster recall of verbatim information in two studies (Beagles-Roos & Gat, 1983; Hayes et al., 1986).

The majority of these studies included thoughtful discussions on the manipulation of the presentation mode, i.e. how the researchers ensured that the semantic content of both presentation modes was the same. In sharp contrast with the earlier introduced and discussed theoretical accounts about the cognitive processes of discourse comprehension, no detailed descriptions of the dependent variable (see sections 1.1 and 1.2) were provided.

Two newer studies (Diergarten & Nieding, 2015; Unsöld & Nieding, 2009) may be excluded from this criticism. Both studies carefully aligned their measures of comprehension to current theories of text comprehension. In addition, they employed online measures of text comprehension, which they assume would capture cognitive processing and comprehension as it happens, and not retrospectively as would recall questions.

In the experiments by Diergarten and Nieding (2015) children either watched scenes from the TV series *Lassie*, or listened to a corresponding audio-version. These scenes all implied that the protagonist, Timmy, would feel a certain positive or negative emotion (for example happiness). Each scene, or audio recording respectively, ended with an auditory statement of Timmy's emotional state, which was either congruent with the emotion implied by the scene ("Timmy feels happy."), incongruent but of the same valence as the actual emotion ("Timmy feels proud."), or incongruent and of the opposite valence ("Timmy feels sad."). The children then had to decide and indicate by pressing a key whether the emotion mentioned in the statement was a positive or a negative feeling. The rationale behind this procedure was that having inferred Timmy's emotion correctly would enable fast reactions in congruent examples. The results indicated that 8-year-old children benefit from an audiovisual, in this case TV, presentation in terms of emotional inferences. Similarly, Unsöld and Nieding (2009) found that 6-year-olds were better at drawing predictive inferences at audiovisual presentation, as indicated by the results of a naming task (see section 1.1.3, p. 7).

Although the above described studies differed in their definition of comprehension and therefore in their outcome measures, and although some of them lack a description of the theoretical foundation of comprehension as the dependent variable, some general conclusions about the three levels of processing (Kintsch, 1998; van Dijk & Kintsch, 1983; Zwaan & Singer, 2003) can be drawn from their results. Contrary to Salomon's (1984) assertion that television hinders comprehension, and more in line with theoretical considerations that rather suggest beneficial effects of adding pictures, audiovisual presentation seems to facilitate situation model construction, as the inference-related outcome measures used by Beagles-Roos and Gat (1983); Pezdek et al. (1984); Ricci and Beal (2002); Gibbons et al. (1986); Diergarten and Nieding (2015); Unsöld and Nieding

(2009) can be allocated in the situation model. The superior memory of story details and characters' activities following audiovisual text presentation may indicate a beneficial effect of audiovisual presentation for memory of the text base. However, the findings of Beagles-Roos and Gat (1983); Hayes et al. (1986), which suggest that auditory presentation fosters recall of verbatim information, contradict this interpretation.

According to the previously described redundancy principle, a general claim that pictures are beneficial for everyone and under every circumstance may be an oversimplification. Plenty of research indicates that adults do not generally seem to benefit from pictures added to verbal information, even when the condition that verbal and pictorial information is presented simultaneously (temporal contiguity) was met. For instance, adults seem to be better at content recall when information had been presented verbally (in written form) rather than audiovisually (Byrne & Curtis, 2000; Corston & Colman, 1997; DeFleur, Davenport, Cronin, & DeFleur, 1992; Furnham, Benson, & Gunter, 1989; Furnham & Gunter, 1987, 1985, 1987; Wicks & Drew, 1991), and no differences were found between verbal, in this case auditory, and audiovisual presentations with regard to the recall of facts (Christie & Collyer, 2008) and local bridging inferences (Tibus, Heier, & Schwan, 2013). Tibus et al. (2013) also found that pictorial information helps university students compensate local coherence breaks in an expository film. This finding emphasizes that when words and pictures are not redundant but supplement one another, pictures can promote understanding.

The role of prior knowledge or expertise during learning with combinations of words and pictures has been extensively studied as a determining factor of the effectiveness of multimedia learning, but not of narrative text comprehension. For example, Mayer and Gallini (1990) asked adult learners to study text-picture combinations about how brakes, pumps, or generators work. Learners with low prior knowledge or expertise on these topics, in particular, improved with regard to both retention and transfer tasks. Further, Mayer (2001) compared effect sizes of experiments that examined beneficial effects of multimedia learning between learners with low and high prior topic knowledge. He found median effect size differences of 0.60 for retention, and 0.80 for transfer tasks (see Kalyuga, 2005). Comparable with the above-mentioned studies with children, the variables included in these studies with adults (fact recall and local bridging inferences) also presumably measured memory for text-base-related information.

1.3.1.6 Developmental aspects of information processing: working memory and perceptual support

From a developmental perspective, there are some indications that the assumed beneficial effects of audiovisual presentation may depend on the recipient's age. According to Baddeley (2002) processing of auditory and visual information is said to occur

in different subsystems of the working memory, the phonological loop and the visuospatial sketchpad, as described above. This structure of the working memory can already be found in children as young as 4-years old (Alloway, Gathercole, & Pickering, 2006; Michalczyk, Malstädt, Worgt, Könen, & Hasselhorn, 2013). However, because the visuospatial sketchpad is more highly developed than the phonological loop in preschoolers before the age of five, they are more efficient in encoding visuospatial than auditory information (Hitch, Halliday, Schaafstal, & Schraagen, 1988; Pickering & Gathercole, 2001). Children seem to preferentially encode information visually until the age of eight, at which point they enter a dual coding stage before ultimately prioritizing verbal coding (S. Palmer, 2000). Younger children are also known to direct their attention to salient perceptual details such as size, shape, color, texture, and movement (Springer, 2001). Originally referred to as "perceptual boundedness" the notation implies this mechanism constitutes a deficiency to children's information processing. However, this mechanism has also been found to be adaptive in providing "perceptual support" and helping children to understand abstract concepts and support conceptual development (Ruggeri & Katsikopoulos, 2013; Springer, 2001).

In line with these considerations, Unsöld and Nieding (2009), for instance, reported that audiovisual presentation improved predictive inferences among 6-year-olds but not among children aged 8 and older.

1.3.1.7 Summary: cognitive processing of verbal information and pictures

Words and pictures constitute two different representational formats. Words can be defined as symbolic signs which do not resemble their referent, while pictures represent a situation based on structural or physical analogy. These differences have been accounted for in several theories that assume that the cognitive processing of words and pictures takes place in different and independently operating cognitive subsystems. Combining words and pictures thus efficiently utilizes the limited capacity of the cognitive system and can result in beneficial effects of added pictures, as shown in prior research with children. Adults however, do not necessarily benefit from a combination of verbal information and pictures, which can be explained by influences of expertise or prior knowledge.

1.3.2 Cognitive processing of different forms of verbal information: auditory and written texts

The obvious difference between reading a text and listening to a text lies in the sensory channel through which verbal information enters the cognitive system. The same information is either perceived as sound waves (auditory text), or as words on paper or

on a screen (written text). Different views exist on whether this difference affects cognitive processing and thus understanding. A unitary process view that assumes the same processes to determine both reading and listening comprehension, and a dual process view that assumes cognitive mechanisms specific to the presentation modality will be discussed in the following.

1.3.2.1 The unitary process view: considerations surrounding the simple view of reading

The simple view of reading (Gough, Hoover, & Peterson, 1996; Gough & Tunmer, 1986; Hoover & Gough, 1990) as a unitary process view assumes a general ability of receptive language comprehension or verbal intelligence and therefore that the same cognitive processes underlie reading and listening comprehension (Bäuerlein, 2014). According to the simple view of reading, reading comprehension is defined as a multiplicative combination of decoding skills and linguistic comprehension skills and both of these skills are considered to be necessary in order to understand the meaning of a text. Decoding skills refer to the ability to efficiently recognize words, and thus to decode the graphic stimuli (letters) into linguistic form (Hoover & Gough, 1990). Linguistic comprehension skills refer to the ability to capture the syntactic structure and the meaning of a sentence, and connect it to preceding and subsequent phrases or sentences. These abilities are basically the processes of situation model construction and coherence formation, described in the sections 1.1 and 1.2. This view is supported by studies showing that listening comprehension skills, which children acquire long before they learn to read, are closely associated with later reading comprehension skills (Kendeou, Bohn-Gettler, White, & van den Broek, 2008; Kendeou et al., 2005; Stanovich, 1991; Stanovich & Siegel, 1994; Stanovich & West, 1989). This linguistic comprehension skill is assumed to be the skill needed to understand both auditory and written texts. If linguistic comprehension skills are highly developed, differences in comprehension of texts presented in different modalities then should presumably arise from differently developed decoding abilities (Hoover & Gough, 1990).

Learning to read initially consists mainly of learning to decode letters into words. Accordingly, from an overview of a number of studies that examined the association of reading comprehension with decoding and linguistic comprehension skills (Curtis, 1980; Jackson & McClelland, 1979; J. Palmer, MacLeod, Hunt, & Davidson, 1985; Stanovich, Cunningham, & Feeman, 1984; Stanovich, Nathan, & Vala-Rossi, 1986), Hoover and Gough (1990) conclude that, although both decoding and linguistic comprehension skills correlate with reading comprehension, at first decoding is more strongly correlated with reading comprehension, while in the further course of development this association weakens and linguistic comprehension becomes more closely associ-

ated with reading comprehension skills. Hence, the simple view of reading would assume that, due to the influence of decoding on readings comprehension, it should not matter whether you read or listen to a text when sufficient decoding skills are available. It follows that until decoding skills are fully automated, a recipient would understand better when verbal information is provided in auditory rather than in written form. This assumption is supported by studies that did not find differences in terms of the recall of text details between written and auditory text presentations in adults who presumably are skilled decoders (Corston & Colman, 1997; Sachs, 1974; Wicks & Drew, 1991).

1.3.2.2 The dual process view and an integrated model

In contrast, the dual process view assumes that cognitive processing and mental representations differ depending on whether someone reads or listens to a text. Advocates of this view argue that different presentation modes make different and specific demands on the recipient, especially in terms of the demands made on working memory (e.g., Danks & End, 1987; Spearritt, 1962). Reading and writing require different processing strategies because style, grammar, and vocabulary differ between spoken and written language. A text written to be read, such as a psychology text book for example, is assumed to be best comprehensible when actually written and not consumed as an audiobook. For a text that is written to be delivered auditorily, such as a politician's speech, it may be the other way around (Danks & End, 1987) because the speaker's gestures, facial expressions and intonation support the semantic content of the speech, or emphasize certain passages. In general, when reading a text to someone intonation or prosody can be specifically used to provide clues to the listener on how to structure and organize the text that are not available in the written version. This is why reading may be assumed to require more efficient metacognitive strategies (see also Bäuerlein, 2014; Rost & Buch, 2010; Danks & End, 1987). However, the same can be achieved by graphic organizing clues in texts, such as headings and paragraphs (Ballstaedt, 1997). Written text is usually stable and the reading speed can be adapted to the individual (linguistic) processing speed and would thus be advantageous over auditory text (Ferreira & Anes, 1994; Imhof, Echternach, Huber, & Knorr, 1996).

Findings that indicate differences between auditory and written text presentations would speak in favor of this dual process view. First of all, short term memory for lists of numbers or words has been found to be better when these were presented auditorily as shown by Penney (1989). This finding indicates that the short term memory is less vulnerable to external disturbances when it is engaged with the processing of auditory information (Paechter, 1997; Baddeley, 1986). In addition, some studies that addressed differences in (higher-level) semantic processing found differences between reading and listening comprehension. Adults have been found to perform better after

a written compared to an auditory text presentation (DeFleur et al., 1992; Furnham et al., 1989; Furnham & Gunter, 1987). These studies were interested in the communicative effectiveness of different media channels and compared how well (college) students were able to recall details or facts from print or radio news or commercials respectively. Studies that found reading to be superior to listening especially when texts are difficult indicate an effect of the assumed advantage in terms of the recipient's control over processing speed that is given when reading but usually not when listening to a text (Kürschner & Schnotz, 2008).

Studies that consider the participants' age when comparing reading and listening comprehension abilities indicate that sufficient decoding skills seem to be necessary in order to benefit from a written text presentation. Sticht and James (1984) state that it may take up to the twelfth grade until decoding skills have become automatic to an extent that reading comprehension exceeds listening comprehension. Up to the seventh grade, listening comprehension exceeds reading comprehension, followed by a period in which neither is superior. In line with these considerations, Hildyard and Olson (1978) found better performance after an auditory text presentation in terms of inference questions in 10- and 12-year-olds.

Other than Hildyard and Olson (1978) and also in contrast with Sticht and James (1984), Diergarten and Nieding (2016) found that 10-year-olds were already better at drawing emotional inferences from written text presentations and that 8-year-olds performed on a comparable level with written and auditory texts. Other than the above cited studies, the children and thus also the study materials in the study by Diergarten and Nieding (2016) were German rather than English. This revelation brings us to another important issue: How fast a beginning reader becomes a fluent decoder depends on the orthographical rules underlying the respective language (Seymour, Aro, & Erskine, 2003). Transparent orthographies, such as the German language, can be characterized by a high grapheme-phoneme correspondence. Children who learn to read a language with a transparent orthography achieve a higher level of decoding fluency earlier than beginning readers of less transparent orthographies such as English (Landerl, Wimmer, & Frith, 1997). Results of a cross-sectional study with children attending the grade levels 1 to 4 of German elementary school (ages 6-7 until 9-10) indicate that while listening comprehension continuously and moderately "increases" between the grade levels 1 to 4, there seems to be a particular large gain in reading comprehension skills between the end of the first school year and the middle of the second school year, that the study authors associate with an automatization of decoding abilities (Marx & Jungmann, 2000). Similarly, a meta-analysis by Florit and Cain (2011) concludes that the influence of linguistic comprehension on reading comprehension exceeds the influence of decoding at an earlier age in transparent than in less transparent orthographies. Con-

sequently, German children should understand auditory and written text equally well or even profit from the benefits of a written text presentation at an earlier age than English-speaking children.

As an attempt to organize the various and partly inconsistent findings, Kürschner and Schnotz (2008) proposed an *integrated model of reading and listening comprehension*. This model is particularly useful as it takes the earlier introduced theories of multiple mental representations (e.g., Kintsch, 1998; see section 1.1) as a basis. Further, it assumes a cognitive architecture that consists of multiple storage systems as originally assumed by Atkinson and Shiffrin (1968). Acoustic as well as visual verbal information is first processed into modality-specific representations that are only available for a very short period of time (e.g., Anderson, 1996). Information that the recipient perceives as relevant, is selected to be further processed within the working memory. Processing within the working memory is assumed to be the often-cited three-step process as assumed by for example Kintsch (1998). The sensory modality of verbal information is assumed to affect only the text surface representation as physical properties of written or spoken language are partly included in the text surface representation (e.g., Baddeley, 1997). By contrast, further semantic processing is in first place concept-driven and hence independent of the initial sensory mode. In accordance with the integrated model of text and picture integration (see p. 27), processing on a propositional text base level and situation model construction is based on the underlying sign system which is the same in auditory and written text. With this regard, Kürschner and Schnotz (2008) make the same assumption as the simple view of reading, namely that a common linguistic comprehension skill underlies both reading and listening comprehension. However, Kürschner and Schnotz (2008) acknowledge that lower-level text processing can be specific to the sensory modality and thus may have a bottleneck function because it determines the amount of information available for higher-level processing.

1.3.2.3 Summary and conclusion: cognitive processing of different forms of verbal information

Theoretical positions differ in their view of whether processing and comprehension differs depending on the mode verbal information is presented in (i.e. auditorily or in written form). Whereas the unitary process view assumes the same underlying cognitive processes and thus no difference between presentation modes, the dual process view assumes processing to be modality specific. Empirical support can be found for both theoretical positions, but overall, existing research does not consistently speak in favor of either the unitary or the dual process view. The fact that prior studies comparing auditory and written text comprehension vary greatly with regard to how they define and operationalize comprehension (as with studies that compare audiovisual

and auditory text) seriously restricts the possibilities to derive a general conclusion. Such an attempt has been made by Kürschner and Schnotz (2008) who assume effects of presentation mode only on lower levels of processing that may in turn affect the amount of information available for higher level processing. However, this model requires further empirical confirmation. Most importantly, there are strong indications that the role modality plays in comprehension depends on the recipient's age and the respective language's underlying orthographic rule set.

1.4 Summary of the Theoretical Background and Outlook on the Empirical Studies

This chapter provided an overview of the theoretical foundation of the empirical studies that will be described and discussed in the following chapter.

The first two sections elaborated the cognitive processes that underlie text comprehension.

Section 1.1 described the broadly accepted theoretical account of text comprehension as a process that involves three levels of mental representation: A mental representation of the text surface that contains the exact wording of a text, which is further processed into a mental representation of the text base that contains the semantic content in form of a propositional structure. Finally, the situation model goes beyond what is explicitly stated in the text and represents the whole situation described. Moreover, the situation model is considered to represent a situation in a functional analogy. Embodied theories of text comprehension go further and assume that a recipient perceptually simulates the situation described in a text. This process would result in an analogous, multidimensional, and modality-specific mental representation. Firstly, empirical studies support the assertion that both children and adults construct three different mental representations when processing and comprehending text. Secondly, a vast amount of studies indicate that adults perceptually simulate object shapes, colors, sounds, and perspectives, while children have been found to mentally simulate object shapes and a protagonist's spatial perspective.

The following section 1.2 elaborated a further pivotal assumption of text processing: Text comprehension is defined as the process of constructing a coherent mental representation of the situation described in the text. As a text cannot describe a situation exhaustively the recipient has to fill the gaps that are left inferentially to achieve coherence. Coherence refers to a local and a global level. Local coherence is achieved by connecting subsequent phrases with one another, while global coherence refers to the process of connecting more widely separated text passages and/or a current phrase with the overall text topic. Different theories that make different assumptions on whether

a recipient routinely generates inferences to achieve both local and global coherence have been discussed. Backed up with comprehensive empirical findings, I concluded that a constructionist position that assumes that a recipient attempts to establish both local and global coherence is most likely to apply. While children have a general idea of causality already early in life, the ability to infer coherence relations and to monitor causality in the course of situation model construction, develops considerably with age.

These two fundamental processes that underlie text comprehension can be assumed to be generally independent of whether a text is presented auditorily or in written form, or whether pictures that illustrate the content are added. Nevertheless, words and picture constitute different signs, and different sensory systems are involved when reading or listening to a text. Section 1.3 therefore elaborated if and to what extent these different underlying sign systems and sensory modes influence how the content is processed and finally understood. Words as symbolic signs and pictures as iconic signs are supposedly processed differently. Pictures consist of iconic signs, which physically resemble their referent; in contrast, words contain symbols that do not resemble their referent. Pictures thus represent a situation in an analogy-relation to its referent. This property is therefore supposed to potentially enhance the perceptual simulation and construction of a modality-specific situation model as assumed by theories of embodied cognition.

Furthermore, different cognitive subsystems for phonological and visual information (Baddeley, 1997), as well as distinct cognitive mechanisms for the processing of verbal and pictorial information are postulated (Paivio, 1986) and against this background it may be assumed that simultaneously presenting words and pictures may be beneficial because this way the limited capacity of the working memory is used efficiently. While a vast amount of studies indicate that children indeed benefit from an audiovisual text presentation (verbal information + pictures), the additional information provided by the picture can be redundant for adults, possibly resulting in negative effects of audiovisual text presentation. With regard to auditory and written text, two different positions on whether reading and listening comprehension are different cognitive operations have been contrasted. Empirical studies have yet produced only inconclusive results. However, there are strong indications that the recipient's age and language play a role.

This overview revealed a knowledge gap that will be addressed in the following empirical studies. The existing body of research that compared memory and/or comprehension between different modes of presentation lacks a solid theoretical foundation in terms of how comprehension is defined and operationalized. Only occasionally, researchers engaging in cross-media research have referred to theories of comprehension that are widely accepted and extensively studied by researchers interested in discourse

and text comprehension. These theoretical considerations are what the empirical studies in this thesis are built upon.

The Studies I and II are based on the idea that text processing and comprehension involves three levels of mental representation. Study I investigates whether the memory of the text surface, the text base, and the situation model differs between auditory, audiovisual, and written text presentation in samples of 8- and 10-year-olds and adults. Similarly, Study II compares memory of these three levels but only between auditory and audiovisual modes of presentation and with different age groups (7-, 9-, and 11-year-olds). Study III further elaborates the idea of text comprehension as the process of building a coherent mental representation and examines in a sample of 9- and 11-year-olds whether the ability to establish local and global coherence differs between auditory, audiovisual and written texts.

2 Empirical Studies

2.1 Study I - Multi-Level Mental Representations: Comparing Written, Auditory, and Audiovisual Text

2.1.1 Research question and assumptions

The aim of Study I[1] was to investigate whether text comprehension differs between written, auditory, and audiovisual (i.e. auditory text combined with pictures) texts. As well-established theories of text comprehension assume that this process involves at least three levels of mental representation, we accordingly compared the memory of the text surface, the text base, and the situation model separately between modes of presentation to assess text comprehension as a whole. Additionally, we included three age groups: 8-year-old second-graders who had just begun learning to read, 10-year-old fourth-graders who had supposedly mastered the transition from learning to read to reading to learn, and young adults who we expected to be experienced readers.

Literature that compares auditory, written, and audiovisual text presentations firstly indicates that the processing of verbal information, irrespective of modality, differs from that of pictorial information. Due to the structural properties that pictures as iconic signs share with reality (e.g., Chun & Plass, 1997) and the additional memory code pictures provide when added to text (Paivio, 1986), an audiovisual presentation, compared to auditory and written text presentations, is assumed to be beneficial for memory and comprehension (the "multimedia effect"; Mayer, 2005a). Research has revealed this effect at least for children up to the age of 12 (e.g., Carney & Levin, 2002). As the effectiveness of the multimedia principle depends on the recipients' prior knowledge, pictures were thus not found to generally promote text comprehension in adults (Kalyuga et al., 2003; Byrne & Curtis, 2000). Based on these considerations, for the 8- and 10-year-olds, we assume that text comprehension in terms of situation model construction is better when a text is presented audiovisually compared to auditory and written text presentations. For the adults, we expect performance in terms of situation model construction after an audiovisual text presentation to be lower compared to auditory and written texts.

[1] This section is based on the following publication: Wannagat, W., Waizenegger, G., & Nieding, G. (2017). Multi-level mental representations of written, auditory, and audiovisual text in children and adults. *Cognitive Processing*, *18*(4), 491-504. doi: 10.1007/s10339-017-0820-y.

In the case of different modalities of verbal information, namely auditory and written texts, results from the scant research available do not consistently speak in favor of either written or auditory text presentations. Consequently, possible differences between auditory and written texts will be addressed exploratorily for all age groups. As multi-level theories of text comprehension assume this process to encompass at least three levels of mental representations, this study also exploratorily examines whether effects of age group and presentation mode occur on the text surface and text base level of processing.

2.1.2 Methods

2.1.2.1 Participants

A total of 108 8-year-olds ($n = 36; M_{age} = 8;1$, $SD_{age} = 0;4$), 10-year-olds ($n = 36; M_{age} = 10;2$, $SD_{age} = 0;7$) and young adults ($n = 36; M_{age} = 21;9$, $SD_{age} = 2;11$) participated in the study. The 8- and 10-year-olds attended the second or the fourth grade of elementary schools in Bavaria, Germany. Data were collected in March and April at the beginning of the second half of the school year and children were tested in the morning during school hours. The adults were psychology undergraduates who received course credit for participating. All participants were either German native speakers or had mastered German at native-speaker level. While sampling the 8- and the 10-year-olds, we excluded those who had been diagnosed with attention deficit hyperactivity disorder and/or dyslexia, as reported by teachers or parents.

2.1.2.2 Memory for the text surface, the text base, and the situation model: The sentence recognition task

To measure memory of information on the levels of text surface, text base, and situation model, we applied a sentence recognition test, which is well established as a valid and informative measure of text comprehension (e.g., C. R. Fletcher & Chrysler, 1990; Narvaez, Radvansky, Lynchard, & Copeland, 2011; Nieding, 2006; Radvansky, Copeland, Berish, & Dijkstra, 2003; Schmalhofer & Glavanov, 1986). The basic idea behind this sentence recognition test is that subjects listen twice to a text; on the second occasion, each sentence is presented either as the original sentence or as one of three distractor sentences, involving surface changes, text base changes, or situation changes (see section 1.1.4). Subjects are asked to decide whether each sentence occurred in the text in exactly the same way or not ("yes" or "no"). The text surface is changed by replacing words with close synonyms or by changing the active voice to passive (or vice versa) and is therefore still within the scope of the text base. A text base change entails a valid inference from the subject matter described in that sentence, in line with the situation

model implicit in the story. A situation change is a sentence that contradicts the situation model. Responses to original sentences and surface changes allow for the determination of how well subjects can discriminate surface changes from original sentences. As surface changes only differ from original sentences in terms of their text surface, the ability to discriminate between these two sentence types indexes memory of the text surface. Discrimination of text base changes from surface changes indexes memory of the text base because surface changes preserve the text base while the inference constituting the text base change does not. Accordingly, the ability to discriminate situation changes from text base changes is a measure of having memory of the situation model.

2.1.2.3 Materials and procedure

The sentence recognition task. For the sentence recognition test applied in this study, we developed 12 narrative stories, each consisting of six sentences. The stories dealt with topics that children were likely to relate to, such as going to the zoo, having a birthday party, or going shopping. For every sentence, we created a version for each distractor type. Across all distractors, the mean surface change, measured by the number of characters changed, was greatest in the surface changed sentences, less in the text base changes, and smallest in the situation changes. This ensured that, for example, better recognition of situation changes as compared to text base changes was not caused by more text surface changes in the situation changes than in the text base changes.

Valid application of the sentence recognition task requires that the participants are familiar with the words that are used in the original and the distractor sentences. In order to test if the words used were at least part of the participants' passive vocabulary, we identified the words that we assumed could be unknown to children of this age. We then asked ten 6- and 7-year-old children for the right definition of these words with single-choice questions and replaced any words that were unknown to any of these children for the final version of the sentence recognition task. We further examined whether the mean frequency of the words involved in the changes in the distractor sentences differed between the original sentences, text surface changes, text bases changes, and situation model changes, assuming that word frequency is associated with familiarity (Brown & Watson, 1987). For this analysis, we referred to the childLex data base (Schroeder, Würzner, Heister, Geyken, & Kliegl, 2015). This data base is based on a corpus of German children's books, comprises of 10 million words, and offers linguistic parameters for three age groups (6- to 8-year-olds, 9- to 10-year-olds, and 11- to 12-year-olds). The mean frequency of the words involved in the changes made in the distractor sentences did not differ significantly for any age group (Fs < 1, ps > .05). This result indi-

cates that differences in familiarity with certain words as an explanation for differences in recognition performance for both the 8- and 10-year-olds are unlikely.

The stories were professionally recorded by a female speaker at the department's recording studio, equipped with a separate control room. We did not use dynamic-range compression to conserve the natural sound of the speaker's voice, but applied loudness adaption. The mean story duration was $M = 46.41$ seconds ($SD = 4.73$ seconds). The following is an example of the stories used (German version in parentheses).

1 Marie lies in bed wide awake, because it will be her seventh birthday tomorrow. (Schlaflos wälzt sich Marie in ihrem Bett, denn morgen wird sie sieben Jahre alt.)

2 She thinks about the many presents she will get and about the dear friends who will come over for cake and cocoa. (Sie denkt an die vielen Geschenke und die lieben Freunde, die zu Kakao und Kuchen kommen.)

3 At six o'clock, she can no longer stand lying in bed, and she gets up. (Um sechs hält sie es in ihrem Bett nicht mehr aus und sie steht auf.)

4 She is just about to take her favorite dress out of the wardrobe, when she hears whispering voices coming from the hallway. (Sie zieht gerade ihr Lieblingskleid aus dem Schrank, als sie auf dem Flur leise Stimmen hört.)

5 The door opens, and her parents enter, holding a very big cake. (Da öffnet sich ihre Zimmertür und ihre Eltern kommen mit einer sehr großen Torte herein.)

6 Marie blows out all candles at the first try, and wishes that every day would be her birthday. (Marie bläst alle Kerzen auf Anhieb aus und wünscht sich ganz fest, dass sie jeden Tag Geburtstag hat.)

Changing the text surface, the last sentence in the original became the following.

> She thinks about the numerous presents she will get and about the nice friends who will come over for cake and cocoa. (Sie denkt über die zahlreichen Geschenke und die netten Freunde nach, die zu Kakao und Kuchen kommen.)

A text base change transformed this sentence as follows.

> She thinks about the many presents she will get and about the dear guests who will come over for cake and cocoa. (Sie denkt an die vielen Geschenke und die lieben Gäste, die zu Kakao und Kuchen kommen.)

Finally, the situation change for this sentence was as follows.

> She thinks about the many games she will play and about the dear friends who will come over for cake and cocoa. (Sie denkt an die vielen Spiele und die lieben Freunde, die zu Kakao und Kuchen kommen.)

For the audiovisual presentation mode, we created a static, colored picture for each sentence to depict the scene described, taking care that the distractor sentence changes were not shown in the pictures. Figure 4 shows the pictures for each sentence of the example story.

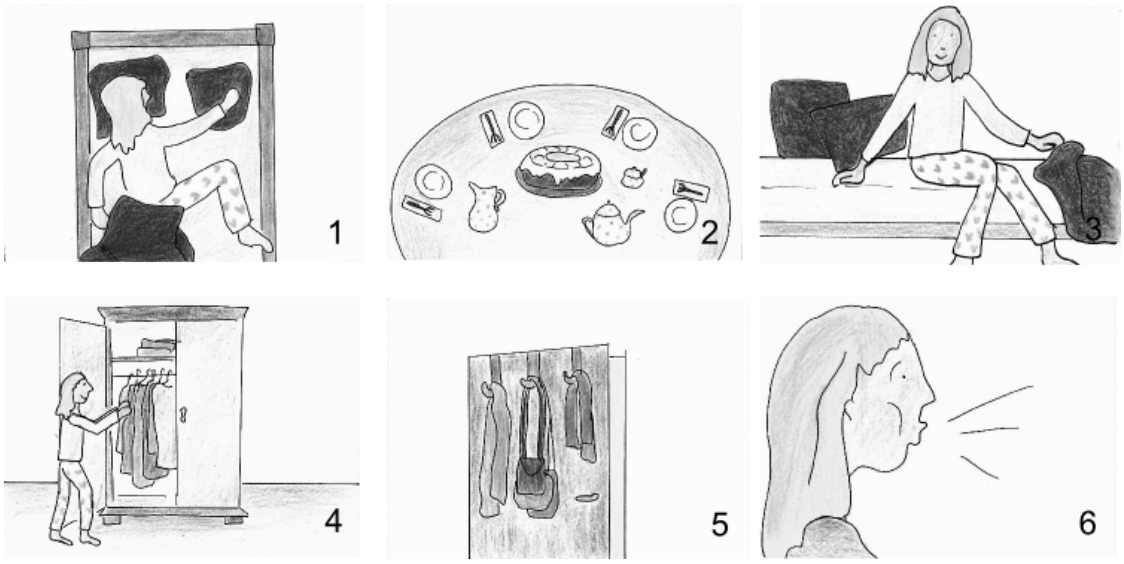


Figure 4. The pictures used in the example story. Numbers in the top left corner of each individual picture indicate the sentence to which the picture belongs. Note that colored versions of these pictures were used in the study.

In the written version, the stories were presented on a screen, appearing sentence after sentence. The letters were black, with a font size of 20, in Arial against a light grey background. In contrast to the auditory and the audiovisual versions, the stories were also presented in written form during the sentence recognition test. The distractor sentences stretched across the same number of lines as the original sentences did. To ensure ecological validity, subjects read the sentences at their own pace with the subsequent sentence appearing after a key press, enabling the subjects to determine their reading pace as it would occur in a non-experimental setting. To maintain the highest level of comparability with the auditory and the audiovisual conditions, it was not possible for the participants to go back to previous sentences that they had already moved on from. The presentation mode was varied within subjects. This means that every subject was presented with three blocks of four stories. One block consisted of four stories in the same mode of presentation. Across all participants of each age group, every story was presented in every mode of presentation as block one, two, and three. In addition, the position of the stories within these blocks was balanced across subjects during the first presentation and during the recognition task.

After every four stories (i.e. after every block), the participants took on the sentence recognition task, in which sentences were presented in a scrambled order within each story. Three of the six sentences in each story were presented as in the original, one involved a surface change, one a text base change, and one a situation change. In this way, 50 % of the sentences (original) required a yes response while 50 % (distractors) required a no response. Across all subjects and within each age group, each sentence was presented equally often in the original form or as each of the three distractor sentences. Additionally, during the recognition task, the position of each sentence within the scrambled order was balanced to ensure that every sentence was presented at every position (1 to 6) in every version (original, surface change, text base change, and situation change) across all subjects within each age group.

Implementing the balancing as described above required at least 18 participants per age group. A power analysis using G*Power (Faul, Erdfelder, Lang, & Buchner, 2007) further suggested 34 participants to detect medium sized effects in pairwise comparisons between presentation modes (assuming small correlations between repeated measures (r = .30), a power of 0.80 and $\alpha = .05$). Thus, we assessed 36 participants of each age group.

The experiment was programmed using PsychoPy (J. W. Peirce, 2007, 2008) and presented on a 17-inch screen. Before commencing the actual experiment, the experimenter issued instructions to ensure that every participant understood the task, which was followed by a practice story to further familiarize them with the sentence recognition task. During both the first presentation and the sentence recognition task, a two-tone tune indicated the beginning of a new story. After the tune, the screen went white for 2 seconds, to be followed by the story. In the auditory condition, a black dot was presented along with the stories. During the sentence recognition task, a cartoon drawing of a person quizzically scratching his head was shown along with the sentences. For the written condition, the cartoon drawing was not shown with the written sentences but was shown for one second before every sentence recognition block. Participants entered their answers by means of a specially prepared keyboard with just two accessible keys: one with a green check to answer "yes" and one with a red cross to answer "no". All participants were seated about 50 centimeters from the screen, with individual variation to allow each individual to observe the screen comfortably and to use the keyboard to enter their answers. The whole sentence recognition task took about 30 minutes to complete and each participant was tested individually. The 8- and 10-year-olds were tested at school during school hours and the adults were tested in the department's laboratory. For the 8- and 10-year-olds, a consent form signed by the parents was required for participation.

2.1.2.4 Data analysis

Our measure for memory of the different levels of mental representation is based on a person's ability to differentiate two stimuli, in our case two sentences, from one another. Therefore, A'-sensitivity values were calculated to index each level of mental representation (Narvaez et al., 2011; Pollack & Norman, 1964; Stanislaw & Todorov, 1999). A' is a non-parametric sensitivity index, which was preferred to the classical signal detection theory-based sensitivity index d' because our data violated the assumption of normal distribution necessary to calculate d'. As a sensitivity measure, A' indicates the ability to discriminate signal from noise, based on hits (correctly answering "yes" when a signal trial is presented) and false alarms (incorrectly answering "yes" when a noise trial is presented). To determine an A'-value as an index for the memory of text surface information, the percentage of yes responses to original sentences was taken as the hit rate, and the percentage of yes responses to surface changes was taken as the false alarm rate. To determine an A'-value as an index for the memory of text base information, the percentage of yes responses to surface changes was taken as the hit rate, and the percentage of yes responses to text base changes was taken as the false alarm rate. Finally, to determine an A'-value as an index for the memory of the situation model, the percentage of yes responses to text-base-changed sentences was taken as the hit rate, and the percentage of yes responses to situation changes was taken as the false alarm rate. A'-values can range from 0 to 1, with values greater than .5 indicating above-chance performance for the respective level of mental representation.

A comparison of A'-values between levels is not appropriate because the A'-values of the different levels are negatively linearly dependent; for example, to determine A' for the text base, the percentage of yes responses from text base changes is subtracted from the text surface changes, and to determine A' for the situation model, situation changes are subtracted from text base changes. For that reason, we examined the effects of age group and presentation mode separately for each level. To determine whether the A'-values differed between audiovisual, auditory, and written presentations, and between the age groups, we conducted a 3 (Presentation Mode: auditory vs. audiovisual vs. written) $\times$ 3 (Age Group: 8 vs. 10 vs. adults)-analysis of variance (ANOVA) with the presentation mode as a within-subject factor, separately for the text surface, the text base, and the situation model. The sphericity assumption was tested by application of Mauchly's sphericity test, and Greenhouse–Geisser-corrected degrees of freedom were considered where the test indicated a violation of this assumption. Furthermore, we square-root-transformed the A'-values to enhance the distributional properties before conducting the analyses (Field, 2013).

2.1.3 Results

Five 10-year-old participants dropped out because they did not finish the experiment. Therefore only 31 10-year-olds were included into the analyses. Similarity to the original sentences gradually decreases from text surface changes, to text base changes, to situation model changes, and recognition performance should therefore decrease accordingly. This could be confirmed for the 8-year-olds, $F(3, 105) = 114.59$, $p < .001$, $\eta_p^2 = .766$; $p < .001$ for all pair-wise comparisons, the 10-year-olds, $F(3, 90) = 214.84$, $p < .001$, $\eta_p^2 = .877$; $p < .001$ for all pair-wise comparisons, and the adults, $F(3, 105) = 240.81$, $p < .001$, $\eta_p^2 = .873$; $p < .001$ for all pair-wise comparisons, as can be seen from the descriptive data displayed in Table 2.

Table 2. *Mean percentages and standard deviations of yes responses to original sentences, surface changes, text base changes, and situation model changes for 8- and 10-year-olds and adults.*

	Original		Surface		Text Base		Situation Model	
Age	*M*(%)	*SD*	*M*(%)	*SD*	*M*(%)	*SD*	*M*(%)	*SD*
8	86.27	9.25	71.30	17.98	52.08	20.16	29.63	21.68
10	82.89	8.58	64.78	14.54	42.74	17.04	13.44	10.47
adults	85.19	10.09	61.34	16.92	36.81	18.41	9.26	9.51

Table 3 shows the mean A'-values for the text surface, the text base, and the situation model, separated by age group and mode of presentation. As A'-values greater than .5 indicate above-chance performance and therefore memory of the respective level of mental representation, as a first step, we tested whether the A'-values for every level and presentation mode (separately for 8-year-olds, 10-year-olds, and adults) were significantly greater than .5 using one-tailed one-sample t-tests. The results of these tests revealed that the A'-values of all subsamples were significantly greater than .5 (ts > 1, ps < .05), indicating that 8-year-olds, 10-year-olds, and adults had some memory of the text surface, the text base, and the situation model after auditory, audiovisual, and written text presentations.

Table 3. *Mean A'-values for the 8-year-olds, the 10-year-olds, and the adults in the auditory, audiovisual, and written presentation modes, and for all three levels. (CI stands for confidence interval).*

	Auditory		Audiovisual		Written	
	M (*SD*)	95 % CI	*M* (*SD*)	95 % CI	*M* (*SD*)	95 % CI
8-year-olds						
text surface	.61 (.22)	[.57, .69]	.62(.13)	[.58, .67]	.61 (.18)	[.55, .67]
text base	.59 (.28)	[.50, .69]	.69 (.18)	[.63, .75]	.64 (.21)	[.57, .71]
situation model	.71 (.20)	[.64, .78]	.74 (.16)	[.68, .79]	.58 (.25)	[.49, .66]
10-year-olds						
text surface	.70 (.14)	[.65, .75]	.58 (.22)	[.51, .66]	.64 (.20)	[.57, .71]
text base	.64 (.20)	[.58, .72]	.69 (.17)	[.63, .75]	.65 (.16)	[.60, .71]
situation model	.73 (.18)	[.66, .79]	.77 (.15)	[.72, .83]	.67 (.20)	[.61, .75]
Adults						
text surface	.65 (.23)	[.57, .73]	.65 (.21)	[.58, .72]	.71 (.20)	[.64, .78]
text base	.67 (.23)	[.60, .75]	.64 (.25)	[.56, .73]	.71 (.18)	[.64, .77]
situation model	.70 (.14)	[.64, .75]	.72 (.20)	[.65, .79]	.74 (.17)	[.69, .80]

2.1.3.1 Differences in memory for the text surface, the text base, and the situation model between presentation modes and age groups

To test whether memory of the text surface, the text base, and the situation model differed between auditory, audiovisual, and written text presentations, and between 8-year-olds, 10-year-olds, and adults, we performed three 3×3 ANOVAs (Presentation Mode $\times$ Age Group) with the square-root-transformed values of the A'-values for the respective levels of mental representation as dependent variables.

Regarding the text surface, neither the main effects for the factors Age Group ($p = .426$) and Presentation Mode ($p = .377$), nor the interaction between these factors ($p = .096$) reached statistical significance. The same was found for the text base ($p = .183$, $p = .506$, $p = .178$). For the situation model, the ANOVA revealed a significant main effect for the Presentation Mode factor, $F(2, 200) = 4.165$, $p = .017$, $\eta_p^2 = .040$, as well as a significant interaction between the Age Group and Presentation Mode factors, $F(2, 200) = 3.641$, $p = .007$, $\eta_p^2 = .068$, while the Age Group factor did not reach statistical significance ($p = .104$). Bonferroni-corrected post hoc tests revealed that the main effect of the Presentation Mode factor is due to a significantly ($p = .015$) greater mean A'-value in the audiovisual ($M = .72$, $SD = .17$) compared to the written condition ($M = .66$, $SD = .22$). Yet, as the interaction indicates, and as the subsequent simple effects analyses show, the differences between the presentation modes can only be found in the 8-year-olds, $F(2, 99) = 9.859$, $p < .001$, $\eta_p^2 = .166$, with significantly greater mean A'-values under the auditory ($p < .001$), as well as under the audiovisual

condition ($p = .002$) compared to the written condition. Simple effects analyses further revealed that only under the written condition, $F(2, 100) = 6.451$, $p = .002$, $\eta_p^2 = .114$, were the A'-values significantly smaller for the 8-year-olds compared to both the 10-year-olds ($p = .046$) and the adults ($p = .002$), for whom no significant difference could be found ($p = 1.00$). In the subsamples of the 10-year-olds, $F(2, 99) = 1.173$, $p = .314$, and the adults, $F(2, 99) = 0.559$, $p = .574$, no effect of presentation mode could be found. Figure 5 summarizes the effects for the situation model.

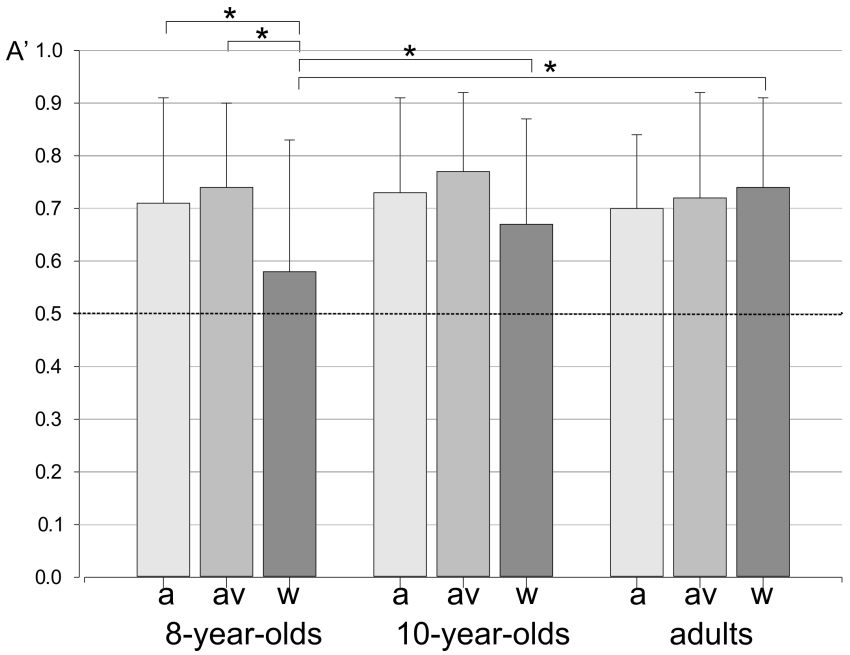


Figure 5. Mean A'-values for the situation model in the auditory (a), audiovisual (av), and written (w) presentation modes for 8- and 10-year-olds and adults. Error bars depict standard deviations. * $p < .05$.

2.1.3.2 Summary of the results

All in all, presentation mode and age group only seem to have an effect on the level of the situation model and not on the text surface or the text base level. Furthermore, memory of the situation model only differed between presentation modes in the 8-year-olds, who displayed less memory of situation model information after written presentations compared to that after both auditory and audiovisual presentations. Additionally, the age groups only differed under the written presentation mode with regards to their memory of the situation model: The adults and the 10-year-olds, who did not differ from each other, outperformed the 8-year-olds.

2.1.4 Discussion

The overall objective of this study was to examine if and how comprehension differs between written, auditory, and audiovisual text presentations in a sample including 8- and 10-year-olds as well as adults. To define and operationalize text comprehension, we refer to multi-level theories of discourse processing according to which the processing and comprehension of text involves at least three levels of mental representation. Therefore, we measured text comprehension with a sentence recognition test that allowed for the separate assessment of the memory of the text surface, text base, and situation model (e.g., Schmalhofer & Glavanov, 1986). We compared the memory of each level between modes of presentation and between age groups.

2.1.4.1 The situation model

As we found a main effect for the presentation mode in terms of memory of situation model information, our data indicate that text comprehension differs depending on whether a text had been presented as a written, auditory, or audiovisual version. We also found an interaction between the age group and presentation mode factors. In accordance with our assumptions, this interaction further indicates that the effect of the presentation mode displays itself differently depending on the recipient's age.

Only 8-year-olds performed differently depending on the presentation mode. As expected, 8-year-olds exhibited better memory of situation model information after an audiovisual text presentation compared to a written text presentation. This finding is in line with previous research and theoretical considerations and confirms that children profit from pictures added to verbal material in terms of text comprehension (Carney & Levin, 2002). We therefore conclude that, at least in contrast to written text, audiovisual text seems to promote the comprehension of narrative texts in 8-year-olds, who, in our case, attended the second grade of elementary school. Additionally, and in line with Sticht and James' (1984) assumption that listening comprehension skills exceed reading comprehension up to the seventh grade, 8-year-olds had better memory of situation model information following an auditory text presentation than following a written text presentation.

In fact, recognition performance was as good with the auditory presentation as it was with the audiovisual text presentation. This means that we cannot conclude from our data that pictures generally enhance the comprehension of verbal information, irrespective of whether verbal information is presented in auditory or written form. This finding may make sense when considering the important role that auditory information plays in children's everyday life. In most cases, children have become accustomed to narrative texts long before entering school and beginning to learn to read. Through

listening to stories that they are told or that are read to them, they supposedly have acquired quite good auditory text comprehension skills that enable them to construct a useful situation model from child-appropriate auditory texts (Artelt et al., 2007; Rosebrock & Nix, 2011). However, this finding that their comprehension of auditory texts seems to be as good as their comprehension of audiovisual texts is in contrast with previous studies that found that there was a beneficial effect from pictures on situation-model-related variables when compared to purely auditory texts (Diergarten & Nieding, 2015; Gibbons et al., 1986). Therefore, we consider it possible that the context in which the reception and processing of different text formats takes place should be taken into account. The current study compared modes of text presentation within subjects. To restrict the overall duration of the experiment to a reasonable time frame, only four stories were presented in each auditory, written, and audiovisual form. Gibbons et al. (1986) as well as Diergarten and Nieding (2015) varied the modes of presentation between participants. Perhaps four stories were not enough for the audiovisual text to unfold and achieve its full potential, especially when considering that the magnitude of the effects of the presentation mode found to date has been small to moderate (for example $\eta_p^2 = .11$, Diergarten & Nieding, 2015) [2].

For the adults and the 10-year-olds, we did not find a presentation mode effect. The memory of situation model information and therefore text comprehension did not differ between written, auditory, and audiovisual text presentations. Accordingly, our initial assumptions for 10-year-olds and adults are not supported by our data. We assumed the adults would show poorer memory of situation model information after audiovisual text presentations compared to verbal-only text presentations (i.e. auditory and written texts), as has been noted in the majority of previous studies (e.g., Byrne & Curtis, 2000). The superiority of verbal compared to audiovisual information is often explained via the effect of "expertise reversal" (Kalyuga et al., 2003), meaning that for someone who has high levels of prior knowledge concerning the topic, the information added by pictures interferes with meaningful processing.

To explain why our data did not indicate an expertise reversal effect, it is important to note that this effect was initially defined with reference to instructional settings and not to narrative text comprehension (Mayer, 1997). Likewise, the studies that found better performance following written presentations asked participants to recall facts from news stories (e.g., DeFleur et al., 1992) or advertisements (Furnham & Gunter, 1987). Therefore, the materials differed from the materials we used in this study. One key difference could be that it was obviously very easy for the adult participants to understand the stories used in this study, as more than 90 % of the sentences that were changed regarding the situation implied by the story were correctly recognized (see Table 2, p.

[2] Benchmark according to Ellis (2010)

48). In addition, text processing in general does not only depend on the semantic content but also on the so-called super structures (Richter & Christmann, 2002), or the text genre (see also Graesser et al., 1997), which initiate cognitive processing depending on the conventionalized structures of different types of text, for example narrative texts, didactic texts, or advertisements. Therefore, it seems plausible that narrative texts are processed differently than news stories or advertisements are in a way that the possible detrimental effects of pictures on text comprehension may not become apparent in adults.

The 10-year-olds' result pattern indicates that, comparable to the adults, their memory of situation model information is comparable across all modes of presentation. In contrast, we expected the 10-year-olds, as was seen with the 8-year-olds, to perform better when pictures were added to the verbal material, as most of the research we reviewed found beneficial effects of pictures on children up to the age of 12. Our findings also indicate that 10-year-olds, like adults, do not seem to differ in terms of their reading and listening comprehension. Although this contradicts Sticht and James' (1984) conclusion that listening comprehension exceeds reading comprehension up to the seventh grade, it is in line with newer research (Gunter, Furnham, & Griffiths, 2000; Walma van der Molen & van der Voort, 2000) that did not find differences in children aged 10 to 12, or that found better performance following written text presentations (Diergarten & Nieding 2016).

To explain their finding, Diergarten and Nieding (2016) refer to a meta-analysis (Florit & Cain, 2011) that concludes that the influence of decoding skills on reading comprehension decreases and the influence of listening comprehension skills increases earlier in transparent orthographies such as the German orthography. As our participants were German native speakers and our materials were thus also German, this explanation may apply to our finding as well. The specifics of the German language may account for our finding that 10-year-olds already deal with written text on a level comparable with adults. This is backed up by our further result indicating that differences between the three age groups were only found in terms of written text presentations. Eight-year-olds displayed poorer memory of situation model information after written text presentations than both the adults and the 10-year-olds did, who did not differ. These written text comprehension skills may eliminate the need for pictures to enhance the verbal materials for the 10-year-olds and may thus explain why, in contrast to our initial assumption, the memory of situation model information was not better after audiovisual text presentations when compared to verbal-only material.

2.1.4.2 Lower levels of text processing

We further exploratorily examined whether the lower levels of processing differed between modes of presentation and between age groups. Our results indicate that age and mode of presentation may only play a role in terms of memory of situation model information, and not in terms of the lower levels of processing. The memory of the text surface or semantic content did not differ between age groups or between written, auditory, and audiovisual text presentations.

2.1.4.3 Limitations and conclusion

One limitation of this study relates to the instructions for the sentence recognition task. We asked participants whether the sentence had occurred in the story in exactly the same way. Some researchers argue that this kind of instruction makes participants pay particular attention to both the text surface and the text base. Therefore, a text verification task has been suggested. This task requires participants to judge whether a sentence matches the state of affairs described in the text and would be more appropriate to measure the memory of the situation model information (Isberner et al., 2013; Maier & Richter, 2013). However, at the same time, this verification task causes the subjects to pay less attention to the text surface and the text base, and thus is not well-suited to measuring the memory of the text surface and text base. Furthermore, someone who must decide whether a certain sentence has occurred in a text also refers to their memory of the text base and the situation model (Reder, 1982). To capture all three levels, we thus decided to apply a sentence recognition task instead of a verification task.

In the group of the 10-year-olds five participants dropped out. We thus had three participants less than suggested by the power analysis. No significant effect of presentation mode was found for the 10-year-olds, $F(2, 99) = 1.173$, $p = .314$, but given the current p-value, we believe that it is highly unlikely that the effect of presentation mode would turn out to be significant by adding another three to five participants. In conclusion, this study offers insights into how different ways of presenting the same semantic content affect the cognitive processing of this information in children aged 8 and 10, and in young adults. Our data indicate that 8-year-old second-graders who have experienced one and a half year of formal reading instruction seem to benefit from audiovisual text presentations in terms of their memory of the situation model information. Consequently, audiovisual text presentation may aid in text comprehension when compared to written texts but not to auditory texts. In contrast, 10- year-olds, like adults, already seem to understand child-appropriate narrative stories to the same extent, regardless of the mode of presentation, although this finding may not hold across more complex orthographies such as the English orthography.

2.2 Study II - Multi-Level Mental Representations: Comparing Auditory, and Audiovisual Text

2.2.1 Research question and assumptions

This study[3] was conducted to account for one of the limitations of Study I. Based on the results of previous cross-media studies and theoretical considerations about the effects of pictures added to verbal materials (see section 1.3.1), we expected the 8- and 10-year-olds to benefit from an audiovisual presentation not only in comparison with a written but also in comparison with an auditory text presentation. In Study I, we compared three modes of text presentation which we manipulated within subjects. This resulted in only 4 stories of each presentation mode, which may have been a too small amount for the expected effect to become apparent.

Consequently, in Study II, we compared the memory of the text surface, the text base, and the situation model between only auditory and audiovisual presentation in samples of 7-, 9-, and 11-year-olds. Focusing on only two modes of presentation allowed presenting 8 stories in each mode of presentation. As in Study I, based on previous cross-media studies (e.g., Carney & Levin, 2002) and considering the idea of perceptually grounded information processing (e.g., Zwaan, 2014), we assumed that audiovisual presentation enhances situation model construction. As the development of the working memory subsystems (S. Palmer, 2000) and the perceptual support hypothesis (Springer, 2001) may indicate that the beneficial effect of pictures may be more pronounced in younger children, we also examined whether the beneficial effect of pictures added to auditory text is more pronounced in the 7-year-olds compared to the older participants, or whether the multimedia principle is effective to the same degree across all of the included age groups. Given the earlier sparse and conflicting results, memory of text surface- and text base-related information was exploratorily compared for audiovisual and auditory presentation modes.

2.2.2 Methods

2.2.2.1 Participants

A total of 106 students, aged 7-years ($n = 35; M_{age} = 7;7$, $SD_{age} = 0;71$), 9-years ($n = 36; M_{age} = 9;9$, $SD_{age} = 0;30$), and 11-years ($n = 35; M_{age} = 11;9$, $SD_{age} = 0;60$) participated in the study. The 7- and 9-year-old participants attended elementary school first and fourth grades, respectively, and the 11-year-olds attended the sixth grade of

[3]This section is based on the following publication: Wannagat, W., Waizenegger, G., & Nieding, G. (2018). Mental Representations of the Text Surface, the Text Base, and the Situation Model in Auditory and Audiovisual Texts in 7-, 9-, and 11-Year-Olds. *Discourse Processes*, *55*(3), 290–304. doi: 10.1080/0163853X.2016.1237246.

secondary schools, in Bavaria (Germany). Sex ratios in the subsamples were more or less balanced; 51.4 % of the 7- and 11-year-olds and 52.8 % of the 9-year-olds were male. All participants were either German native speakers or understood German at native-speaker-level. Students with ADHD or dyslexia were excluded before data collection. Seven- and nine-year-olds were tested at school during school hours. Eleven-year-olds were tested in the department laboratory in the afternoon because the secondary schools the participants attended did not allow testing during school hours. Only students whose parents had signed a consent form were allowed to participate.

2.2.2.2 Materials and procedure

To measure the memory of the text surface, the text base, and the situation model we employed a sentence recognition task (e.g., Schmalhofer & Glavanov, 1986) as described in Study I (see section 2.1, Section 2.1.2.3). For the purposes of this study, we developed 16 narrative stories (the 12 stories used in Study II plus four addional stories). Again, across all distractors, the mean surface change, measured by the number of characters changed, was greatest in the surface changed sentences, less in the text base changes, and smallest in the situation changes. This ensured that, for example, better recognition of situation changes as compared to text base changes was not caused by more text surface changes in the situation changes than in the text base changes.

With the additional four stories, we took the same precautions as with the materials used in Study I to ensure the validity of the sentence recognition task (see section 2.1.2.3). To examine if the words used in the materials are part of participants' at least passive vocabulary, we identified the words in the four additional stories, which we assumed could be unknown to children of this age. We then asked 10 six- and seven-year-old children for the right definition of these words with single-choice questions and replaced words unknown to any of these children for the final version of the sentence recognition task.

As in Study I, we examined, this time for the 16 stories used in this study, whether the mean frequency of the words involved in the changes in the distractor sentences differed between the sentence types. For this, we referred to the childLex data base (Schroeder et al., 2015). This data base is based on a corpus of German children's books and comprises 10 million words and offers linguistic parameters for three age groups (6- to 8-year-olds, 9- to 10-year-olds, and 11- to 12-year-olds). For neither age group, the mean frequency of the words involved in the changes made in the distractor sentences differs significantly ($F(3, 626) = 0.13$, $p = .944$, for the 7-year-olds; $F(3, 633) = 0.14$, $p = .934$ for the 9-year-olds; $F(3, 629) = 0.10$, $p = .958$ for the 11-year-olds). These findings indicate that the conclusion we made from the respective analysis in Study I, also applies when considering the materials of this study: Differ-

ences in familiarity with certain words as an explanation for differences in recognition performance are unlikely.

Again, the stories were professionally recorded by a female speaker at the department's recording studio. Mean story duration was $M = 47.08$ seconds ($SD = 4.93$ seconds).

Presentation mode was manipulated within subjects. First, eight stories were presented auditorily, and afterwards, eight stories were presented audiovisually (or vice versa). Across all subjects and within each age group, each story was presented in both auditory and audiovisual mode within the first and the second half of the sixteen stories.

After two stories (for 7-year-olds) or four stories (for 9- and 11-year-olds), the participants took on the sentence recognition task, in which sentences were presented in a scrambled order within each story. The balancing of the frequency and the positioning of the distractor sentences was conducted as in Study I: Three of the six sentences in each story were presented as per the original; one involved a surface change, one a text base change, and one a situation change. In this way, 50 % of the sentences (original) required a yes response while 50 % (distractors) required a no response. Across all subjects and within each age group, each sentence was presented equally often in the original form and as each of the three distractor sentences. Additionally, during the recognition task, the position of each sentence within the scrambled order was balanced to ensure that every sentence was presented at every position (1-6) in every version (original, surface change, text base change, and situation change) across all subjects within each age group. The position of the stories within the blocks of two or four stories after which the recognition test followed was also balanced across subjects within each age group.

For the programming of the experiment, presentation, and procedure we retained the proceedings of Study I. The experiment was programmed using PsychoPy (J. W. Peirce, 2007, 2008) and presented on a 17-inch screen. Before the actual experiment commenced, the experimenter issued instructions to ensure that every participant understood the task, followed by a practice story to further familiarize them with the sentence recognition task. During both the first presentation and the sentence recognition task, a two-tone tune indicated the beginning of a new story. After the tune, the screen went white for 2 seconds, followed by presentation of the story. In the auditory condition, a black dot was presented along with the stories. During the sentence recognition task, a cartoon drawing of a person quizzically scratching his head was shown along with the sentences. Participants entered their answers by means of a specially-prepared keyboard with just two accessible keys: one with a green check to answer "yes" and one with a red cross to answer "no". All participants were seated about 50 cm from the screen, with individual variation to allow each individual to observe the screen com-

fortably and to use the keyboard to enter their answers. The whole experiment took about 45 minutes to complete, and each participant was tested individually.

2.2.2.3 Data analysis

As in Study I (see section 2.1.2.4 for a detailed description), we calculated A'-sensitivity values to index each level of mental representation (Narvaez et al., 2011; Pollack & Norman, 1964; Stanislaw & Todorov, 1999). To determine whether A'-values differed between audiovisual and auditory presentation, and between 7- , 9-, and 11-year-olds on each level, we conducted three, 2 (Presentation Mode: auditory vs. audiovisual) $\times$ 3 (Age Group: 7 vs. 9 vs. 11)-ANOVAs with presentation mode as a within-subject factor. The sphericity assumption was tested by application of Mauchly's sphericity test, and Greenhouse-Geisser-corrected degrees of freedom were considered where the test indicated violation of this assumption.

2.2.3 Results

Before looking at the A'-values, the following table (Table 4) provides an overview of the percentages of yes-responses to the original sentences, the surface changes, the text base changes, and the situation model changes for each age group. As similarity to the original sentences gradually decreases from the original sentences to the situation model changes, recognition should decrease accordingly. This can be confirmed for the 7-year-olds, $F(3, 102) = 203.59$, $p < .001$, $\eta_p^2 = .86$; $p < .004$ for all pairwise comparisons, the 9-year-olds, $F(3, 105) = 290.70$, $p < .001$, $\eta_p^2 = .89$; $p < .001$ for all pairwise comparisons, and the 11-year-olds, $F(3, 102) = 293.07$, $p < .001$, $\eta_p^2 = .85$; $p < .001$ for all pairwise comparisons.

Table 4. *Mean percentages and standard deviations of yes-responses to original sentences, surface changes, text base changes and situation model changes for 7-, 9-, and 11-year-olds.*

	Original	Surface	Text Base	Situation Model
Age	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
7	90 % (.09)	78 % (.17)	69 % (.14)	30 %(.14)
9	86 % (.08)	63 % (.16)	44 % (.17)	13 % (.11)
11	80 % (.09)	63 % (.17)	41 % (.17)	11 % (.10)

Table 5 summarizes the mean A'-values for each age group, level of mental representation, and presentation mode.

Table 5. *Mean A'-values (overall and by age group) for auditory and audiovisual presentation modes and all three levels. (CI stands for confidence interval.)*

	Overall		7-year-olds		9-year-olds		11-year-olds	
	M (*SD*)	95 % CI	*M* (*SD*)	95 % CI	*M* (*SD*)	95 % CI	*M* (*SD*)	95 % CI
Auditory								
text surface	.60 (.22)	[.56, .64]	.53 (.23)	[.45, .60]	.64 (.23)	[.56, .72]	.64 (.16)	[.58, .70]
text base	.64 (.17)	[.61, .68]	.65 (.18)	[.59, .71]	.65 (.17)	[.60, .71]	.63 (.17)	[.57, .69]
situation model	.73 (.16)	[.70, .76]	.73 (.16)	[.67, .78]	.74 (.15)	[.69, .79]	.73 (.17)	[.67, .79]
Audiovisual								
text surface	.66 (.22)	[.62, .70]	.67 (.29)	[.47, .67]	.71 (.15)	[.65, .76]	.70 (.18)	[.64, .76]
text base	.58 (.21)	[.54, .62]	.52 (.24)	[.44, .60]	.65 (.17)	[.59, .71]	.58 (.20)	[.51, .64]
situation model	.78 (.14)	[.75, .80]	.81 (.11)	[.77, .85]	.76 (.14)	[.72, .81]	.77 (.17)	[.71, .83]

To test whether A'-values were significantly greater than .5 (indicating memory of all three levels of representation in every age group in both auditory and audiovisual presentation modes), we performed one-sample t-tests. Among the 9- and 11-year-olds, all A'-values differed significantly from .5 (ts > 1; ps < .05). For the 7-year-olds, A'-values did not differ significantly from .5 for text surface in auditory presentation mode, or for text base in audiovisual presentation mode (ts < 1). However, across all age groups, the mean A'-values proved to be significantly greater than .5 (ts > 1; ps < .05), indicating that all three levels of mental representation were constructed to some extent in both auditory and audiovisual presentation modes. However, 7-year-olds seem to have had difficulty in constructing mental representations of the text surface when a text was presented audiovisually, and with the text base when a text was presented audiovisually.

2.2.3.1 Effects of presentation mode and age group for each level

To examine whether the corresponding A'-values differ between age groups and between auditory and audiovisual presentation separate analyses were conducted for each of the three levels of mental representation. An overview of the significant effects is presented in Figure 6.

For the text surface level, a 2 × 3-ANOVA (Presentation Mode × Age Group) revealed significant main effects for the factors Presentation Mode, $F(1, 103) = 4.34$, $p = .040$, $\eta_p^2 = .04$, and Age Group, $F(2, 103) = 7.02$, $p = .001$, $\eta_p^2 = .12$. The interaction of these two factors did not reach statistical significance ($p = .936$). As shown in Figure 6, A, the mean A'-value for the text surface level was greater for audiovisual presentation mode than the respective value for auditory presentation mode. Additionally, Bonferroni-corrected post-hoc tests showed that, irrespective of the presentation mode,

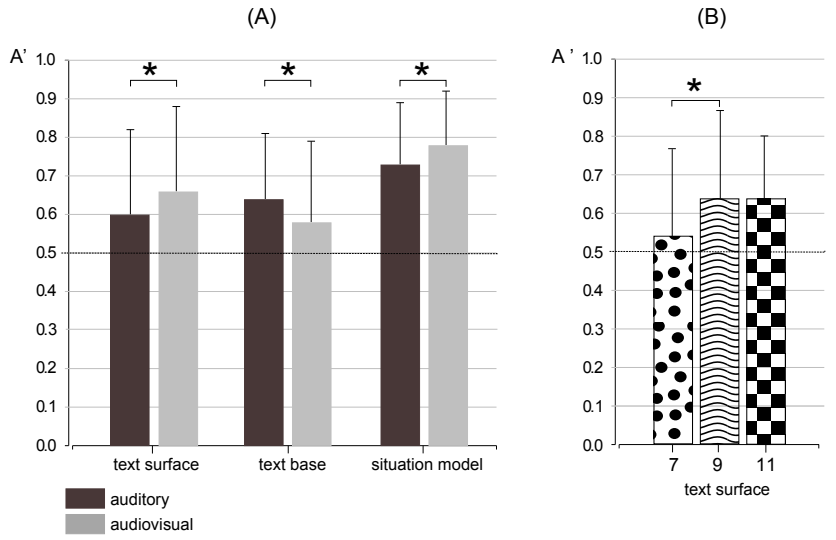


Figure 6. Mean A'-values for text surface, text base, and situation model in auditory and audiovisual presentation modes across all age groups (A), and for 7-year-olds, 9-year-olds, and 11-year-olds for the text surface (B). Error bars depict standard deviations. * $p < .05$.

the mean A'-value was lower in 7-year-olds than in both 9- ($p = .004$) and 11-year-olds ($p = .005$), who did not differ significantly ($p = 1.00$; see Figure 6, B).

For the text base level, we found a significant main effect for the factor Presentation Mode, $F(2, 103) = 6.03$, $p = .016$, $\eta_p^2 = .06$, but in contrast to the findings for text surface, the mean A'-value for text base was greater for the auditory than for the audiovisual presentation mode. Both the factor Age Group ($p = .140$) and the interaction of Age Group $\times$ Presentation Mode ($p = .131$) failed to reach statistical significance.

For the situation model, a 2 $\times$ 3-ANOVA (Presentation Mode $\times$ Age Group) revealed a significant main effect for the factor Presentation Mode, $F(2, 103) = 5.10$, $p = .026$, $\eta_p^2 = .05$. Here, as for text surface, the mean A'-value was greater in the audiovisual condition. Both the factor Age Group ($p = .689$) and the interaction of Age Group $\times$ Presentation Mode ($p = .530$) failed to reach statistical significance.

In sum, this result pattern indicates that mode of presentation affects all of the three levels of mental representation constructed during text comprehension. While memory of the text surface and situation model was better when the texts were presented audiovisually, memory of the text base, in contrast, was better when texts were presented in auditory mode. Age had an effect only on text surface level, where 7-year-olds exhibited poorer memory than 9- and 11-year-olds.

2.2.4 Discussion

The purpose of this study was to compare comprehension of auditory and audiovisual text presentation within a sample of 7-, 9-, and 11-year-olds. As in Study I, we relied on well-established, multi-level theories of discourse processing (Kintsch, 1998; van Dijk & Kintsch, 1983; Zwaan & Radvansky, 1998) to define and operationalize comprehension. For this reason, we applied a sentence recognition task, enabling separate assessment of memory of text surface, text base, and situation model (Schmalhofer & Glavanov, 1986).

2.2.4.1 The situation model

Recently, an embodied account of text comprehension has been proposed as an addition to Kintsch and van Dijk's original assumptions (Barsalou, 1999, 2008). This account assumes that situation model construction is a process of perceptual simulation, resulting in a situation model that is analogous to perception in resembling the representation of a real-life experience. Pictures, which constitute an integral part in any form of audiovisual text, represent information by means of signs that resemble their referent (Salomon, 1979; Schnotz, 2002). Consequently, they are more like a real-life experience than the exclusively verbal information of auditory text, which explains the potential of pictures to enhance perceptual simulation and situation model construction. On that basis, we assumed, as we did for Study I, that audiovisual presentation would result in better memory of the situation model than auditory presentation of the same information.

This time, with only the relevant presentation modes included and a larger set of materials, we found that children across all of the included age groups showed better memory of the situation model when texts were presented audiovisually. This finding indicates that children's situation model construction—and therefore understanding of a text—is enhanced by audiovisual text presentation.

Previous studies found that children performed better on tasks requiring inferences, and that children included more inferences when retelling a story following audiovisual presentation (Beagles-Roos & Gat, 1983; Gibbons et al., 1986; Pezdek et al., 1984; Ricci & Beal, 2002). Because inferences are an important component in the process of situation model construction, we interpret these findings to indicate the beneficial effects of audiovisual text presentation for comprehension. While our results align with previous studies, they also represent a valuable extension to the existing body of knowledge. In particular, use of the sentence recognition task allowed us to directly and explicitly examine the effect of presentation mode on situation model construction as widely acknowledged in the field of text comprehension.

Moreover, the pictures in our audiovisual texts did not show any element that was changed in the distractor sentences. Therefore, our findings indicate that the beneficial

effect of audiovisual presentation cannot be explained by the dual-coding assumption alone. Instead, audiovisual presentation seems to enhance perceptual simulation of the whole situation described in a story rather than only those elements actually represented in both verbal and pictorial form.

Furthermore, we examined whether the beneficial effect of pictures added to auditory text was more pronounced in the 7-year-olds compared to the older participants, in contrast to the multimedia principle being effective to the same degree across all age groups. Some research indicates that younger children prefer visual to phonological encoding until the age of eight (S. Palmer, 2000), and that perceptual details like shape, size, and color ("perceptual support") help children process information. Also, Unsöld and Nieding (2009) found that audiovisual text presentation improved predictive inferencing – a process that occurs on situation model level – only for 6-year-olds and not for children aged eight years and older.

Unlike Unsöld and Nieding (2009), we did not find an interaction of age and presentation mode. Instead, all age groups seemed to benefit equally from audiovisual presentation with respect to memory of the situation model. This may be explained by the different outcome measures used in both studies as Unsöld and Nieding (2009) examined predictive inferences, a specific aspect of situation model processing. According to the constructionist theory of inferences during text comprehension (Graesser et al., 1994), predictive inferences count as elaborative inferences and are thus not automatically generated, nor are they always necessary in understanding a text (see also Section 1.2.2). Unsöld and Nieding (2009) argued that predictive inferences made by 6-year-old children in audiovisual presentation mode may result from younger children's tendency to overrepresent information in the situation model (Fincher-Kiefer, 2001) while older participants construct a more parsimonious situation model. Such an information processing strategy does not necessarily contradict our finding that 7-, 9-, and 11-year-old children generally display better situation model memory in the audiovisual presentation mode, as against a highly specific process such as drawing predictive inferences, when a text is presented audiovisually. We assume that this finding reflects how the picture component of an audiovisual text triggers perceptual simulation, which seems effective irrespective of developmental status.

2.2.4.2 The text surface and the text base

Regardless of presentation mode, we found an age effect with respect to the text surface, which 7-year-olds were less able to remember than 9- and 11-year-olds, regardless of presentation mode. Our data further suggest that, when a text is presented auditorily, 7-year-olds may have difficulty in remembering any of the text surface whatsoever. This finding is compatible with what is known about the development of various memory

components, which substantially improve between the ages 6 and 12 (Schneider, 2015). At first sight, it might therefore seem contradictory that no age effect was found in terms of the higher levels of text processing. However, Reyna and Kiernan (1994) also found that 6-year-olds exhibited poorer surface memory than 9-year-olds, even though recall of the gist of stories used in their experiment did not differ between age groups. They explain this finding as a confirmation of fuzzy-trace theory (Brainerd & Reyna, 1992) that states that surface (verbatim) and gist memory are independent from one another, which offers the possibility that either one is more pronounced than the other one.

In addition to this age effect, we also found an effect of presentation mode for both the text surface and the text base. As in the results relating to the situation model, memory of text surface information was also better when the text was presented audiovisually. As compared to auditory presentation, audiovisual presentation enhanced not only text comprehension but also the ability to recall the exact wording of the text.

Unlike the findings for text surface and situation model, memory of the text base was better when the texts were presented auditorily rather than audiovisually. The results also indicate that 7-year-olds may only have weak memory of the text base at all following audiovisual presentation, as the relevant A'-value does not differ significantly from .5. All in all, the overall result pattern seems a bit odd at first glance and seems to contradict the prevailing view that audiovisual presentation has generally positive effects (e.g., Carney & Levin, 2002). Audiovisual presentation enhances memory of text surface and, as expected, situation model, but not of text base. As the outcome measures used in those cross-media studies discussed earlier (e.g., Beagles-Roos & Gat, 1983) did not assess the different levels of processing separately, it is difficult to integrate our findings with the existing research. However, one possible theoretical explanation for this result pattern can be derived from the assumptions of the integrated model of text and picture comprehension (see section 1.3.1.3, Schnotz & Bannert, 2003; Schnotz, 2014). This model proposes a theoretical framework describing the integrated cognitive processing of text and pictures. It is based on the assumption that texts and pictures are different kinds of external representation containing different sign types. Thus, two different paths for simultaneous but separate processing of verbal and pictorial information are assumed. Information from both paths is later integrated into a situation model that combines the information from both text and picture. While this model assumes the same three-step process of auditory and written text comprehension as Kintsch and van Dijk, pictures can be said to offer a more direct path to situation model construction. Cognitive processing of a picture begins with a (mental) visual perceptual representation of the picture, equivalent to the surface representation of a text. Based on this visual perception, a situation model is constructed by means of analog structure-mapping (Falkenhainer, Forbus, & Gentner, 1989; Gentner, 1989; Schnotz & Bannert, 2003). According to this

model, a propositional representation is not necessarily constructed as a prestage of the situation model during picture comprehension. Consequently, during audiovisual text processing, the verbal information is probably less important for situation model construction, because pictures are available as an alternative source of information and according to the IPTC, the text base is not needed as a "data base" for the situation model during the process of picture comprehension (Graesser et al., 1997; Schnotz & Baadte, 2015). This would explain our finding of poorer memory of text base information following audiovisual presentation. Assuming this analog structure-mapping mechanism, poorer text base memory following audiovisual text presentation does not indicate any detrimental effect of audiovisual texts with regard to the situation model, as a situation model is still constructed but via a different path.

Further, when the text is presented auditorily, and no support from pictures is given, continuous concentration for parsing and semantic analysis is required. This process results in propositions, which, once extracted, make memory for the text surface (the exact wording) redundant. This offers a possible explanation for the finding that text surface memory was lower at auditory text presentation. In addition, if according to the framework of the integrated model of text and picture comprehension, verbal information is less important at audiovisual text presentation, more working memory resources may be available. These can be invested in encoding text surface information, resulting in better memory of text surface information at audiovisual text presentation.

2.2.4.3 Limitations and conclusion

As the sentence recognition task applied in this study was essentially the same as in Study I, the same limitation in terms of the instruction can be discussed here as well: We asked participants whether the sentence had occurred in the story in exactly the same way. It has been argued that such a task makes participants pay special attention to the text surface and the text base. In contrast, a text verification task, which requires participants to judge whether a sentence matches the state of affairs described in the text, would be better suited to measure memory for the situation model (Isberner et al., 2013; Maier & Richter, 2013). However, at the same time, this verification task causes the subject to pay less attention to the text surface and the text base, and so is not well-suited to measuring memory of text surface and text base. Furthermore, someone who must decide whether a certain sentence has occurred in a text also refers to their memory of the text base and situation model (Reder, 1982). To capture all three levels, we therefore opted for the sentence recognition task rather than a verification task.

Another limitation concerns the difference in the recognition task between 7-year-olds and 9- and 11-year-olds. Seven-year-olds performed the recognition task every two stories and 9- and 11-year olds every four stories. We decided to proceed this way,

because we assumed that the same testing interval for all age groups would result in either floor effects (in case of the 7-year-olds with a recognition task every four stories) or ceiling effects (in case of the older participants with a recognition task every two stories). A'-values greater than .5, which indicate that two sentences can be discriminated at an above-chance level, are a prerequisite for any further meaningful comparisons of A'-values between conditions. The 7-year-olds had difficulties discriminating surface-changed sentences from original sentence in the auditory condition ($M_{A'} = .53$), and text base-changed sentences from surface-changed sentences in the audiovisual condition ($M_{A'} = .52$), although their recognition task was presumably easier. In our opinion, this finding confirms that we made the right decision in using a different testing interval in the different age groups.

It should also be noted that in combining auditory text with static pictures, we investigated the simplest possible form of audiovisual texts, and further research extending to animated pictures or even film would be of interest.

In conclusion, we found that mode of presentation affects all of the three levels of text comprehension in 7-, 9-, and 11-year-olds. While audiovisual presentation supported memory of the text surface and thus the exact wording of a text, auditory presentation proved better for text base memory. Most importantly, memory of situation model information was better in audiovisual than in auditory presentation mode. The findings confirm that, until the age of at least 11, children's text comprehension benefits more from audiovisual than from auditory texts.

2.3 Study III - Local and Global Coherence: Comparing Auditory, Audiovisual, and Written Texts

2.3.1 Research question and assumptions

Although constructing an especially globally coherent mental representation is essential for text comprehension and can be challenging for younger children, previous cross-media studies have rarely included measures of the inferences necessary to construct local and/or global coherence (see sections 1.2.4 and 1.3.1). The aim of the current study is to contribute to closing this gap and compare the construction of local and global coherence between auditory, written, and audiovisual presentation in a sample of elementary and secondary school children. For this purpose, we conducted a reaction-time-based word-recognition task by adapting a procedure that has previously been employed by Rizzella and O'Brien (1996, see section 1.2.3) and Unsöld (2008). Participants listened to or read short narrative texts in which the protagonists pursue goals that they either reach via a subordinate goal or do not because an obstacle occurs. This text structure was chosen because various findings indicate that goals help organize a narrative and

provide coherence (see section 1.2.4).
In a word recognition task, the reaction times to target words that are associated with the protagonist's superordinate (global target words) or subordinate goal or obstacle, respectively (local target words), serve as indicators of the inferences drawn to construct coherence on a global and a local level. These reaction times are compared between auditory and audiovisual presentation and between auditory and written text presentation. The evidence from these reaction times is supplemented by additional questions about text details that were either explicitly or implicitly mentioned in the texts.

Other than in Study I (see section 2.1), we leave the comparison of written and audiovisual text aside. In contrast to a comparison of written and audiovisual text presentation, in a comparison of auditory and audiovisual text, the only difference between these two is that at audiovisual presentation pictures as a further sign are introduced while the sensory modality of the verbal information remains the same. We address the effect of the sensory channels by which verbal information is perceived by comparing auditory and written text.

We investigate two age groups: 9-year-olds who attended the fourth grade of elementary school and 11-year-olds who attended the sixth grade of secondary school in Germany. These two age groups capture two important steps in the development of reading competence. Fourth-graders are transitioning from a phase in which they engage with text to learn how to read to a phase where they read for the purpose of knowledge acquisition. Sixth-graders are assumed to have successfully managed this transition (Bäuerlein, 2014). In line with the multimedia principle and the numerous studies (including our own, see section 2.2) that found that children benefit from pictures in terms of memory and comprehension of text, we assume that constructing both local and global coherence is easier in texts presented audiovisually than in auditory texts. We accordingly expect reaction times to local and global target words to be faster and the number of correctly answered comprehension questions to be higher in the audiovisual than in the auditory presentation mode across both age groups.

Given the inconsistent findings of previous research that compared written with auditory text (see section 1.3.2), we do not make any assumptions on whether children in both age groups are better at constructing a locally and globally coherent mental representation when they read or when they listen to a text. Therefore, we exploratorily compare between auditory and written text presentation the reaction times to local and global target words and the number of correctly answered comprehension questions.

2.3.2 Methods

2.3.2.1 Participants

A total of 155 9-year-olds ($n = 88, M_{age} = 9;10, SD_{age} = 0;4$) and 11-year-olds ($n = 67, M_{age} = 11;8, SD_{age} = 0;5$) participated in the study. The 9-year-olds attended the fourth grade of elementary school and the 11-year-olds attended the sixth grade of secondary school in Germany. We collected data in November and December, therefore in the middle of the first term of the respective school years. Table 6 gives an overview of the distribution of participants over the different conditions (presentation modes). The mean age does not differ between participants with regard to the written, auditory, and audiovisual conditions in either age group (p = .306 for the 9-year-olds; p = .126 for the 11-year-olds). All participants were either native German speakers or had mastered German at a native-speaker level. We did not include students who had been diagnosed with attention deficit and hyperactivity disorder and/or dyslexia as reported by parents or teachers.

Table 6. *Number of participants and mean age (standard deviations) for the written, auditory, audiovisual, and overall (sub)sample for both 9- and 11-year-olds.*

	Written		Auditory		Audiovisual		Overall	
	N	$M_{age}(SD)$	N	$M_{age}(SD)$	N	$M_{age}(SD)$	N	$M_{age}(SD)$
9-year-olds	28	9;10 (0;4)	28	9;10 (0;4)	32	9;8 (0;4)	88	9;10 (0;4)
11-year-olds	23	11;7 (0;4)	20	11;10 (0;6)	24	11;7 (0;4)	67	11;8 (0;5)

2.3.2.2 Word-recognition task to measure local and global coherence

For the purpose of this study, we developed 40 narrative stories. These stories all have the same underlying structure. They differ only in terms of the topics they deal with (for example, getting a new pet, a birthday party, or going on a vacation). A protagonist pursues a goal that he either reaches via a subordinate goal in half of the stories (for an example, see Table 7, left column), or that he does not reach in the other half of the stories, because an obstacle occurs (see Table 7, right column). The first sentence introduces the general setting of the story. In the second sentence, a superordinate goal is introduced, which is further elaborated in the third sentence. In the example story about Christoph, it is his goal to win the trophy. Then, in the following two sentences (sentences 4 and 5), the story goes on without further direct reference to the superordinate goal. Sentences 6 and 7 deal with a subordinate goal, which is for Christoph to score a goal. In the other example story, an obstacle occurs (rain). Scoring a goal is a subordinate goal because Christoph's team has to score a goal to win the trophy. In

the eighth and final sentence, an emotion of the protagonist is described. The reader or listener can attribute this emotion to the protagonist either reaching the superordinate goal or the subordinate goal. Christoph can either be happy because he has reached the superordinate goal and won the trophy, or he can be happy just because he has scored a goal. The negative emotion in the example story about Lena and Finn can be accordingly attributed to the protagonists not reaching the superordinate goal (collect mushrooms) or to the obstacle (rain). The words in italics (trophy /goal; mushrooms/rain) as global and local target words are associated with the superordinate or global goal and the subordinate or local goal (or with the obstacle, respectively). We conducted a pilot study to validate the association of the global and local target words with the super- and subordinate goals and the obstacles. Ten students read the stories and were asked to name two reasons for the protagonist experiencing the emotion described in the last sentence (in the example story: "Why does Christoph smile happily?" Please name two reasons).

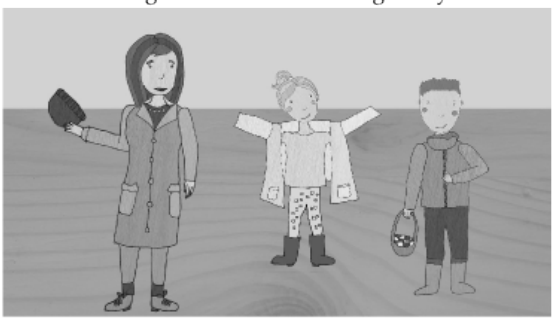
Depending on the answers we revised the target words in some cases. During the experiment either the global or the local target word was presented after each of the 20 stories. It was the participants' task to answer as fast and as accurately as possible whether or not this word had occurred in the respective story. The mean reaction times to the local and global target words served as indicators of the extent to which the participants have drawn inferences to construct and maintain local and global coherence. As explained in the introduction, the inferences to obtain a coherent mental representation of the situation described in a text are predominantly causal inferences (e.g., Bohn-Gettler et al., 2011). Thus, someone who has been presented with the example story would automatically integrate the last sentence (Christoph smiles happily) into the situation model constructed thus far by trying to answer the question regarding why Christoph is happy. When coherence is, for example, obtained on a global level, the answer to this question would be that Christoph is happy because his team has won a trophy. This information, which is active in memory, thus is reflected in a fast reaction to the global target word. To create the auditory and audiovisual versions, we professionally recorded the stories with a female speaker at the department's recording studio. The mean story duration was $M = 38.88$ seconds ($SD = 5.92$ seconds).

For the audiovisual versions, we additionally created five colored pictures that illustrate the content. Figure 7 shows the pictures that illustrate the example story. Not every sentence was illustrated with its own picture, but the first and the last sentence of every story and at least every second sentence were illustrated. The target words were not shown in any of the pictures. When no picture was presented along with the sentence, a blank (white) screen was shown instead. In the written version, the stories were presented on a white screen with black letters in Arial font. Participants silently

read one sentence at a time, skipping to the next sentence by pressing the space key that was marked with an arrow pointing to the right. It was not possible to go back to previous sentences. This enabled the participants to read at their own pace, as would occur in a non-experimental setting.

1. Lena and her older brother Finn are happy because their grandmother is visiting today.

2. At lunch time, Lena, Finn, and their mother decide that they will have mushrooms for dinner:



4. So, they put on their jackets and cheerfully start walking.

6. They quickly hide under a big fir tree, and heavy rain starts pouring down that just would not stop.

8. Back home, they all are in a very bad mood.

Figure 7. The pictures that illustrate the example story in which the protagonists do not reach their goal (right column of Table 7) with the corresponding sentences written underneath. Colored versions of these pictures were used in the study.

Table 7. *The story on the left is an example of a story in which the protagonist reaches his goal. The story on the right is an example of a story in which the protagonists do not reach their goal.*

	Goal reached	Goal not reached
1	Today is the last game of the season, and Christoph tightly laces his soccer shoes. (Heute ist das letzte Spiel der Saison und Christoph schnürt seine Fußballschuhe fest zu.)	Lena and her older brother Finn are happy because their grandmother is visiting today. (Lena und ihr großer Bruder Finn freuen sich, denn die Oma kommt heute zu Besuch.)
2	He wants to perform at his best, because the winning team will take home a nice trophy. (Er will sein Bestes geben, denn in diesem Jahr darf die Siegermannschaft sogar einen tollen Pokal mit nach Hause nehmen.)	At lunch time, Lena, Finn, and their mother decide that they will cook mushrooms for dinner. (Mittags überlegen sich Lena, Finn und die Mutter, dass es zum Abendessen Pilze geben könnte.)
3	During the game all players eagerly fight for the ball, because they all want to get the trophy. (Während des Spiels kämpfen alle Spieler verbissen um den Ball, denn alle wollen unbedingt den Pokal holen.)	Lena and Finn suggest collecting the mushrooms in the woods behind their house. (Lena und Finn schlagen vor, die Pilze im Wald hinter ihrem Haus zu sammeln.)
4	Soon before the game is over, both teams are tied 1:1. (Kurz vor Spielende sind beide Mannschaften gleich auf und es steht 1:1.)	So, the three of them put on their jackets and cheerfully start walking. (So ziehen die drei sich ihre Jacken an und laufen frohen Mutes los.)
5	Christoph gets the ball and runs across the field. (Christoph bekommt den Ball und rennt über das Spielfeld).	For a long time, they walk through rustling leaves, when suddenly dark clouds appear in the sky. (Lange laufen sie durch raschelndes Laub, als auf einmal dunkle Wolken aufziehen.)
6	Directly in front of the goal, he lunges and kicks the ball. (Kurz vor dem Tor holt er kräftig aus und schießt.)	They quickly hide under a big fir tree, and heavy rain starts pouring down that just would not stop. (Schnell stellen sie sich unter eine große Tanne und prompt setzt starker Regen ein, der einfach nicht aufhören will.)
7	The ball hits the goal, and after that the referee immediately blows the final whistle. (Der Ball geht ins Tor und sofort danach pfeift der Schiedsrichter das Spiel ab.)	So, the mother takes Lena's hand and the three of them start running home through the rain as fast as they can. (So nimmt die Mutter Lena an die Hand und die drei rennen so schnell sie können durch den Regen zurück nach Hause.)
8	Christoph smiles happily. (Christoph strahlt über das ganze Gesicht.)	Back home, they are in a very bad mood. (Im Haus angekommen sind alle drei ziemlich schlecht gelaunt.)

Procedure. The experiment was programmed using PsychoPy (J. W. Peirce, 2007, 2008). Figure 8 illustrates the procedure. The 40 stories we included in the experiment resulted in 40 trials. Each trial started with a blue dot appearing on the screen for 1,000 ms followed by a period of 750 ms during which the screen was blank (white). After that, the stories were presented in one of the three presentation modes. Then, a black cross was shown for 1,000 ms, again followed by a period of 750 ms with a blank screen. Then, a word was presented either auditorily, as in auditory and audiovisual text presentation, or appeared on the screen. The participants, who had been instructed to answer as fast and as accurately as possible, then had to decide whether this word had occurred in the story or not. They entered their answers by means of a specially-prepared keyboard with two marked keys: one with a green check to answer "yes" and one with a red cross to answer "no". Whether the "yes" key was on the left or the right was balanced across participants for each presentation mode in both age groups.

Twenty stories were experimental stories. In ten of these stories, the query word was the global target word; in the other ten, it was the local target word. The stories in which the protagonist reaches his or her superordinate goal and the ones in which he or she does not were equally distributed over the stories, after which a global or a local target word was queried. Across all participants within each presentation mode, the global and the local target words were presented equally often for each story. The other 20 stories were filler stories. The query words in four of these stories were words that had appeared in the middle of the story. Sixteen query words did not appear in the respective stories at all, therefore requiring a "no" answer. This resulted in a 60:40 ratio of required yes and no answers. Half of the filler stories consisted of only seven instead of eight sentences. The order of the 40 trials was randomly determined by the computer for each individual participant.

2.3.2.3 Additional questions of text comprehension

We included an additional measure of text comprehension within the above-described reaction time experiment. After 22 of the stories, we asked a question that addressed the content of the respective story but that did not address the super- and subordinate goals or obstacles. In the example story about Christoph described above, this question is *Which team wins?* (*Welche Mannschaft gewinnt?*). To answer correctly, the participants had to either draw an inference, as in this example question, or remember a story detail. The questions were recorded by the same speaker as in the stories and played automatically after the respective stories. The participants answered orally to the experimenter, who noted the answers on a separate protocol sheet. We did not ask questions after all of the stories to reduce the overall duration of the experiment. This measure served two purposes. Firstly, we obtained a second measure of text comprehension to validate

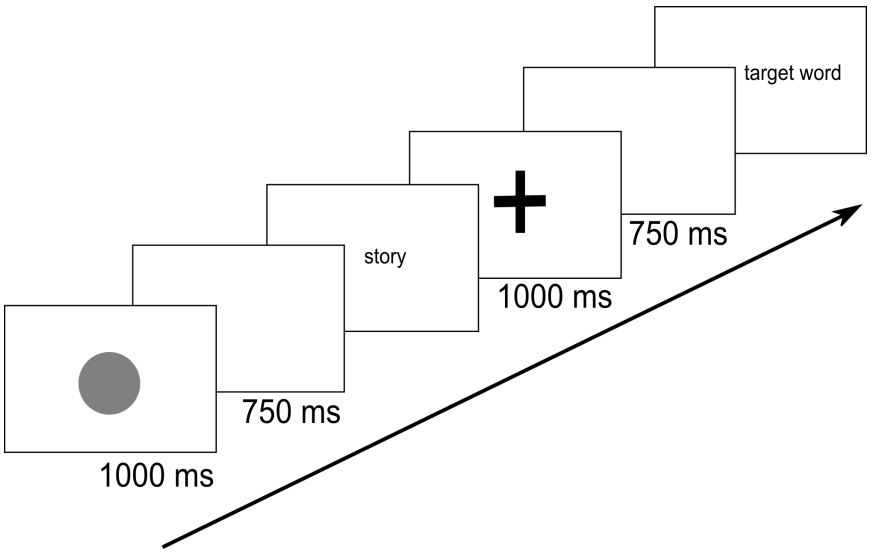


Figure 8. Sequence of events in an example trial. The dot was blue.

and supplement the reaction times to global and local target words. Secondly, we motivated the participants to pay attention during the whole story and not only during the sentences in which the target words appeared.

2.3.2.4 Overall procedure

Participants were randomly assigned to the written, the auditory, or the audiovisual presentation mode, which was thus manipulated between subjects. The reaction time experiment to measure global and local coherence, including the additional questions for text comprehension, took approximately 45 minutes. Every participant was tested individually. Before the actual experiment started, participants read or listened to (depending on whether they were assigned to the written, auditory, or audiovisual condition) two practice stories—one with a query word that had occurred in the story and another one in which the query word had not occurred. In these practice stories, participants received automated feedback about the correctness of their answers.

2.3.2.5 Data analysis

The main goal of the current study was to compare—between different modes of text presentation—the ability to draw inferences to obtain global and local coherence. A faster reaction to the global target words with audiovisual text presentation than with auditory text presentation thus means that inferences to obtain global coherence are drawn to a greater extent with the audiovisual text presentation than with the auditory text presentation. The reaction times to the local target words are interpreted accord-

ingly. To examine whether the reaction times to the global and the local words of both age groups differ between the auditory and audiovisual presentation and between the auditory and written text presentation, we conducted a 3 (Presentation Mode) × 2 (Age Group) × 2 (Target Word)-ANOVA.

First, we excluded reaction times when the answer was inaccurate in the word recognition task. In addition, one 11-year-old participant was completely excluded because of more than 25 % inaccurate responses. Secondly, to identify and eliminate outliers, we first removed reaction times greater than 5,000 ms and then eliminated response times that lay three standard deviations above or below the mean of each age group and each presentation mode. Ultimately, 96.1 % of all reactions of the 9-year-olds and 96.4 % of all reactions of the 11-year-olds were included in the analyses.

Finally, to calculate the scores of text comprehension, as measured by the additional questions, one point was awarded for each correct answer, thus resulting in a maximum of 22 points that could be obtained. Answers that were neither entirely correct nor entirely incorrect were awarded half a point. We expect this score to be negatively correlated with the reaction times to both the local and the global target words, which, as indicators of the extent to which local and global coherence is obtained, also capture an aspect of text comprehension. We conducted a 3 (Presentation Mode) × 2 (Age Group)-ANOVA to test whether text comprehension differs between presentation modes and age groups.

2.3.3 Results

An overview of the mean reaction times to both types of target words in all age groups and presentation modes can be found in the first two columns of Table 8.

To examine whether the reaction times to the global and the local words of both age groups differed between the different modes of presentation, we conducted a 3 (Presentation Mode) × 2 (Age Group) × 2 (Target Word)-ANOVA. This analysis revealed a significant main effect of the factor Age Group, $F(1, 149) = 28.699$, $p < .001$, $\eta_p^2 = .162$. The 11-year-old participants, irrespective of target word and mode of presentation, reacted significantly faster than the 9-year-olds. This finding replicates the well-examined effect that general processing speed substantially increases during the course of childhood (Kail, 2000).

Furthermore, we found significant main effects of the factors Target Word, $F(1, 149) = 76.872$, $p < .001$, $\eta_p^2 = .340$, and Presentation Mode, $F(1, 149) = 8.692$, $p < .001$, $\eta_p^2 = .104$, and a significant interaction of these two factors, $F(2, 149) = 3.704$, $p = .027$, $\eta_p^2 = .047$. Subsequent simple effect analyses showed that the effect of the factor Target Word can be found across all three presentation modes (all *p*s < .001). Reaction times to local target words were generally faster than to global target words.

Table 8. *Mean (M) reaction times (and standard deviations, SD) to global target words, local target words, and mean scores obtained in the additional measure of text comprehension (Comprehension) with written, auditory, and audiovisual presentation.*

	RT_{global}(ms) *M* (*SD*)	RT_{local}(ms) *M* (*SD*)	Comprehension *M* (*SD*)
Overall			
Written	1436 (298)	1319 (273)	17.6 (2.6)
Auditory	1687 (331)	1494 (298)	16.3 (2.8)
Audiovisual	1484 (243)	1386 (204)	17.7 (2.2)
9-year-olds			
Written	1499 (330)	1393 (315)	17.4 (2.6)
Auditory	1807 (341)	1630 (289)	16.4 (3.0)
Audiovisual	1557 (264)	1450 (233)	17.0 (2.3)
11-year-olds			
Written	1358 (237)	1229 (179)	17.7 (2.6)
Auditory	1519 (234)	1304 (188)	16.1 (2.6)
Audiovisual	1386 (172)	1301 (164)	18.7 (1.7)

However, the effect of presentation mode displayed itself differently depending on the type of target word. Participants across both age groups reacted faster to global target words in the audiovisual ($p = .002$, $d = .714$) and the written ($p < .001$, $d = .806$) mode in comparison to the auditory mode of presentation, $F(2, 149) = 9.975$, $p < .001$, $\eta_p^2 = .118$. In terms of the local target words, $F(2, 149) = 5.316$, $p = .006$, $\eta_p^2 = .067$, participants reacted faster in the written ($p = .004$, $d = .619$) compared to the auditory presentation mode. In contrast, the comparison of the audiovisual with the auditory presentation mode did not become significant ($p = .160$). We applied Bonferroni-corrections for multiple tests for all post-hoc comparisons. The interactions of the factors Target Word and Age Group, $F(1, 149) = 0.186$, $p = .667$, Presentation Mode and Age Group, $F(2, 149) = 1.638$, $p = .198$, and the three-way interaction Target Word $\times$ Presentation Mode $\times$ Age Group, $F(2, 149) = 0.347$, $p = .708$, did not reach statistical significance. Figure 9 displays the significant effects of presentation mode.

The mean scores obtained by the 9- and 11-year-olds, and both age groups taken together, when answering the additional questions for text comprehension are displayed in the third column of Table 8. We calculated a 3 (Presentation Mode) $\times$ 2 (Age Group)-ANOVA to examine whether there are differences between the written, the auditory, and the audiovisual mode of text presentation and between the 9- and 11-year-olds. This analysis revealed a significant main effect only for the factor Presentation Mode, $F(2, 147) = 5.462$, $p = .005$, $\eta_p^2 = .069$. Subsequent Bonferroni-corrected post-hoc tests showed the same result pattern as those for the reaction times (see Figure 10): Participants across both age groups scored higher with both the written ($p = .034$, $d = .487$)

and the audiovisual ($p = .015$, $d = .567$) mode in comparison to the auditory mode of text presentation.

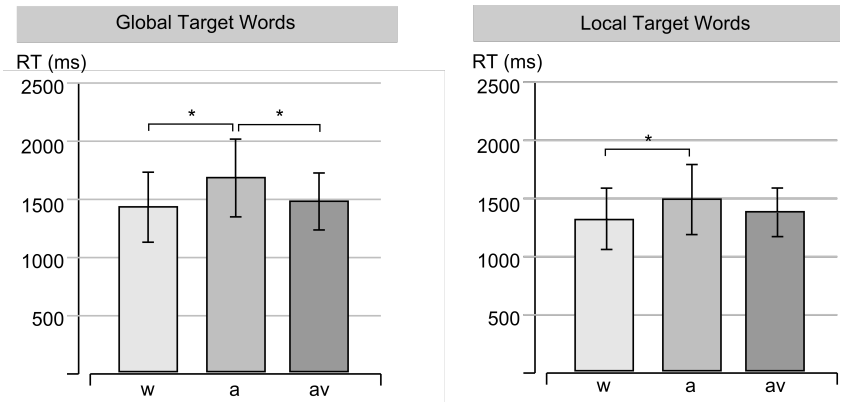


Figure 9. The mean reaction time (RT) in milliseconds (ms) to global and local target words with written (w), auditory (a), and audiovisual (av) text presentation. Error bars depict standard deviations. * $p < .05$.

Furthermore, the score obtained in this measure (not differentiated by age group, because there is no significant main effect of the factor Age Group) was negatively correlated with reaction times to both global ($r = -.198$, $p = .014$) and local target words ($r = -.204$, $p = .011$). These correlations indicate that the faster the participants reacted to the target words associated with the superordinate and the subordinate goals, the better they scored in terms of the text comprehension measure.

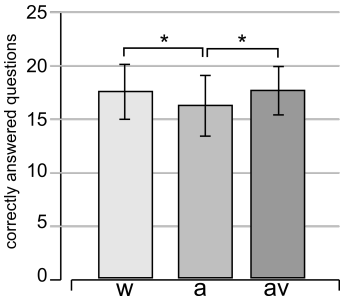


Figure 10. The mean scores obtained in the additional questions for text comprehension across both age groups. Error bars depict standard deviations. * p < .05.

2.3.3.1 Summary of the results

The reactions to local target words were faster than to the global target words. Additionally, in line with our assumption that children would benefit from an audiovisual text presentation, we found faster reactions to global target words at the audiovisual compared to the auditory mode of presentation. Also, the reaction times to global target words were faster at the written than at the auditory mode of text presentation. The same was found in terms of the additional questions for text comprehension.

In terms of the local target words, we found that reactions were faster at the written than the auditory mode of presentation. Other than for the global target words, the reactions to local target words were not significantly faster at audiovisual than at auditory text presentation.

From the negative correlations of both target words with the text comprehension score we conclude that the reaction times to the local and global target words are a valid measure of text comprehension. Additionally, the moderate magnitude of these correlations classifies this measure as an informative supplement to the reaction times to global and local target words as indicators of text comprehension.

2.3.4 Discussion

The purpose of this study was to compare—between different modes of text presentation—the extent to which 9- and 11-year-olds construct inferences to establish local and global coherence. Following well-established theoretical accounts of text comprehension, it is the construction of a mental representation that is coherent on a local and especially a global level that defines text comprehension (e.g., Graesser et al., 1994; Rizzella & O'Brien, 1996). To measure global and local inferences, we employed a reaction-time-based word-recognition task. The reaction times to target words associated with a protagonist's super- or subordinate goals, or obstacles preventing the protagonist from reaching the superordinate goal, serve as indicators of the extent to which coherence has been constructed on a global or a local level, respectively. In addition, the participants answered questions about the text content that required either remembering a text detail or making an inference. These additional questions serve the purpose of validating and at the same time supplementing the reaction times as a measure of text comprehension.

We found that both the 9- and the 11-year-olds generally reacted faster to the target words associated with the subordinate goals and obstacles (local target words) than to the target words associated with the superordinate goals (global target words). Establishing local coherence is probably easier than establishing global coherence, as connecting consequent text pieces imposes lower demands on the working memory than connecting more widely separated text pieces. Also, the local target word, by definition, appeared in the final section of the stories and was therefore obviously more active in memory than the global target words at the time of the recognition task. Consequently, this effect which is in line with prevailing theory, presumably is confounded with the position within the text in which the local and global target words were presented.

The main purpose of this study was to examine for each of the 9- and the 11-year-olds whether making inferences to maintain coherence on a local and a global level is eas-

ier with audiovisual than auditory text presentation and whether there are differences between written and auditory text presentation.

2.3.4.1 Audiovisual and auditory text

Both the 9- and the 11-year-olds achieved higher scores in terms of the additional text comprehension questions when texts were presented audiovisually rather than auditorily. To correctly answer the questions, the participants were required to either remember explicitly mentioned text details or to draw inferences. As assumed, this finding aligns very well with a vast amount of previous research that employed comparable measures to capture text comprehension and found in children up to the age of 12 that they benefit from audiovisual text presentation (e.g., Beagles-Roos & Gat, 1983; Gibbons et al., 1986).

We also assumed the reactions to local and global target words as indicators of the processes for achieving local and global coherence to be faster with audiovisual than with auditory presentation. This could be confirmed for the global target words. Both age groups reacted faster with the audiovisual presentation mode, thus indicating that achieving global coherence may be facilitated when texts are presented audiovisually in comparison to auditory versions of the same texts. Consequently, it seems that the general claim, termed the multimedia effect, that people learn better from words and pictures than from words alone seems to apply to the specific aspect of achieving a globally coherent mental representation of a text.

However, the reaction times to local target words did not differ between the auditory and the audiovisual mode of presentation. Therefore, we cannot conclude that an audiovisual text presentation promotes the achievement of local coherence equally. This different effect of audiovisual text presentation on local and global coherence formation may be traced back to the levels of processing each is associated with and to some general assumptions on how pictures affect the processing of verbal information.

Some authors argue that local connections may be more closely linked to the text itself and that they are to some degree already incorporated into the representation of the text base (Graesser & McNamara, 2011). In Study II, as in terms of local coherence in this study, we found that an audiovisual presentation did not promote the memory of the text base. In turn, memory of the situation model was better with audiovisual presentation. We explained this finding with recourse on the integrated model of text and picture comprehension (Schnotz & Bannert, 2003; Schnotz, 2014, see section 2.2.4.2). This model assumes that verbal information is processed by the same three-level process as assumed by Kintsch and van Dijk. In contrast, pictures are assumed to offer a more direct path to situation model construction. Beginning with a (mental) visual perceptual representation of a picture, equivalent to the surface representation of a text, a situation

model is directly constructed by means of analogue structure-mapping (Falkenhainer et al., 1989; Gentner, 1989; Schnotz & Bannert, 2003). Consequently, The text base is not necessarily included as a pre-stage or to serve as a "data base" for situation model construction on the basis of pictures (Schnotz & Baadte, 2015). From these considerations it follows that, in an audiovisual text that consists of verbal information and pictures, the text base may not play the same role as in auditory-only text. This may affect not only the memory of the text base as found in Study I, but also the inferences made in the process of local coherence formation.

Additionally, the multimedia effect implies that verbal and pictorial information is processed independently from one another in parallel processes, which results in dual-coding of the same content (Mayer, 2001; Paivio, 1986). However, in our materials, the pictures did not show the target words but rather depicted the context of the scenario described in the texts. It therefore seems that pictures may not only be effective via mechanisms of dual-coding but also seem to enhance the perceptual simulation of the whole situation described in a story rather than only those elements actually represented in both verbal and pictorial form, as assumed by embodied accounts of language comprehension (Barsalou, 1999; Glenberg & Kaschak, 2002). Consequently, global coherence in the sense of representing the text macrostructure, in this case a protagonist's superordinate goal, may be promoted by an audiovisual presentation to a greater extent than local coherence.

2.3.4.2 Auditory and written text

With the written compared to the auditory mode of text presentation, we found that participants across both age groups scored higher in terms of the questions of text comprehension. This finding indicates that both age groups perform better in terms of making inferences and in terms of the recall of text details with written text presentation.

In addition and more importantly, as indicated by the faster reactions to both local and global target words, children of both age groups seem to be better at making causal connections on a local and a global level and thus at constructing a coherent mental representation when reading a text.

As it has been explained in section 1.3.2.2, written text can have some advantages over auditory text. In contrast to auditory text, written text is usually stable, affording a reader the advantage of far greater control over the speed of processing compared to a listener (Ferreira & Anes, 1994; Imhof et al., 1996). To really benefit from this advantage, the reader is required to have sufficient decoding skills so that the process of decoding itself does not consume most of the cognitive resources. It has already been outlined that children growing up with a language with a relatively transparent orthography, such as German, supposedly become fluent decoders at an earlier age than

children growing up with less transparent orthographies such as English (see sections 1.3.2.2 and 2.1.4.1). This may explain why the participating children aged 9 and 11 benefit from a written text presentation, although previous studies with English-speaking participants found that it may take up the 12th grade until decoding skills become fully automated (Sticht & James, 1984; Hildyard & Olson, 1978). In line with our current findings, also Diergarten and Nieding (2016) found that by the age of 10 German children performed better at written text presentation in terms of emotional inferences during text comprehension.

Participants in the Diergarten and Nieding study were instructed to read the texts aloud, which has been found to be advantageous because, in addition to keeping the reader's attention high, the information is perceived both visually and auditorily (Prior et al., 2011). In the current study, however, we did not give specific instructions on whether the participants should read the texts out loud or in silence. Therefore, the participants read the texts the way they felt most comfortable. However, the processing of written text, also when read silently, involves processes of grapheme-phoneme conversion, resulting in a sort of "inner speech" and therefore is additionally encoded phonologically to some degree even by skilled readers (Rieben & Perfetti, 1991; Shankweiler, 1999). This form of "dual-coding" of the verbal information may be another reason why written text seems to be beneficial for achieving both local and global coherence and for the ability to answer detail and inference questions..

2.3.4.3 Limitations and conclusion

In section 1.2.3, we briefly discussed working memory capacity as a relevant influence upon the process of constructing a coherent mental representation of a text. However, we did not include a measure of working memory, because our main focus lay in comparing three different modes of presentation. Future studies comparing different forms of verbal materials and their combination with pictures could include a measure of working memory to further break down the mechanisms at work when processing different modes of text presentation. As working memory develops over the course of childhood (Schneider, 2015), including younger children may also be worth considering. However, the task we employed in this study may be too demanding for younger children, because it requires concentration over quite a long period of time. Participants who received the written materials also reacted to target words that were presented to them in written form, while the participants in the auditory and the audiovisual presentation mode were presented the target words auditorily. One could argue that when target words are presented in written form, participants' reading speed may have influenced the reaction times in contrast to an auditory presentation of the target words. However, there are several reasons why we think that this does not severely limit the

validity of our conclusion about global and local coherence with written and auditory modes of presentation. As reading speed develops with age, an interaction between the factors Age Group and Presentation Mode, which we did not find, would have indicated such an influence. Instead, we found an effect of presentation mode across both age groups. At least for the global target words, we found faster reactions with both the written and the audiovisual modes compared to the auditory mode of presentation. Assuming reading speed as a relevant influence, one would also expect a difference between the written and the audiovisual condition [4]. Furthermore, the additional questions for text comprehension, which are not based on reaction times and thus reading speed, support the findings of the reaction time task.

In conclusion, the current study confirms and extends the findings of existing cross-media studies. Both age groups' performances in answering questions that required either the recall of text details or making an inference was better with the audiovisual and the written modes than with the auditory mode of text presentation. However, our main focus, lay in measuring inferences necessary to construct global and local coherence, which have been identified as crucial processes for text comprehension. When comparing auditory with audiovisual presentation, the current study suggests that pictures added to auditory text have a beneficial effect on the ability to establish global coherence. To establish coherence on a local level, pictures do not seem to provide an added value. Written text, in turn, seems to be advantageous over auditory text for achieving both local and global coherence. This result pattern shows that whether and how different modes of presentation affect text comprehension depends on the specific cognitive process.

[4] This comparison was not part of the results section because we did not make any assumptions about this. The 3 (Presentation Mode) $\times$ 2 (Age Group) $\times$ 2 (Target Word)-ANOVA reported in section 2.3.3 revealed a p-value of $p > .999$ when comparing reaction times to global target words between audiovisual and written presentation. The respective p-value was $p = .496$ in terms of local target words.

3 General Discussion

The three studies described in this thesis all compared processes of text comprehension between different modes of presentation in children of different age groups and, as in Study I, adults. To define and operationalize comprehension, this thesis refers to well-established theories of text comprehension that assume comprehension to involve constructing a coherent mental representation of the subject matter described in the text (e.g., Kintsch, 1998). Moreover, it is agreed upon that this process includes three levels of mental representations: a representation of the text surface, a representation of the text base, and finally a situation model.

Studies I and II focused on the latter aspect of text comprehension and examined whether the memory of the text surface, the text base, and the situation model differed between written, auditory, and audiovisual text presentation in 8- and 10-year-olds and adults (Study I, section 2.1), and between auditory and audiovisual text presentation in 7-, 9-, and 11-year-olds (Study II, section 2.2). Study III (section 2.3) further explored the aspect of coherence and examined whether the ability to achieve local and global coherence of a text differs between written, auditory, and audiovisual text presentation in 9- and 11-year-olds.

3.1 Summary and Evaluation of the Main Findings

What are the conclusions to be drawn in terms of the comparisons between auditory and audiovisual and between auditory and written text presentation, the main findings of three studies considered as a whole?

Table 9 provides an overview of the included presentation modes, age groups, the addressed aspects of text processing, and the main findings of the three studies. The following discussion is separated into two major sections. The first section (3.1.1) summarizes and integrates the main findings of all three studies regarding the comparison between audiovisual and verbal, meaning auditory and written, presentation. The second section (3.1.2) deals with the comparison between written and auditory presentation and therefore deals with the Studies I and III.

Table 9. *Overview of the main findings of the three studies.*

	Presentation modes	Age Groups	Outcome	Main Findings
Study I	w, a, av	8, 10, adults	Memory of TS, TB, and SM	Situation model: 8: a, av > w 10, adults: a = av = w
Study II	a, av	7, 9, 11	Memory of TS, TB, and SM	Across all age groups: SM, TS: av > a TB: a > av
Study III	w, a, av	9, 11	Local and global coherence, detail and inference questions	Across both age groups: global: av > a; w > a local: av = a; w > a questions: av > a; w > a

Note: The numbers in the "Age-Groups"-column refer to the age of the participants, e.g. "8" stands for 8-year-old children. "Outcome" refers to the specific aspect of cognitive text processing that was addressed in the respective studies. "TS" stands for text surface, "TB" for text base, and "SM" for situation model. "Av" stands for audiovisual, "a" for auditory, and "w" for written. ">" signifies that the respective presentation mode was superior. Note that for local and global coherence, in this overview I refer to the underlying process and not to the reaction times.

3.1.1 Audiovisual and verbal-only (auditory and written) presentation

The following overview of the findings regarding the comparison between audiovisual and verbal-only text presentation will first deal with the results of Study I (see 2.1), in which an audiovisual mode of presentation was contrasted with both a written and an auditory condition. The second section discusses the findings of the Studies II and III (see sections 2.2 and 2.3). Other than Study I, in these studies we analyzed and discussed comparisons of audiovisual and auditory (but not written) text presentation.

3.1.1.1 Study I

First of all, the results of Study I revealed an effect of the presentation mode on the level of situation model, but not on the levels of text surface and text base. Moreover, the results of this study indicate that only 8-year-olds, but not the 10-year-olds and the adults, differ in terms of their memory of the situation model depending on the mode of text presentation.

Eight-year-olds displayed better memory of the situation model at audiovisual than at written text presentation. This finding indicates that adding pictures to verbal materials facilitates situation model construction and therefore comprehension, and corresponds to findings of previous cross-media studies and theoretical considerations about the beneficial effects of audiovisual materials (see section 1.3.1). Based on these considerations, it was expected that the 8-year-olds would also benefit from an audiovisual

presentation in comparison with an *auditory* presentation of the verbal materials. However, this assumed effect did not become apparent. The 8-year-olds' memory of the situation model did not differ between auditory and audiovisual text presentations. This finding may be explained with the experimental setup of this study: The mode of presentation was manipulated within subjects and in order to restrict the overall duration of this experiment to a reasonable time frame, only four stories of each presentation mode were presented. This may have been a too small amount for the assumed beneficial effect of audiovisual over auditory presentation to become apparent.

Also the assumptions regarding the adults and the 10-year-olds were not supported by the data. The adults' memory of the situation model does not seem to differ between written, auditory, and audiovisual presentation. According to the redundancy or expertise reversal effect it was instead expected that the pictures at the audiovisual presentation would interfere with meaningful processing and that therefore the adults would display less memory of the situation model at audiovisual presentation. To evaluate this finding it is important to note that the redundancy and expertise reversal effects were initially proposed with reference to instructional settings, but not narrative text comprehension. Moreover, studies that found effects used different materials (e.g., news stories or advertisement DeFleur et al., 1992; Furnham & Gunter, 1987). These types of texts are supposedly processed differently than narrative texts, with more emphasis on actually understanding more complex factual connections or abstracts concepts, in contrast to simply following a narrative about everyday activities. Understanding these child-appropriate narrative texts may require so little effort from an adult recipient that, although the picture contain redundant details, this does not have detrimental effects. Thus, possible detrimental effects of pictures may not be generalized over narrative texts, which put other demands on the recipient.

As found for the adults, the 10-year-olds' memory of the situation model did not differ between modes of presentation, although it was expected that they (like the 8-year-olds) would benefit from an audiovisual presentation. Maybe the same explanation that was considered for not finding a difference between auditory and audiovisual presentation may apply here: Only four stories in each presentation mode were too little for each one to unfold its assumed effect on the recipient. Other than for the 8-year-olds, this study did not reveal a difference between written and audiovisual presentation mode. Instead, the 10-year-olds' memory of the situation model at written presentation was at the same level as the adults' and significantly greater than the 8-year-olds'. Maybe these advanced written text comprehension skills eliminate the need for pictures among the 10-year-olds.

3.1.1.2 Studies II and III

To account for the supposedly methodological explanation for not finding the assumed beneficial effect of audiovisual over auditory presentation in the 8- and 10-year-olds in Study I, Study II was conducted. Study II was limited to a comparison between an auditory and an audiovisual mode of text presentation and included 7-, 9-, and 11-year-old participants. This limitation to only two modes of presentation in a within-subjects comparison allowed presentation of eight stories in each mode of presentation (thus twice as many as in Study I).

The findings of Study II suggest that, across all of the included age groups, the memory of all three levels differs between auditory and audiovisual text presentation. In line with our assumptions, memory of the situation model was better at audiovisual than at auditory text presentation. Thus, audiovisual presentation seems to be advantageous for text comprehension, not only over a written version of the same semantic content - as indicated by the results of Study I, but also over an auditory presentation and thus over verbal-only information in general.

The results of Study III point into the same direction. These results indicate that children aged 9 and 11 draw inferences to establish global coherence to a greater extent at audiovisual than at auditory text presentation. Establishing global coherence is a defining aspect of the situation model and thus text comprehension, as this process involves integrating related information over widely separated text pieces and integrating every new incoming piece of information with the text's macro structure. Furthermore, both 9- and 11-year-olds were better at answering detail and inference questions at an audiovisual than at an auditory text presentation (Study III). This finding replicates the previously often-found effect that audiovisual materials enhance memory and comprehension (see section 1.3, p. 22 for an overview).

Also children's memory of the text surface was better at audiovisual than at auditory presentation (Study II). Memory of the text base of these children aged 7, 9, and 11, in contrast, was greater at auditory presentation. Also, contrary to the initial assumption, and other than in terms of global coherence, findings of Study III do not suggest that 9- and 11-year-olds seem to draw inferences to establish local coherence to a greater extent at audiovisual compared to an auditory text presentation.

It may seem counter-intuitive at first glance that, while memory of the situation model and the text surface is greater at audiovisual presentation, the memory of the text base, in contrast is greater at auditory presentation, as these levels supposedly are constructed in a hierarchical order (Kintsch, 1998). We explained this result pattern can be explained in the context of the *Integrated Model of Text and Picture Comprehension* (Schnotz & Bannert, 2003; Schnotz, 2005, 2014, see section 1.3.1.3, p. 27). This model assumes the same three levels of mental representations to be involved in the processing

of verbal information as Kintsch and van Dijk (see section 1.1). Pictures, in contrast, are assumed to be processed without necessarily involving a propositional representation (or "text" base). Instead, the information necessary to construct a situation model of the subject matter displayed in a picture is read off a visual perceptual representation (equivalent to the text surface) via processes of analog structure-mapping (Falkenhainer et al., 1989; Gentner, 1989; Schnotz & Bannert, 2003). Thus, propositional structures presumably play a less important role in the processing of audiovisual text, in which verbal information is enhanced with pictures. This may explain why memory of the text base was greater at auditory presentation. Furthermore, when less cognitive resources are involved in text base construction, these resources may in turn be available to remember text surface information at audiovisual presentation. The finding that local coherence was supported by an audiovisual text presentation also makes sense when considering the assumptions of the Integrated Model of Text and Picture Comprehension. While of course local connections are included into the situation model, some authors argue that these local connections are to some extend already made on the level of the text base (Graesser et al., 2001). Thus, local inferences may not necessarily be promoted when adding pictures because pictures according to the assumptions of the integrated model of text and picture comprehension rather have an effect on a macro-structural processing level, such as global coherence processes on the level of the situation model.

Do younger children benefit from audiovisual information even more? These just discussed effects of presentation mode (auditory versus audiovisual) have been found across all of the age groups included in Studies II and III. However, in Study II the effect of presentation mode was expected to be more pronounced in the 7-year-olds. Instead, it was found that 7-, 9- and 11-year-olds seem to benefit from an audiovisual presentation to a comparable extend. This assumption was in first place motivated by theoretical considerations with only one - in a broader sense comparable - other study, which reported that an audiovisual presentation promotes predictive inferences among children aged 6, but not among children aged 8 and older (Unsöld & Nieding, 2009). As some authors argue, predictive inferences are not routinely generated and are not always necessary for understanding (Graesser et al., 1994). Unsöld and Nieding (2009) themselves suggest that their finding may indicate that the 6-year-olds use the extra information provided by the pictures of the audiovisual materials to rather "over-represent" information in the situation model while the older children represent information more parsimoniously and strategically. It is plausible to assume that this mechanism may apply to the highly specific process of predictive inference making. Memory of the situation model in general seems to be promoted by audiovisual materials irrespective of the chil-

dren's developmental status. Hence, there is no indication of a mechanism of (further) perceptual support to be effective in children's comprehension of audiovisual texts.

Similarly, Hauf (2016) reports that 6-year-old children, in contrast to older children, seem to benefit from an audiovisual presentation with regard to the mental representation of object shapes, which may also count as elaborative inferences (Graesser et al., 1994). She also argues that audiovisual presentation may trigger these processes in younger children, for whom a detailed representation that contains presumably "unnecessary" information has an adaptive function because they need this augmented representation to gain understanding at all. In contrast, older children employ more efficient processing strategies and represent information parsimoniously and therefore do not use the additional information offered by the pictures to further enrich their situation model.

3.1.1.3 How do pictures become effective?

Although this question was not the central question to be addressed by these three studies and the experiments were not specifically designed to give insights into the cognitive processes of picture comprehension, some conclusions may be derived nevertheless. In all three studies, it was taken care that the pictures served as a useful illustration of the text content, but at the same time did not display the aspects that were involved in the distractor sentences of the sentence recognition task in Studies I and II, or the target words in Study III. Thus, the beneficial effects of pictures we found in these three studies cannot be accounted to the fact that the participants in the audiovisual had been presented the relevant information twice - as a word and as a picture. Instead, our findings indicate that audiovisual presentation seems to enhance a perceptual simulation of the whole situation and not just the elements actually shown in the pictures. Thus, a picture seems to support processing on a macro-structural or gist-level. Under this premise, it is plausible that lower-level processes, such as local coherence or mental representations of the propositional structure, are not supported at an audiovisual text presentation (see also section 3.1.1.2).

The role of perceptual simulations: an outlook. As I briefly discussed in section 1.1.3.2, the idea of perceptual simulation as the integral mechanism of text comprehension is not undisputed. Currently, theoretical positions emerge that emphasize the importance of propositional representations besides perceptual simulations (e.g., Kaup et al., 2015). Constructing mental representations that contain less perceptual features may even be considered the more efficient way of text processing (Rommers, Meyer, & Huettig, 2013). This corresponds to the findings and discussions by Hauf (2016) that I already outlined in the previous section. These considerations also disclose another per-

spective on the role pictures play for text comprehension among recipients of different ages or with different cognitive abilities respectively. Assuming that older children or adults engage in less perceptual simulation during text comprehension, pictures which have been discussed to promote perceptual simulations should have less of an effect in adults. This may offer an alternative explanation, besides expertise reversal (discussed in section 2.1.4.1), for why adults do not benefit from an audiovisual presentation to the same extent as children.

I suggest that further research on the question on how pictures become effective, should thus consider different age groups, further cognitive abilities that are associated with efficient processing, such as inhibition (see Hauf, 2016, p. 165), and a broad range of tasks that also measure processes that Graesser et al. (1994) would rate as "elaborative".

3.1.1.4 Conclusion: audiovisual and verbal presentation

All in all, the current findings align well with prior research and make sense against the background of theoretical accounts about text comprehension and cognitive processing of pictures. Firstly, by finding that answering questions that required remembering text details or making inferences was easier for 9- and 11-year-olds at audiovisual text presentation (Study III), the findings of a vast amount of previous cross-media studies that assessed text comprehension by employing similar measures (e.g., Carney & Levin, 2002) were replicated. More importantly, by systematically deriving the primary measures of text comprehension from established theories about the cognitive processes that underlie text comprehension (Kintsch, 1998), as it was the overarching goal of these three studies, the overall result pattern offers more differentiated insights. On the one hand, for children at least to the age of 12 both memory of the situation model and establishing global coherence, which can be summarized as processing of the text's gist or macro structure, are facilitated by an audiovisual text presentation. On the other hand, it seems that both the memory of the text base and local coherence, which may be more closely linked to the text base than to situation model processing, do not benefit from additional pictures, probably because the cognitive processing of pictures does not necessarily include mentally representing a propositional structure (i.e. text base) of the content (e.g., Schnotz & Bannert, 2003).

3.1.2 Auditory and written presentation

First of all, the 8-year-olds (Study I), in line with the findings by Sticht and James (1984) and also further findings and considerations about the impact of the language in this context by Diergarten and Nieding (2016) as well as Florit and Cain (2011), display bet-

ter memory of situation model information when they listen to a text as opposed to reading. Consequently, it seems that their approximately one and a half year of formal reading instruction is not enough experience to have developed the decoding skills necessary for their reading ability approximating or surpassing their auditory comprehension abilities.

In contrast, the 10-year-olds and adults that were included in this study did not differ in terms of their memory of the situation model depending on whether they had read or listened to the texts. Furthermore, the 10-year-olds and adults did not differ regarding this aspect. These findings indicate that already 10-year-old fourth graders seem to be skilled enough decoders to understand German narrative texts equally well at auditory and written presentation. The findings of Study III point into the same direction and suggest that a written presentation may even be superior to an auditory presentation. Both 9-year-olds, who as the 10-year-olds in Study I attended the fourth grade, and 11-year-olds (sixth-grade students) seem to be better at written than at auditory presentation with regard to drawing inferences to establish local and global coherence and answering detail and inference questions about the same texts.

Two possible reasons why fluent decoders may not only comprehend texts equally well at both auditory and written presentation but may also even perform better at written presentation were discussed (see section 2.3.4.2). Firstly, not only reading aloud, but also silent reading involves processes of grapheme-phoneme conversion and thus an additional phonological encoding (Rieben & Perfetti, 1991; Shankweiler, 1999). This sort of "inner speech" results in a dual coding of the processed information, which may benefit the recipient's memory and comprehension of the text content. Furthermore, written text, in contrast to auditory text is usually stable, which provides the reader control over the processing speed that can be adjusted to the information processing speed. This is at least the case when the reading time is self-paced as it was in our studies.

Although, it may be generally concluded that the findings of both studies point into the same direction, the question remains why we found an effect of presentation mode in terms of local and global coherence as well as the detail and inference questions (Study III), but not in terms of the memory of the situation model (Study I). Once again, I refer to the experimental set up of Study I to provide a possible explanation. Presenting only four studies in every mode of presentation, especially without any breaks in between, may be a too small amount for effects of presentation mode to become apparent.

3.1.2.1 Conclusion: auditory and written presentation

Two positions on whether processing of auditory and written text differ have been discussed in section 1.3.2. Unitary process views, such as the simple view of reading, assume the same cognitive processes to underlie reading and listening comprehension. From this follows that once sufficient decoding skills have been acquired, it should not matter whether a text is presented in written or auditory form. In contrast, dual process views assume differences between written and auditory texts that affect the way these are processed. Empirical support can be found for both positions, but all in all, we concluded that the findings of available research did not allow to make precise predictions in terms of the studies included into this thesis.

Overall, the findings of the Studies I and III taken together suggest that, in line with the simple view of reading, the recipients' age and thus presumably their decoding ability seems to mediate the effect of presentation mode on the various aspects of text comprehension assessed in the three studies. 8-year-olds showed better memory of the situation model at auditory than at written text presentation.

Older children (10, and 11) and adults in turn, apparently either do not benefit from either form of presentation, as shown in terms of memory of the situation model (Study I), or perform better after reading a text than after listening to it, when it comes to establishing local and global coherence and remembering text details and drawing inferences (Study III). Consequently, the result pattern produced by these studies does not speak in favor of either the unitary or the dual process view.

Furthermore, when comparing our findings with previous research with mostly English materials, it seems to be important to consider the language of the materials before generalizing the results across different languages.

3.2 Discussion of the Methods

To reach a valid general conclusion of the previously described main findings of the studies included in this thesis it is crucial to discuss the limitations and strengths of the applied methods. Some limitations that are specific to the individual studies have been discussed in the limitations sections of the respective studies and will thus not be repeated here.

3.2.1 How to assess text comprehension

The vast majority of studies that compared memory and/or comprehension between different modes of presentation either instructed participants to retell the content or asked specific open- or closed-ended questions (see section 1.3.1, p. 22 for an overview).

These tasks require the recipient to reconstruct what he or she explicitly remembers. This is a disadvantage, especially when studying children, because verbal abilities can be assumed as a confounding factor.

In contrast, the sentence recognition task (Studies I and II) as well as the reaction-time-based word recognition task (Study III) did not require any verbal expression from the participants but a key press by which they indicated whether they thought that a sentence or word respectively had occurred in the previously received text or not. These measures can be classified as implicit or quasi-online measures (Nieding & Ohler, 2004; Graesser et al., 2001). Besides not relying on the participants' verbal expressions, they have the further advantage that they do not consider the final product of comprehension, but capture the underlying processes that are relevant for text comprehension (Nieding & Ohler, 2004), such as memory of the three different levels of mental representation and establishing local and global coherence. Additionally including detail and inference questions as an additional offline measure in Study III further completed the overall picture of the effects of different modes of text presentation on comprehension.

In contrast to many other previous studies, the current studies did not refer to naturally occurring story materials to determine how (well) these are processed and comprehended, but we generated our own texts. Using experimenter-generated materials has sometimes been criticized because non-natural texts may not measure text-processing as it would naturally occur (Singer & Leon, 2013). For our specific purposes, I believe, generating our own materials was the only practicable option because of the vast amount of highly comparable texts that were needed and because of the very specific demands when constructing the sentence recognition task and the word-recognition task. The sentence recognition task for example, as described in section 2.1, required three different versions of each original sentence.

Of course, one may argue that the generalizability of the results may be limited. However, when considering internal validity, this alleged limitation can be discussed as one of the major strengths of this series of studies.

3.2.2 Varying presentation mode within or between subjects

In the first two studies presentation mode was varied *within subjects*. Within-subjects designs have several advantages. It is an efficient way to achieve relatively large experimental power with smaller numbers of participants. Furthermore, we can rule out that the effects of presentation mode that were found in these two studies are confounded by third variables such as working memory capacity or intelligence (to only name a few). This makes these studies stand out from many of the studies included in Carney and Levin's (2002) review study. However, the result pattern of Study I revealed a disad-

vantage. While trying to keep the duration of the experiment within a reasonable limit, only four stories of each mode of presentation could be presented and the assumed difference between auditory and audiovisual presentation did not become apparent. We discussed that only four stories, especially without any breaks in between, may simply have been too small an amount for the effect to become apparent (see Section 2.1, p.51). For this reason, presentation mode was varied between subjects in Study III, in which again compared three different presentation modes were compared. Although assignment to conditions was random, third variables may have influenced the results (see the next section for a detailed discussion of this issue). However, the findings of Study III align well with the findings of Studies I and II.

3.3 Relevant Covariates and Implications for Further Research

These three studies examined how the mode of presentation affects cognitive processing and comprehension of narrative texts. By comparing written, auditory, and audiovisual versions of the same texts, we varied the sensory channel through which the information was perceived (visual and auditory), and the signs used to present the information (verbal and pictorial). Of course, further variables on the level of the text as well as on the level of the recipient can be assumed to have an effect on text comprehension.

A relevant aspect on the level of the text may be its goal structure. As it has been outlined in section 1.2.4, protagonists' goals, as a first narrative organizing principles, have been found to provide coherence for children who retell narrative texts (e.g., Lynch et al., 2008). Therefore, it may be interesting to systematically manipulate the extend to which protagonists' goals are emphasized in a replication of Study III and examine the effect of goal emphasis on the extend to which local and global coherence is established. However, special attention has to be paid on how this goal emphasis may be best implemented, as indicated by the findings of a study by Diergarten and Nieding (2016) who found that goal emphasis helped adults but not 8- and 10-year-old children build emotional inferences. They suggest that incorporating the additional information provided by the goal emphasis exceeded the children's mental capacities. To avoid this confounding of emphasis and the amount of propositions to be processed, it is imperative to think of a more parsimonious way to emphasize protagonists' goals.

As already discussed previously (see for example, section 2.2.4.3, p.64), the pictures included in the audiovisual versions were static pictures. To further approximate the audiovisual materials to reality, animated pictures could be used. According to the assumption that situation model construction includes a perceptual simulation it may be assumed that animated pictures may further facilitate this process. As hinted in section 1.1.3.2, there is evidence of a process called motor resonance, meaning that verbally de-

scribed actions are motorically simulated (Glenberg & Kaschak, 2002; Zwaan & Taylor, 2006; Zwaan, Taylor, & de Boer, 2010). Furthermore, findings from studies employing neuroimaging methods imply that the processing of verbally described actions is associated with the same activation pattern in motor and premotor areas as real action or the observation of action (Pulvermüller, 2005; Aziz-Zadeh, Wilson, Rizzolatti, & Iacoboni, 2006). These findings may allow the assumption that observing action presented by an animated picture or a film may support comprehension of a corresponding text describing this action via mirror neuron structures (see also Seger, Wannagat, & Nieding, in preparation). Furthermore, a meta-analysis on the effectiveness of animations for instructional purposes when learning from expository text indicates a beneficial effect of animations over static pictures (Höffler & Leutner, 2007).

Furthermore, it may be worth researching whether the findings presented in this thesis can be generalized over different text genres. Especially expository texts, as one of the primary media used for learning in schools, are of interest. In the context of expository text it would also be interesting to not only include realistic drawings but also different types of pictures. As described in section 1.3.1, charts and graphs, although they do not require knowledge of an arbitrary symbol system, such as language, represent a concept based on structural and not physical similarity. This implies that, contrary to realistic pictures, some knowledge about representational conventions is required to properly interpret a line graph for example.

On the side of the recipient, prior knowledge and reading ability have already been discussed as relevant influences on text comprehension as they were indirectly included both by investigating different age groups who supposedly differ with respect to their prior knowledge and the reading abilities simply because they differ in the amount of formal reading instruction received so far. Prior knowledge as an important determinant of comprehension has been extensively discussed in the literature and experimental (e.g., Hyönä & Lorch, 2004; Lorch, 1989) and quasi-experimental studies (e.g., Schneider, Körkel, & Weinert, 1989; Nieding, 2006) have shown that context cues, such as headings, and domain specific prior knowledge facilitate comprehension. Controlling prior knowledge involves the risk that an assessment anticipates what is later asked for in a comprehension task. By comparing presentation modes within subjects (Studies I and II) and by choosing topics from children's everyday lives we minimized the risk of domain knowledge as a confounding influence. An experimental manipulation of text signaling devices as a further determinant for text comprehension may be worth considering in subsequent studies.

Of course it would be best to not only infer children's reading ability from their grade level, but to assess it using standardized tests. According to my opinion, this would only work for an investigation of effects within age groups because none of the existing

tests, to my knowledge, is suited for comparisons between age groups and would result in either floor effects in less skilled readers or ceiling effects in highly skilled readers. Using an age appropriate test for every age group would require the exact same underlying construct and operationalization and is therefore problematic.

Furthermore, executive functions of the working memory, especially updating capacity, is discussed as a relevant prerequisite for successful text comprehension. Updating describes the process of continuous substitution of working memory content according to a relevant criterion (Morris & Jones, 1990). Likewise, in text comprehension, a recipient who follows a narrative has to constantly evaluate for every new incoming pieces of information whether it is relevant for the story or not. The working memory, due to its limited capacity, is constantly updated to maintain coherence within the situation model (Morris & Jones, 1990). Several studies have accordingly found that good and poor comprehenders differ in terms of their updating capacity and their ability to suppress no longer relevant working memory content and replace it with new information (Palladino, Cornoldi, De Beni, & Pazzaglia, 2001; Carretti, Cornoldi, De Beni, & Romanò, 2005; Diergarten & Nieding, 2016). For that reason, including updating capacity as a further predictor of the ability to achieve global coherence may be a promising research question.

Research has also shown that media sign literacy, besides intelligence, predicts verbal and mathematical academic precursor skills in preschool children (Nieding et al., 2017). According to Nieding and Ohler (2008), media sign literacy corresponds to Potter's (1998) concept of "rudimentary skills" of media literacy, which are a subcomponent of his media literacy model and describes the ability to understand the symbol systems that occur in different information and entertainment media. Further, media sign literacy predicted children's learning from audiovisual media, such as an educational film and a hypermedia learning environment (Diergarten, Möckel, Nieding, & Ohler, 2017). Most importantly in the context of this thesis, media sign literacy has been shown to support 5-year-old children's ability to make exact emotional inference and to thus to accurately monitor a character's emotional state within the situation model (Diergarten & Nieding, 2015).

Considering this discussion of different factors that influence text comprehension it becomes obvious that, firstly, much more than just the presentation mode is to be considered when trying to figure out how to present information in order to achieve maximum benefit for the recipient. Secondly, incorporating all these factors in one study, which would provide the complete picture, is a very ambitious endeavor.

3.4 Overall Conclusion and Practical Implications

All in all, this series of studies demonstrates that, in part depending on the recipient's age, the presentation mode affects how (well) texts are processed and understood. Although this research can be classified as rather basic research, the overall message to be derived from this has far-reaching practical consequences: Whenever you want to share information, it is decisive to carefully consider whether you present it as a written version, produce an auditory version, or even add pictures. This applies to various situations in which an audience is to be instructed, informed, or simply entertained to ensure social and cultural participation for everyone. This overall conclusion is not new information, but by introducing new ways of operationalizing and measuring comprehension that are closely oriented on the prevailing view on the processes involved in text comprehension, I achieved a much more detailed picture. Which mode of presentation is beneficial seems to depend on the specific cognitive process. Children aged 7 to 12 display better memory of the situation model and are better at drawing inferences to achieve global coherence when static pictures that illustrate the content are added to an auditorily presented text. These processes are associated with deep processing and thus what text comprehension is essentially about. In terms of lower level processes, associated with more text-base related processing, such as remembering text base information and drawing inferences to maintain local coherence, the general claim of beneficial effects of an audiovisual presentation cannot be upheld. Furthermore, in terms of the comparison of auditory and written presentation, the findings of the experiments included in this thesis, indicate that, in contrast to many previous research with English-speaking participants, already fourth-graders can benefit from written text, at least in languages with a transparent orthography as German.

References

Ackerman, B. P. (1986). Referential and causal coherence in the story comprehension of children and adults. *Journal of Experimental Child Psychology*, *41*(2), 336–366.

Ackerman, B. P., & McGraw, M. (1991). Constraints on the causal inferences of children and adults in comprehending stories. *Journal of Experimental Child Psychology*, *51*(3), 364–394.

Albrecht, J. E., & O'Brien, E. J. (1993). Updating a mental model: Maintaining both local and global coherence. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(5), 1061–1070.

Alloway, T. P., Gathercole, S. E., & Pickering, S. J. (2006). Verbal and visuospatial short-term and working memory in children: Are they separable? *Child Development*, *77*(6), 1698–1716.

Anderson, J. R. (1996). *Kognitive Psychologie*. Heidelberg: Spektrum.

Artelt, C., McElvany, N., Christmann, U., Richter, T., Groeben, N., Köster, J., ... Ring, K. (2007). *Förderung von Lesekompetenz – Expertise* (BMBF, Ed.). Bonn: BMBF.

Atkinson, R., & Shiffrin, R. (1968). Human memory: A proposed system and its control processes. In K. Spence (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2, pp. 89–195). New York: Academic.

Aziz-Zadeh, L., Wilson, S. M., Rizzolatti, G., & Iacoboni, M. (2006). Congruent embodied representations for visually presented actions and linguistic phrases describing actions. *Current Biology*, *16*(18), 1818–1823.

Baddeley, A. (1986). *Working memory*. Oxford, England: Clarendon Press.

Baddeley, A. (1997). *Human memory: Theory and practice*. Hove: Psychology Press.

Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences*, *4*(11), 417–423.

Baddeley, A. (2002). Is working memory still working? *European Psychologist*, *7*(2), 851–864.

Ballstaedt, S. P. (1997). *Wisssensvermittlung*. Weinheim: Beltz.

Barsalou, L. W. (1999). Language comprehension: Archival memory or preparation for situated action? *Discourse Processes*, *28*(1), 61–80.

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, *59*, 617–645.

Bates, E., Masling, M., & Kintsch, W. (1978). Recognition memory for aspects of dialogue. *Journal of Experimental Psychology: Human Learning and Memory*, *4*(3), 187–197.

Beagles-Roos, J., & Gat, I. (1983). Specific impact of radio and television on children's

story comprehension. *Journal of Educational Psychology*, *75*(1), 128–137.

Bohn-Gettler, C. M., Rapp, D. N., van den Broek, P., Kendeou, P., & White, M. J. (2011). Adults' and children's monitoring of story events in the service of comprehension. *Memory & Cognition*, *39*(6), 992–1011.

Brainerd, C. J., & Reyna, V. F. (1992). Explaining "memory free" reasoning. *Psychological Science*, *3*(6), 332–339.

Bransford, J. D., Barclay, J. R., & Franks, J. J. (1972). Sentence memory: A constructive versus interpretive approach. *Cognitive Psychology*, *3*(2), 193–209.

Brown, G. D., & Watson, F. L. (1987). First in, first out: Word learning age and spoken word frequency as predictors of word familiarity and word naming latency. *Memory & Cognition*, *15*(3), 208–216.

Brunyé, T. T., Ditman, T., Mahoney, C. R., Walters, E. K., & Taylor, H. A. (2010). You heard it here first: Readers mentally simulate described sounds. *Acta Psychologica*, *135*(2), 209–215.

Bäuerlein, K. (2014). *Leseverständnisdiagnostik in der Sekundarstufe-Theoretische Grundlagen sowie Konstruktion und empirische Erprobung der Lesetests LESEN 6-7 und LESEN 8-9*. Würzburg: Würzburg University Press.

Byrne, M., & Curtis, R. (2000). Designing health communication: Testing the explanations for the impact of communication medium on effectiveness. *British Journal of Health Psychology*, *5*(2), 189–199.

Carney, R. N., & Levin, J. R. (2002). Pictorial illustrations still improve students' learning from text. *Educational Psychology Review*, *14*(1), 5–26.

Carretti, B., Cornoldi, C., De Beni, R., & Romanò, M. (2005). Updating in working memory: A comparison of good and poor comprehenders. *Journal of Experimental Child Psychology*, *91*(1), 45–66.

Christie, B., & Collyer, J. (2008). Do video clips add more value than audio clips? presenting industrial research and development results using multimedia. *Behaviour & Information Technology*, *27*(5), 395–405.

Chun, D. M., & Plass, J. L. (1997). Research on text comprehension in multimedia environments. *Language Learning & Technology*, *1*(1), 60–81.

Clement, C. A., & Gentner, D. (1991). Systematicity as a selection constraint in analogical mapping. *Cognitive Science*, *15*(1), 89–132.

Corston, R., & Colman, A. M. (1997). Modality of communication and recall of health-related information. *Journal of Health Psychology*, *2*(2), 185–194.

Curtis, M. E. (1980). Development of components of reading skill. *Journal of Educational Psychology*, *72*(5), 656–669.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*(4), 450–466.

Danks, J. H., & End, L. J. (1987). Processing strategies for reading and listening. In

R. Horowitz & S. J. Samuels (Eds.), *Comprehending oral and written language* (pp. 271–294). San Diego, CA: Academic Press.

DeFleur, M. L., Davenport, L., Cronin, M., & DeFleur, M. (1992). Audience recall of news stories presented by newspaper, computer, television and radio. *Journalism & Mass Communication Quarterly*, *69*(4), 1010–1022.

De Saussure, F. (1960). *Course in general linguistics*. London: Peter Owen.

Diergarten, A. K., Möckel, T., Nieding, G., & Ohler, P. (2017). The impact of media literacy on children's learning from films and hypermedia. *Journal of Applied Developmental Psychology*, *48*, 33–41.

Diergarten, A. K., & Nieding, G. (2015). Children's and adults' ability to build online emotional inferences during comprehension of audiovisual and auditory texts. *Journal of Cognition and Development*, *16*(2), 381–406.

Diergarten, A. K., & Nieding, G. (2016). Online emotional inferences in written and auditory texts: a study with children and adults. *Reading and Writing*, *29*(7), 1383–1407.

Dopkins, S., Klin, C., & Myers, J. L. (1993). Accessibility of information about goals during the processing of narrative texts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(1), 70–80.

Ellis, P. D. (2010). *The essential guide to effect sizes: Statistical power, meta-analysis, and the interpretation of research results*. Cambridge University Press.

Engelen, J. A., Bouwmeester, S., de Bruin, A. B., & Zwaan, R. A. (2011). Perceptual simulation in developing language comprehension. *Journal of Experimental Child Psychology*, *110*(4), 659–675.

Estevez, A., & Calvo, M. G. (2000). Working memory capacity and time course of predictive inferences. *Memory*, *8*(1), 51–61.

Eysenck, M., & Keane, M. (2000). *Cognitive Psychology* (4th ed.). Hove: Psychology Press.

Eysenck, M., & Keane, M. (2005). *Cognitive Psychology* (5th ed.). Hove: Psychology Press.

Falkenhainer, B., Forbus, K. D., & Gentner, D. (1989). The structure-mapping engine: Algorithm and examples. *Artificial Intelligence*, *41*(1), 1–63.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G* power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191.

Ferreira, F., & Anes, M. (1994). Why study spoken language? In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 34–36). San Diego, CA: Academic Press.

Field, A. (2013). *Discovering statistics using ibm spss statistics*. London: Sage.

Fincher-Kiefer, R. (2001). Perceptual components of situation models. *Memory & Cognition*, *29*(2), 336–343.

Fletcher, C. R., & Bloom, C. P. (1988). Causal reasoning in the comprehension of simple narrative texts. *Journal of Memory and language*, *27*(3), 235–244.

Fletcher, C. R., & Chrysler, S. T. (1990). Surface forms, textbases, and situation models: Recognition memory for three types of textual information. *Discourse Processes*, *13*(2), 175–190.

Fletcher, J., & Tobias, S. (2005). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The cambridge handbook of multimedia learning* (pp. 117–134). New York, NY: Cambridge Univ Press.

Florit, E., & Cain, K. (2011). The simple view of reading: Is it valid for different types of alphabetic orthographies? *Educational Psychology Review*, *23*(4), 553–576.

Furnham, A., Benson, I., & Gunter, B. (1989). Memory for television commercials as a function of the channel of communication. *Social Behaviour*, *2*(2), 105–112.

Furnham, A., & Gunter, B. (1985). Sex, presentation mode and memory for violent and non-violent news. *Journal of Educational Television*, *11*(2), 99–105.

Furnham, A., & Gunter, B. (1987). Effects of time of day and medium of presentation on immediate recall of violent and non-violent news. *Applied Cognitive Psychology*, *1*(4), 255–262.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, *7*(2), 155–170.

Gentner, D. (1989). The mechanisms of analogical learning. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 197–241). Cambridge: Cambridge Univ Press.

Gernsbacher, M. (1990). *Language comprehension as structure building*. Hillsdale: Erlbaum.

Gibbons, J., Anderson, D. R., Smith, R., Field, D. E., & Fischer, C. (1986). Young children's recall and reconstruction of audio and audiovisual narratives. *Child Development*, *57*(4), 1014–1023.

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review*, *9*(3), 558–565.

Gough, P., Hoover, W., & Peterson, C. (1996). Some observations on a simple view of reading. In C. Cornoldi & J. Oakhill (Eds.), *Reading comprehension difficulties: Processes and intervention* (pp. 1 –13). Mahwah, NJ: Erlbaum.

Gough, P., & Tunmer, W. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, *7*(1), 6–10.

Graesser, A. C., & McNamara, D. S. (2011). Computational analyses of multilevel discourse comprehension. *Topics in Cognitive Science*, *3*(2), 371–398.

Graesser, A. C., Millis, K. K., & Zwaan, R. A. (1997). Discourse comprehension. *Annual Review of Psychology*, *48*(1), 163–189.

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narra-

tive text comprehension. *Psychological Review*, *101*(3), 371–395.

Graesser, A. C., Wiemer-Hastings, P., & Wiemer-Hastings, K. (2001). Construction inferences and relations during text comprehension. In T. Sanders, J. Schilperoord, & W. Spooren (Eds.), *Human cognitive processing: Vol. 8. text representation. linguistic and psycholinguistic aspects* (pp. 249 – 271). Amsterdam: Benjamins.

Greenhoot, A. F., & Semb, P. A. (2008). Do illustrations enhance preschoolers' memories for stories? age-related change in the picture facilitation effect. *Journal of Experimental Child Psychology*, *99*(4), 271–287.

Gunter, B., Furnham, A., & Griffiths, S. (2000). Children's memory for news: A comparison of three presentation media. *Media Psychology*, *2*(2), 93–118.

Guttmann, J., Levin, J. R., & Pressley, M. (1977). Pictures, partial pictures, and young children's oral prose learning. *Journal of Educational Psychology*, *69*(5), 473–480.

Hauf, J. (2016). *Experimentelle Untersuchung der Entwicklung der kognitiven Textverarbeitung unter Berücksichtigung unterschiedlicher Darbietungsmodalitäten* (Dissertation). Universität Würzburg, Fakultät für Humanwissenschaften.

Hayes, D. S., Kelly, S. B., & Mandel, M. (1986). Media differences in children's story synopses: Radio and television contrasted. *Journal of Educational Psychology*, *78*(5), 341–346.

Hegarty, M., Carpenter, P., & Just, M. (1996). Diagrams in the comprehension of scientific texts. In R. Barr, M. L. Kamil, P. Mosenthal, & P. D. Pearson (Eds.), *Handbook of Reading Research. Vol. II* (pp. 641–668). Mahwah, NJ: Lawrence Erlbaum Associates.

Helder, A., van Leijenhorst, L., & van den Broek, P. (2016). Coherence monitoring by good and poor comprehenders in elementary school: Comparing offline and online measures. *Learning and Individual Differences*, *48*, 17–23.

Hemforth, B., & Konieczny, L. (2008). Saetze und Texte verstehen und produzieren. In J. Muesseler & W. Prinz (Eds.), *Lehrbuch Allgemeine Psychologie* (pp. 504–552). Heidelberg: Spektrum.

Hess, D. J., Foss, D. J., & Carroll, P. (1995). Effects of global and local context on lexical processing during language comprehension. *Journal of Experimental Psychology: General*, *124*(1), 62–82.

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

Hitch, G. J., Halliday, S., Schaafstal, A. M., & Schraagen, J. M. C. (1988). Visual working memory in young children. *Memory & Cognition*, *16*(2), 120–132.

Höffler, T. N., & Leutner, D. (2007). Instructional animation versus static pictures: A meta-analysis. *Learning and Instruction*, *17*(6), 722–738.

Hoover, W., & Gough, P. (1990). The simple view of reading. *Reading and Writing*, *2*(2), 127–160.

Horton, W. S., & Rapp, D. N. (2003). Out of sight, out of mind: Occlusion and the

accessibility of information in narrative comprehension. *Psychonomic Bulletin & Review*, *10*(1), 104–110.

Huitema, J. S., Dopkins, S., Klin, C. M., & Myers, J. L. (1993). Connecting goals and actions during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(5), 1053–1060.

Hyönä, J., & Lorch, R. F. (2004). Effects of topic headings on text processing: Evidence from adult readers' eye fixation patterns. *Learning and Instruction*, *14*(2), 131–152.

Imhof, M., Echternach, B., Huber, S., & Knorr, S. (1996). Hören und Sehen. Behaltensrelevante Effekte von Illustrationen beim Zuhören. *Unterrichtswissenschaft*, *24*(4), 329–342.

Isberner, M.-B., Richter, T., Maier, J., Knuth-Herzig, K., Horz, H., & Schnotz, W. (2013). Comprehending conflicting science-related texts: graphs as plausibility cues. *Instructional Science*, *41*(5), 849–872.

Jackson, M. D., & McClelland, J. L. (1979). Processing determinants of reading speed. *Journal of Experimental Psychology: General*, *108*(2), 151–181.

Johnson-Laird, P. (1980). Mental models in cognitive science. *Cognitive Science*, *4*(1), 71–115.

Johnson-Laird, P. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness* (No. 6). Cambridge, MA: Harvard University Press.

Johnson-Laird, P. (1987). The comprehension of discourse and mental models. In M. Nagao (Ed.), *Language and artificial intelligence* (pp. 253–261). Amsterdam: North-Holland.

Kail, R. (2000). Speed of information processing: Developmental change and links to intelligence. *Journal of School Psychology*, *38*(1), 51–61.

Kalyuga, S. (2005). Prior knowledge principle in multimedia learning. In R. Mayer (Ed.), *The cambridge handbook of multimedia learning* (pp. 325–337). New York, NY: Cambridge Univ Press.

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist*, *38*(1), 23–31.

Kaup, B., de la Vega, I., Strozyk, J., & Dudschig, C. (2015). The role of sensorimotor processes in meaning composition. In M. Fischer & Y. Coello (Eds.), *Conceptual and interactive embodiment: Foundations of embodied cognition* (pp. 46–70). New York, NY: Routledge.

Keefe, D. E., & McDaniel, M. A. (1993). The time course and durability of predictive inferences. *Journal of Memory and Language*, *32*(4), 446–463.

Kendeou, P., Bohn-Gettler, C., White, M. J., & van den Broek, P. (2008). Children's inference generation across different media. *Journal of Research in Reading*, *31*(3), 259–272.

Kendeou, P., Lynch, J. S., van den Broek, P., Espin, C. A., White, M. J., & Kremer, K. E. (2005). Developing successful readers: Building early comprehension skills

through television viewing and listening. *Early Childhood Education Journal*, *33*(2), 91–98.

Kintsch, W. (1974). *The representation of meaning in memory*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: a construction-integration model. *Psychological Review*, *95*(2), 163.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge university press.

Kintsch, W., Kozminsky, E., Streby, W. J., McKoon, G., & Keenan, J. M. (1975). Comprehension and recall of text as a function of content variables. *Journal of Verbal Learning and Verbal Behavior*, *14*(2), 196–214.

Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, *85*(5), 363–394.

Kintsch, W., & Vipond, D. (1979). Reading comprehension and readability in educational practice and psychological theory. In L. G. Nilsson (Ed.), *Perspectives on memory research* (pp. 329–365). Hillsdale, NJ: Erlbaum.

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

Landerl, K., Wimmer, H., & Frith, U. (1997). The impact of orthographic consistency on dyslexia: A german-english comparison. *Cognition*, *63*(3), 315–334.

Larkin, J. (1983). The role of problem representation in physics. In D. Gentner & A. Stevens (Eds.), *Mental models* (pp. 75–98). Hillsdale, NJ: Lawrence Erlbaum.

Levie, W. H., & Lentz, R. (1982). Effects of text illustrations: A review of research. *ECTJ*, *30*(4), 195–232.

Lorch, R. F. (1989). Text-signaling devices and their effects on reading and memory processes. *Educational Psychology Review*, *1*(3), 209–234.

Lynch, J. S., & van den Broek, P. (2007). Understanding the glue of narrative structure: Children's on-and off-line inferences about characters' goals. *Cognitive Development*, *22*(3), 323–340.

Lynch, J. S., van den Broek, P., Kremer, K. E., Kendeou, P., White, M. J., & Lorch, E. P. (2008). The development of narrative comprehension and its relation to other early reading skills. *Reading Psychology*, *29*(4), 327–365.

MacWhinney, B. (1999). The emergence of language from embodiment. In B. MacWhinney (Ed.), *The emergence of language* (pp. 213–256). Mahwah, NJ: Lawrence Erlabaum.

Magliano, J. P., Baggett, W. B., Johnson, B. K., & Graesser, A. C. (1993). The time course of generating causal antecedent and causal consequence inferences. *Discourse Processes*, *16*(1-2), 35–53.

Maier, J., & Richter, T. (2013). How nonexperts understand conflicting information on social science issues: The role of perceived plausibility and reading goals. *Journal of Media Psychology: Theories, Methods, and Applications*, *25*(1), 14–26.

Marx, H., & Jungmann, T. (2000). Abhängigkeit der Entwicklung des Leseverstehens von Hörverstehen und grundlegenden Lesefertigkeiten im Grundschulalter: eine Prüfung des Simple View of reading-Ansatzes. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *32*(2), 81–93.

Mayer, R. E. (1997). Multimedia learning: Are we asking the right questions? *Educational Psychologist*, *32*(1), 1–19.

Mayer, R. E. (2001). *Multimedia learning*. Cambridge University Press.

Mayer, R. E. (2005a). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The cambridge handbook of multimedia learning* (pp. 31–48). New York, NY: Cambridge Univ Press.

Mayer, R. E. (2005b). Introduction to multimedia learning. In R. E. Mayer (Ed.), *The cambridge handbook of multimedia learning* (pp. 1–16). New York, NY: Cambridge Univ Press.

Mayer, R. E., & Gallini, J. K. (1990). When is an illustration worth ten thousand words? *Journal of Educational Psychology*, *82*(4), 715–726.

McKoon, G., & Ratcliff, R. (1992). Inference during reading. *Psychological Review*, *99*(3), 440–466.

McKoon, G., & Ratcliff, R. (2008). Meanings, propositions, and verbs. *Psychonomic Bulletin & Review*, *15*(3), 592–597.

Michalczyk, K., Malstädt, N., Worgt, M., Könen, T., & Hasselhorn, M. (2013). Age differences and measurement invariance of working memory in 5-to 12-year-old children. *European Journal of Psychological Assessment*, *29*(3).

Morris, N., & Jones, D. M. (1990). Memory updating in working memory: The role of the central executive. *British Journal of Psychology*, *81*(2), 111–121.

Myers, J. L., O'Brien, E. J., Albrecht, J. E., & Mason, R. A. (1994). Maintaining global coherence during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(4), 876–886.

Narvaez, D., Radvansky, G. A., Lynchard, N. A., & Copeland, D. E. (2011). Are older adults more attuned to morally charged information? *Experimental Aging Research*, *37*(4), 398–434.

Nieding, G. (2006). *Wie verstehen Kinder Texte? Die Entwicklung kognitiver Repräsentationen*. Lengerich: Pabst Science Publ.

Nieding, G., & Ohler, P. (2004). Laborexperimentelle Methoden. In R. Mangold, P. Vorderer, & G. Bente (Eds.), *Lehrbuch Medienpsychologie* (p. 355-376). Göttingen: Hogrefe.

Nieding, G., & Ohler, P. (2008). Mediennutzung und Medienwirkung bei Kindern und Jugendlichen. In B. Batinic & M. Appel (Eds.), *Medienpsychologie* (pp. 379–402).

Heidelberg: Springer.

Nieding, G., Ohler, P., Diergarten, A. K., Möckel, T., Rey, G. D., & Schneider, W. (2017). The development of media sign literacy—a longitudinal study with 4-year-old children. *Media Psychology*, *20*(3), 401–427.

Nieding, G., Ohler, P., & Thußbas, C. (1996). The cognitive development of temporal structures: How do children make inferences with temporal ellipses in films? In P. Vorderer, H. J. Wulff, & M. Friedrichsen (Eds.), *Suspense. conceptualizations, theoretical analyses and empirical explorations* (pp. 301–328). Hillsdale, NJ: Lawrence Erlbaum.

Paechter, M. (1997). Auditive und visuelle Texte in Lernsoftware. *Unterrichtswissenschaft*, *3*(25), 223–240.

Paivio, A. (1969). Mental imagery in associative learning and memory. *Psychological Review*, *76*(3), 241.

Paivio, A. (1975). Perceptual comparisons through the mind's eye. *Memory & Cognition*, *3*(6), 635–647.

Paivio, A. (1986). *Mental representations. a dual coding approach.* New York: Oxford University Press.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*, *45*(3), 255–287.

Palladino, P., Cornoldi, C., De Beni, R., & Pazzaglia, F. (2001). Working memory and updating processes in reading comprehension. *Memory & Cognition*, *29*(2), 344–354.

Palmer, J., MacLeod, C. M., Hunt, E., & Davidson, J. E. (1985). Information processing correlates of reading. *Journal of Memory and Language*, *24*(1), 59–88.

Palmer, S. (2000). Working memory: A developmental study of phonological recoding. *Memory*, *8*(3), 179–193.

Peirce, C. (1906). Prolegomena to an apology for pragmatism. *The Monist*, 492–546.

Peirce, J. W. (2007). Psychopy—psychophysics software in python. *Journal of Neuroscience Methods*, *162*(1), 8–13.

Peirce, J. W. (2008). Generating stimuli for neuroscience using psychopy. *Frontiers in Neuroinformatics*, *2*.

Penney, C. G. (1989). Modality effects and the structure of short-term verbal memory. *Memory & Cognition*, *17*(4), 398–422.

Perfetti, C. A., Landi, N., & Oakhill, J. (2005). The acquisition of reading comprehension skill. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook. blackwell handbooks of developmental psychology* (p. 227-247). Malden: Blackwell Publishing.

Pezdek, K., Lehrer, A., & Simon, S. (1984). The relationship between reading and cognitive processing of television and radio. *Child Development*, *55*(6), 2072-2082.

Piaget, J. (1929). *The child's conception of the world*. New York: Harcourt, Brace.

Piaget, J. (1930). *The child's conception of physical causality*. London: Keagan Paul.

Pickering, S., & Gathercole, S. E. (2001). *Working memory test battery for children (wmtb-c)*. Psychological Corporation.

Pollack, I., & Norman, D. A. (1964). A non-parametric analysis of recognition experiments. *Psychonomic Science*, *1*(1-12), 125–126.

Potter, W. J. (1998). *Media literacy* (1st ed.). Thousand Oaks, CA: Sage Publications.

Poynor, D. V., & Morris, R. K. (2003). Inferred goals in narratives: Evidence from self-paced reading, recall, and eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(1), 3–9.

Prior, S. M., Fenwick, K. D., Saunders, K. S., Ouellette, R., O'Quinn, C., & Harvey, S. (2011). Comprehension after oral and silent reading: Does grade level matter? *Literacy Research and Instruction*, *50*(3), 183–194.

Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews Neuroscience*, *6*(7), 576–582.

Radvansky, G. A., Copeland, D. E., Berish, D. E., & Dijkstra, K. (2003). Aging and situation model updating. *Aging, Neuropsychology, and Cognition*, *10*(2), 158–166.

Rall, J., & Harris, P. L. (2000). In cinderella's slippers? story comprehension from the protagonist's point of view. *Developmental Psychology*, *36*(2), 202–208.

Rapp, D. N., van den Broek, P., McMaster, K. L., Kendeou, P., & Espin, C. A. (2007). Higher-order comprehension processes in struggling readers: A perspective for research and intervention. *Scientific Studies of Reading*, *11*(4), 289–312.

Ratcliff, R., & McKoon, G. (1978). Priming in item recognition: The organization of propositions in memory for text. *Journal of Verbal Learning and Verbal Behavior*, *17*(4), 403–417.

Reder, L. M. (1982). Plausibility judgments versus fact retrieval: Alternative strategies for sentence verification. *Psychological Review*, *89*(3), 250–280.

Reyna, V. F., & Kiernan, B. (1994). Development of gist versus verbatim memory in sentence recognition: Effects of lexical familiarity, semantic content, encoding instructions, and retention interval. *Developmental Psychology*, *30*(2), 178–191.

Ricci, C. M., & Beal, C. R. (2002). The effect of interactive media on children's story memory. *Journal of Educational Psychology*, *94*(1), 138–144.

Richter, T., & Christmann, U. (2002). Lesekompetenz: Prozessebenen und interindividuelle Unterschiede. In N. Groeben & B. Hurrelmann (Eds.), *Lesekompetenz: Bedingungen, Dimensionen, Funktionen* (pp. 25–58). Weinheim: Juventa.

Richter, T., & Zwaan, R. A. (2009). Processing of color words activates color representations. *Cognition*, *111*(3), 383–389.

Rideout, V. (2013). *Zero to eight: Children's media use in america.* (Tech. Rep.). San Francisco: Common Sense Media. Retrieved from

https://www.commonsensemedia.org/file/zero-to-eight-2013pdf-0/download

Rieben, L., & Perfetti, C. (1991). *Learning to read: basic research and its implications*. Hillsdale, NJ: Lawrence Erlbaum.

Riedel, R., Büsching, U., & Brand, M. (2017). *Medienstudie BLIKK* (Tech. Rep.). Rheinische Fachhochschule Köln. Retrieved from http://www.rfh-koeln.de/aktuelles/meldungen/2017/medienstudie$_b$likk/index$_g$er.html

Rinck, M. (2000). Situationsmodelle und das Verstehen von Erzähltexten: Befunde und Probleme. *Psychologische Rundschau*, *51*(3), 115–122.

Rizzella, M. L., & O'Brien, E. J. (1996). Accessing global causes during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(5), 1208–1218.

Rommers, J., Meyer, A. S., & Huettig, F. (2013). Object shape and orientation do not routinely influence performance during language processing. *Psychological Science*, *24*(11), 2218–2225.

Rosebrock, C., & Nix, D. (2011). *Grundlagen der Lesedidaktik und der systematischen schulischen Leseförderung*. Hohengehren: Schneider-Verlag.

Rost, D., & Buch, S. (2010). Leseverständnis. In D. Rost (Ed.), *Handwörtbuch Pädagogische Psychologie* (Vol. 4, pp. 507–520). Weinheim: Beltz.

Ruggeri, A., & Katsikopoulos, K. V. (2013). Make your own kinds of cues: When children make more accurate inferences than adults. *Journal of experimental child psychology*, *115*(3), 517–535.

Sachs, J. S. (1967). Recognition memory for syntactic and semantic aspects of connected discourse. *Perception and Psychophysics*, *2*, 437–442.

Sachs, J. S. (1974). Memory in reading and listening to discourse. *Memory & Cognition*, *2*(1), 95–100.

Salomon, G. (1979). Media and symbol systems as related to cognition and learning. *Journal of Educational Psychology*, *71*(2), 131–148.

Salomon, G. (1984). Television is "easy" and print is "tough": The differential investment of mental effort in learning as a function of perceptions and attributions. *Journal of Educational Psychology*, *76*(4), 647-658.

Schmalhofer, F., & Glavanov, D. (1986). Three components of understanding a programmer's manual: Verbatim, propositional, and situational representations. *Journal of Memory and Language*, *25*(3), 279–294.

Schneider, W. (2015). *Memory development from early childhood through emerging adulthood*. Springer International Publishing.

Schneider, W., Körkel, J., & Weinert, F. E. (1989). Domain-specific knowledge and memory performance: A comparison of high- and low-aptitude children. *Journal of Educational Psychology*, *81*(3), 306-312.

Schnotz, W. (1993). On the relation of dual coding and mental models in graphics comprehension. *Learning and Instruction*, *3*(3), 247–249.

Schnotz, W. (1994). *Aufbau von Wissensstrukturen*. Weinheim: Beltz, Psychologie Verlangs Union.

Schnotz, W. (2002). Towards an integrated view of learning from text and visual displays. *Educational Psychology Review*, *14*(1), 101–120.

Schnotz, W. (2005). An integrated model of text and picture comprehension. In R. Mayer (Ed.), *The cambridge handbook of multimedia learning* (pp. 49–69). New York, NY: Cambridge Univ Press.

Schnotz, W. (2014). Integrated model of text and picture comprehension. In R. Mayer (Ed.), *The cambridge handbook of multimedia learning* (pp. 72–103). New York, NY: Cambridge Univ Press.

Schnotz, W., & Baadte, C. (2015). Surface and deep structures in graphics comprehension. *Memory & Cognition*, *43*(4), 605–618.

Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning and Instruction*, *13*(2), 141–156.

Schroeder, S., Würzner, K.-M., Heister, J., Geyken, A., & Kliegl, R. (2015). childlex: A lexical database of german read by children. *Behavior Research Methods*, *47*(4), 1085–1094.

Seger, B., Wannagat, W., & Nieding, G. (in preparation). How pictures and picture animations contribute to multi-level mental representations of auditory text in 7-, 9-, and 11-year-old children.

Seymour, P. H., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in european orthographies. *British Journal of Psychology*, *94*(2), 143–174.

Shankweiler, D. (1999). Words to meanings. *Scientific Studies of Reading*, *3*(2), 112–127.

Singer, M., & Leon, J. (2013). Psychological studies of higher language processes: Behavioral and empirical approaches. In F. Schmalhofer & C. Perfetti (Eds.), *Higher level language processes in the brain: Inference and comprehension* (p. 9-25). New York: Psychology P.

Sodian, B., & Thoermer, C. (2006). Theory of Mind. In W. Schneider & B. Sodian (Eds.), *Enzyklopädie der Psychologie. Serie Entwicklungspsychologie. Band 2: Kognitive Entwicklung* (pp. 495–608). Hogrefe: Hogrefe.

Spearritt, D. (1962). *Listening comprehension: A factorial analysis* (No. 76). Melbourne: Australian Council for Education Research.

Springer, K. (2001). Perceptual boundedness and perceptual support in conceptual development. *Psychological Review*, *108*(4), 691–708.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, *31*(1), 137–149.

Stanovich, K. E. (1991). Discrepancy definitions of reading disability: Has intelligence

led us astray? *Reading Research Quarterly*, *26*(1), 7–29.

Stanovich, K. E., Cunningham, A. E., & Feeman, D. J. (1984). Intelligence, cognitive skills, and early reading progress. *Reading Research Quarterly*, *19*(3), 278–303.

Stanovich, K. E., Nathan, R. G., & Vala-Rossi, M. (1986). Developmental changes in the cognitive correlates of reading ability and the developmental lag hypothesis. *Reading Research Quarterly*, *21*(3), 267–283.

Stanovich, K. E., & Siegel, L. S. (1994). Phenotypic performance profile of children with reading disabilities: A regression-based test of the phonological-core variable-difference model. *Journal of Educational Psychology*, *86*(1), 24–53.

Stanovich, K. E., & West, R. F. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly*, *24*(4), 402-433.

Sticht, T., & James, J. (1984). Listening and reading. In P. Pearson, R. Barr, M. Kamil, & P. Mosenthal (Eds.), *Handbook of reading research* (pp. 293–317). New York: Longman.

Suh, S., & Trabasso, T. (1993). Inferences during reading: Converging evidence from discourse analysis, talk-aloud protocols, and recognition priming. *Journal of Memory and Language*, *32*(3), 279–300.

Tardif, T., & Craik, F. I. (1989). Reading a week later: Perceptual and conceptual factors. *Journal of Memory and Language*, *28*(1), 107–125.

Thompson, J. G., & Myers, N. A. (1985). Inferences and recall at ages four and seven. *Child Development*, *56*(5), 1134–1144.

Tibus, M., Heier, A., & Schwan, S. (2013). Do films make you learn? inference processes in expository film comprehension. *Journal of Educational Psychology*, *105*(2), 329–340.

Trabasso, T., & van den Broek, P. (1985). Causal thinking and the representation of narrative events. *Journal of Memory and Language*, *24*(5), 612–630.

Unsöld, I. (2008). *Die Bildung von Inferenzen bei der kognitiven Verarbeitung medialer Texte: eine Untersuchung an Kindern und Erwachsenen*. Hamburg: Verlag Dr. Kovac.

Unsöld, I., & Nieding, G. (2009). Die Bildung prädiktiver Inferenzen von Kindern und Erwachsenen bei der kognitiven Verarbeitung audiovisueller und auditiver Texte. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *41*(2), 87–95.

van der Schoot, M., Reijntjes, A., & van Lieshout, E. C. (2012). How do children deal with inconsistencies in text? an eye fixation and self-paced reading study in good and poor reading comprehenders. *Reading and Writing*, *25*(7), 1665–1690.

van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.

Vosniadou, S., & Ortony, A. (1989). *Similarity and analogical reasoning*. New York: Cambridge University Press.

Walma van der Molen, J. H., & van der Voort, T. H. (2000). The impact of television,

print, and audio on children's recall of the news. *Human Communication Research*, *26*(1), 3–26.

Wassenburg, S. I., Beker, K., van den Broek, P., & van der Schoot, M. (2015). Children's comprehension monitoring of multiple situational dimensions of a narrative. *Reading and Writing*, *28*(8), 1203–1232.

Whiteman, M. (1967). Children's conceptions of psychological causality. *Child Development*, *38*(1), 143-155.

Whitney, P., Ritchie, B. G., & Clark, M. B. (1991). Working-memory capacity and the use of elaborative inferences in text comprehension. *Discourse Processes*, *14*(2), 133–145.

Wicks, R. H., & Drew, D. G. (1991). Learning from news: Effects of message consistency and medium on recall and inference making. *Journalism & Mass Communication Quarterly*, *68*(1-2), 155–164.

Wrage, H. (2010). Jene Fabrik der Bücher. Über Lesesucht, ein Phantasma des medialen Ursprungs und die Kinder-und Jugendliteratur der Aufklärung. *Monatshefte*, *102*(1), 1–21.

Ziegler, F., Mitchell, P., & Currie, G. (2005). How does narrative cue children's perspective taking? *Developmental Psychology*, *41*(1), 115–123.

Zwaan, R. A. (2004). The immersed experiencer: Toward an embodied theory of language comprehension. In B. Ross (Ed.), *Psychology of learning and motivation* (Vol. 44, pp. 35–62). New York: Academic Press.

Zwaan, R. A. (2014). Embodiment and language comprehension: reframing the discussion. *Trends in Cognitive Sciences*, *18*(5), 229–234.

Zwaan, R. A., Magliano, J. P., & Graesser, A. C. (1995). Dimensions of situation model construction in narrative comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(2), 386–397.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, *123*(2), 162–185.

Zwaan, R. A., & Singer, M. (2003). Text comprehension. In A. Graesser, M. Gernsbacher, & S. Goldman (Eds.), *Handbook of discourse processes* (pp. 83–121). Mahwah, NJ: Erlbaum.

Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science*, *13*(2), 168–171.

Zwaan, R. A., & Taylor, L. J. (2006). Seeing, acting, understanding: motor resonance in language comprehension. *Journal of Experimental Psychology: General*, *135*(1), 1–11.

Zwaan, R. A., Taylor, L. J., & de Boer, M. (2010). Motor resonance as a function of narrative time: Further tests of the linguistic focus hypothesis. *Brain and Language*, *112*(3), 143–149.

List of Figures

List of Tables

Appendix

A Materials Studies I and II

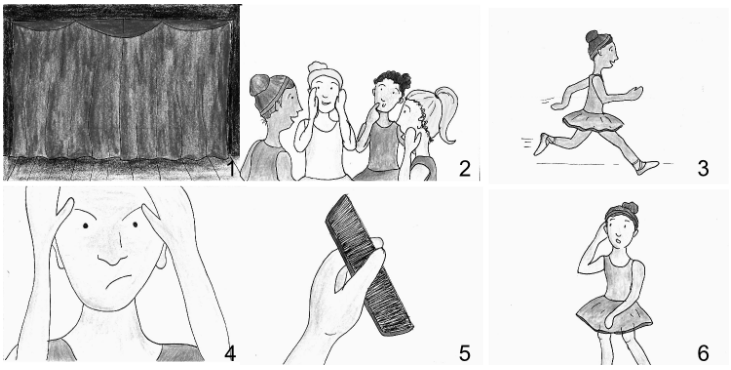
Practice Story

Anne-Marie freut sich, denn ihre große Schwester spielt mit ihr.
Zuerst verkleiden sie sich mit den Kleidern ihrer Mutter und schminken sich vor dem großen Spiegel.
Danach gehen sie raus in den Garten und schaukeln ganz wild auf der Schaukel.

Distractors

Anne-Marie freut sich, denn ihre Cousine spielt heute mit ihr.
Danach gehen sie raus in den Garten und schaukeln ganz wild auf der Schaukel.
Zuerst verkleiden sie sich mit den Klamotten ihrer Mutter und schminken sich vor dem großen Spiegel.

Note: Colored versions of the following pictures have been used in the experiments.



1. Anna und das Ballett

Original
Kurz vor dem Auftritt der Ballettgruppe ist der Vorhang noch zu und Anna und ihre Freunde laufen hinter der Bühne wild durcheinander.
Jeder geht noch einmal seine Rolle durch und alle sind sehr aufgeregt.
Anna rennt schnell noch einmal in den Flur und sieht ein letztes Mal in den Spiegel.
Sie bemerkt, dass sich eine Haarsträhne gelöst hat und ärgert sich deswegen sehr.
Sie versucht zu retten, was zu retten ist und nimmt schnell einen Kamm zur Hand, aber es ist zu spät.
Sie hört ihre Musik und so tritt sie wohl oder übel hinaus ins Scheinwerferlicht.

Text Surface Changes
Kurz bevor die Ballettgruppe auftritt, ist der Vorhang noch zu und Anna und ihre Freunde laufen hinter der Bühne wild durcheinander.
Jeder wiederholt noch einmal seine Rolle und alle sind sehr aufgeregt.
Anna rennt schnell noch einmal in den Gang und schaut ein letztes Mal in den Spiegel.
Sie sieht, dass sich eine Haarsträhne gelockert hat und ärgert sich daher sehr.
Sie probiert zu retten, was zu retten ist und nimmt schnell einen Kamm zur Hand, aber es ist zu spät.
Sie hört ihre Musik und so geht sie wohl oder übel hinaus ins Rampenlicht.

Text Base Changes
Kurz vor dem Auftritt der Ballettgruppe ist der Vorhang noch zu und Anna und die anderen laufen hinter der Bühne wild durcheinander.
Jeder geht noch einmal seine Tanzschritte durch und alle sind sehr aufgeregt. Anna rennt schnell noch einmal in den Flur und prüft ein letztes Mal ihr Spiegelbild.
Sie bemerkt, dass sich ihre Frisur gelöst hat und ärgert sich deswegen sehr.
Sie versucht zu retten, was zu retten ist und nimmt schnell einen Kamm zur Hand, aber da muss sie los.
Sie hört ihre Musik und so tritt sie wohl oder übel hinaus auf die Bühne.

Situation Changes
Kurz vor der Probe der Ballettgruppe ist der Vorhang noch zu und Anna und ihre Freunde laufen hinter der Bühne wild durcheinander.
Jeder geht noch einmal seine Texte durch und alle sind sehr aufgeregt.
Anna rennt schnell noch einmal in den Flur und sieht zum vorletzten Mal in den Spiegel.
Sie bemerkt, dass sich der Haarschmuck gelöst hat und ärgert sich deswegen sehr.
Sie versucht zu retten, was zu retten ist und nimmt schnell einen Kamm zur Hand, aber es ist zu schwer.
Sie hört ihre Musik und so tritt sie wohl oder übel in die Garderobe.

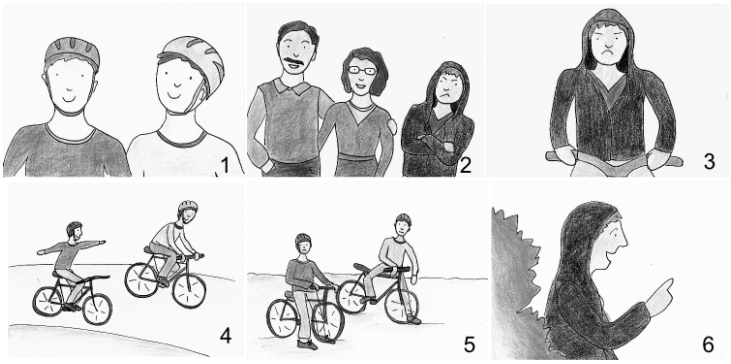
2. Im Stadion

Original
Gespannt sitzen Basti und sein Vater auf der Tribüne, während sich die Fußballspieler auf dem Rasen warm machen.
Fast alle Zuschauer tragen schwarz-gelbe Schals und auch Basti bindet sich den Schal seiner Lieblingsmannschaft um.
Als er sich genauer umschaut, entdeckt er nur einige Fans der anderen Mannschaft.
Vom lauten Anfeuern bekommt er einen ganz trockenen Hals und so reicht ihm der Vater einen Becher Saft.
Auch der Vater hat schon einen ganz roten Kopf, weil heute so ein wichtiger Tag ist.
Der Anpfiff ertönt und Basti drückt seiner Mannschaft ganz fest die Daumen.

Text Surface Changes
Gespannt sitzen Basti und sein Vater auf der Tribüne, während sich die Fußballer auf dem Rasen aufwärmen.
Fast jeder Zuschauer trägt einen schwarz-gelben Schal und auch Basti bindet sich den Schal seiner Lieblingsmannschaft um.
Als er sich genauer umschaut, entdeckt er nur einige Anhänger der anderen Mannschaft.
Vom lauten Anfeuern bekommt er einen ganz trockenen Hals und so gibt der Vater ihm einen Becher Saft.
Auch der Vater hat schon einen ganz roten Kopf, weil der heutige Tag so wichtig ist.
Der Anpfiff erfolgt und Basti drückt seiner Mannschaft ganz fest die Daumen.

Text Base Changes
Gespannt sitzen Basti und sein Vater auf der Tribüne, während sich die Fußballspieler auf dem Rasen warm laufen.
Fast alle Zuschauer haben schwarz-gelbe Schals besorgt und auch Basti bindet sich den Schal seiner Lieblingsmannschaft um.
Als er sich genauer umschaut, entdeckt er nur einige Fans in anderen Farben.
Vom lauten Anfeuern bekommt er einen ganz trockenen Hals und so reicht ihm der Vater etwas zu trinken.
Auch der Vater hat schon einen ganz roten Kopf, weil heute so ein wichtiges Spiel ist.
Das Spiel beginnt und Basti drückt seiner Mannschaft ganz fest die Daumen.

Situation Changes
Gespannt sitzen Basti und sein Vater auf der Tribüne, während sich die Fußballspieler in der Kabine warm machen.
Fast alle Spieler tragen schwarz-gelbe Schals und auch Basti bindet sich den Schal seiner Lieblingsmannschaft um.
Als er sich genauer umschaut, entdeckt er sehr viele Fans der anderen Mannschaft.
Vom lauten Anfeuern bekommt er einen ganz trockenen Hals und so reicht ihm der Vater einen Becher Eis.
Auch der Vater hat schon einen ganz roten Kopf, weil heute so ein sonniger Tag ist.
Der Gong ertönt und Basti drückt seiner Mannschaft ganz fest die Daumen.



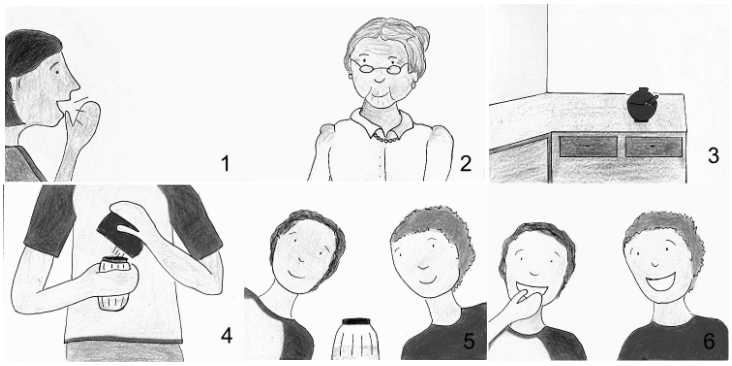
3. Die Fahrradtour

Original
Daniel und Tobi haben schon die Helme auf, denn sie möchten ihre neuen Fahrräder ausprobieren.
Die Eltern bitten ihren großen Bruder Karsten, die Brüder zu begleiten, obwohl er Schwimmtraining hat.
So fahren die drei los und während Karsten schlecht gelaunt voraus rast, haben Daniel und Tobi viel Spaß.
Als Daniel Tobi zeigt, wie gut er schon freihändig fahren kann, bemerken sie plötzlich, dass Karsten verschwunden ist.
Besorgt halten sie an und rufen laut nach ihm.
Als sie schon kurz vorm Weinen sind, springt der gemeine Karsten hinter einem Busch hervor und lacht die beiden aus.

Text Surface Changes
Daniel und Tobi haben schon die Helme auf, denn sie möchten ihre neuen Fahrräder ausprobieren.
Die Eltern bitten ihren großen Bruder Karsten, die Brüder zu begleiten, obwohl er Schwimmtraining hat.
So fahren die drei los und während Karsten schlecht gelaunt voraus rast, haben Daniel und Tobi viel Spaß.
Als Daniel Tobi zeigt, wie gut er schon freihändig fahren kann, bemerken sie plötzlich, dass Karsten verschwunden ist.
Besorgt halten sie an und rufen laut nach ihm.
Als sie schon kurz vorm Weinen sind, springt der gemeine Karsten hinter einem Busch hervor und lacht die beiden aus.

Text Base Changes
Daniel und Tobi haben schon die Helme auf, denn sie möchten mit ihren neuen Fahrrädern fahren.
Die Eltern bitten ihren großen Bruder Karsten, die Brüder zu begleiten, obwohl er keine Zeit hat.
So beginnt die Fahrradtour und während Karsten schlecht gelaunt voraus rast, haben Daniel und Tobi viel Spaß.
Als Daniel Tobi zeigt, wie gut er schon freihändig fahren kann, bemerken sie plötzlich, dass Karsten nicht mehr zu sehen ist.
Besorgt halten sie an und suchen nach ihm.
Als sie schon kurz vorm Verzweifeln sind, springt der gemeine Karsten hinter einem Busch hervor und lacht die beiden aus.

Situation Changes
Daniel und Tobi haben schon die Helme auf, denn sie möchten ihre neuen Fahrräder putzen.
Die Eltern bitten ihren großen Bruder Karsten, die Brüder zu begleiten, obwohl er Kopfschmerzen hat.
So fahren die drei los und während Karsten schlecht gelaunt voraus rast, haben Daniel und Tobi viel zu tun.
Als Daniel Tobi zeigt, wie gut er schon freihändig fahren kann, bemerken sie plötzlich, dass Karsten müde ist.
Besorgt halten sie an und rufen ihn an.
Als sie schon kurz vorm Umkehren sind, springt der gemeine Karsten hinter einem Busch hervor und lacht die beiden aus.



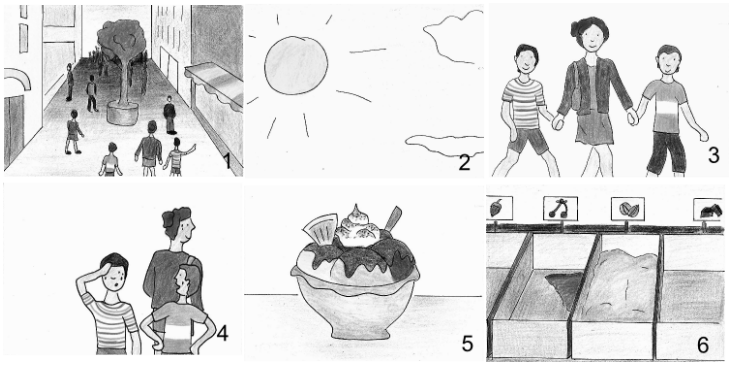
4. Beim Essen

Original
Die Mutter ruft nach Max und Sascha, doch eigentlich wollen die beiden gerade ihre Lieblingsserie anschauen.
Die Oma kommt zum Mittagessen vorbei und Max und Sascha sollen deshalb den Tisch decken.
Als sie widerwillig Teller und Gläser aus der Küche herbeitragen, sieht Max auf einmal eine rote Zuckerdose auf dem Schränkchen stehen.
Grinsend zwinkert er Sascha zu und schüttet dann den Zucker aus der roten Dose in den Salzstreuer hinein.
Bald sitzen alle am Tisch und Max und Sascha lassen den Salzstreuer nicht aus den Augen.
Als der Vater dann endlich den Salzstreuer nimmt, können die beiden ihr Lachen nicht mehr unterdrücken.

Text Surface Changes
Die Mutter ruft nach Max und Sascha, doch eigentlich wollen die beiden gerade ihre Lieblingsserie ankucken.
Die Oma kommt zum Mittagessen vorbei und Max und Sascha müssen deshalb den Mittagstisch decken.
Als sie widerstrebend Teller und Gläser aus der Küche herbeibringen, sieht Max auf einmal eine rote Zuckerdose auf dem Schränkchen stehen.
Grinsend zwinkert er seinem Bruder zu und schüttet dann den Zucker aus der roten Dose in den Salzstreuer hinein.
Bald sitzt jeder am Esstisch und Max und Sascha lassen den Salzstreuer nicht aus den Augen.
Als der Vater dann endlich nach dem Salzstreuer greift, können die beiden ihr Lachen nicht mehr unterdrücken.

Text Base Changes
Die Mutter ruft nach Max und Sascha, doch eigentlich wollen die beiden gerade lieber Fernseh schauen.
Die Oma kommt zum Mittagessen vorbei und die beiden sollen deshalb mithelfen.
Als sie widerwillig Geschirr aus der Küche herbeitragen, sieht Max auf einmal eine rote Zuckerdose auf dem Schränkchen stehen.
Begeistert zwinkert er Sascha zu und schüttet dann den Zucker aus der roten Dose in den Salzstreuer hinein.
Bald sitzen alle beim Essen und Max und Sascha lassen den Salzstreuer nicht aus den Augen.
Als der Vater dann endlich nachsalzen will, können die beiden ihr Lachen nicht mehr unterdrücken.

Situation Changes
Die Mutter ruft nach Max und Sascha, doch eigentlich wollen die beiden gerade ihr Lieblingsbuch anschauen.
Die Oma kommt zum Frühstück vorbei und Max und Sascha sollen deshalb den Tisch decken.
Als sie eifrig Teller und Gläser aus der Küche herbeitragen, sieht Max auf einmal eine rote Zuckerdose auf dem Schränkchen stehen.
Grinsend winkt er Sascha zu und schüttet dann den Zucker aus der roten Dose in den Salzstreuer hinein.
Bald stehen alle am Buffet und Max und Sascha lassen den Salzstreuer nicht aus den Augen.
Als die Oma dann endlich den Salzstreuer nimmt, können die beiden ihr Lachen nicht mehr unterdrücken.



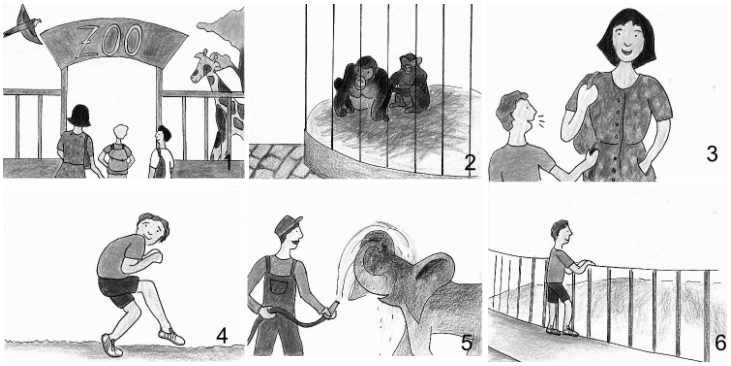
5. Im der Stadt

Original
Fabian und Tim sind mit ihrer Mutter in der Stadt, weil sie neue Schuhe brauchen.
Aber da die Sonne heiß vom Himmel scheint, wollen die beiden lieber Eis essen.
Sie fragen ihre Mutter so lange, bis sie schließlich einwilligt und sie zum Eiscafé gehen.
Schon von weitem sehen sie eine lange Schlange vor dem Eiscafé und schwitzend stellen sie sich an.
Während sie in der Schlange stehen, erzählen sich Fabian und Tim, welche riesigen Eisbecher sie gleich essen werden.
Als sie dann dran sind, gibt es nur noch Pistazieneis, genau das Eis, das beiden überhaupt nicht schmeckt.

Text Surface Changes
Fabian und Tim sind mit ihrer Mutter in der Stadt, weil sie neue Schuhe benötigen.
Aber da die Sonne heiß vom Himmel scheint, wollen die beiden lieber Eis schlecken.
Sie fragen ihre Mutter so lange, bis sie endlich ja sagt und sie zum Eiscafé gehen.
Schon aus der Ferne sehen sie eine lange Schlange vor der Eisdiele und schwitzend stellen sie sich an.
Während sie in der Reihe anstehen, erzählen sich Fabian und Tim, welche riesigen Eisbecher sie gleich essen werden.
Als sie dann dran sind, gibt es nur noch Pistazieneis, genau das Eis, das beide gar nicht mögen.

Text Base Changes
Fabian und Tim sind mit ihrer Mutter in der Stadt, um neue Schuhe zu kaufen.
Aber da die Sonne heiß vom Himmel scheint, wollen die beiden lieber Eis holen.
Sie nerven ihre Mutter so lange, bis sie schließlich einwilligt und sie zum Eiscafé gehen.
Schon von weitem sehen sie viele Menschen vor dem Eiscafé und schwitzend stellen sie sich an.
Während sie.... warten, erzählen sich Fabian und Tim, welche riesigen Eisbecher sie gleich essen werden,
Als sie dann dran sind, gibt es nur noch Pistazieneis, genau das Eis, das beide überhaupt nicht wollen.

Situation Changes
Fabian und Tim sind mit ihrer Mutter in der Stadt, um neue Schuhe zu kaufen.
Aber da die Sonne heiß vom Himmel scheint, wollen die beiden lieber Eis holen.
Sie nerven ihre Mutter so lange, bis sie schließlich einwilligt und sie zum Eiscafé gehen.
Schon von weitem sehen sie viele Menschen vor dem Eiscafé und schwitzend stellen sie sich an.
Während sie.... warten, erzählen sich Fabian und Tim, welche riesigen Eisbecher sie gleich essen werden.
Als sie dann dran sind, gibt es nur noch Pistazieneis, genau das Eis, das beide überhaupt nicht wollen.



6. Im Zoo

Original
Die Klasse macht einen Ausflug in den Zoo und Felix ist besonders auf die Elefanten gespannt.
Zuerst gehen sie zu den Affen und beobachten fasziniert, wie die Schimpansen sich lausen.
Schon bald wird Felix ungeduldig und fragt die Lehrerin, wann sie denn zu den Elefanten gehen.
Da der Rest der Klasse aber noch die Affen fotografieren möchte, schleicht Felix sich heimlich davon.
Kurz darauf steht er endlich vor dem Elefantengehege, in dem der Wärter gerade einen Babyelefanten abduscht.
Zufrieden schaut Felix zu und hofft, dass die Lehrerin seine Flucht nicht bemerkt.

Text Surface Changes
Die Schulklasse unternimmt einen Ausflug in den Zoo und Felix ist vor allem auf die Elefanten gespannt.
Zunächst gehen sie zu den Affen und beobachten begeistert, wie die Schimpansen sich lausen.
Bald schon wird Felix unruhig und fragt die Lehrerin, wann sie denn zu den Elefanten gehen.
Da die übrige Klasse jedoch noch die Affen fotografieren möchte, schleicht Felix sich heimlich davon.
Kurze Zeit später ist er endlich vor dem Elefantengehege, in dem der Wärter gerade einen Babyelefanten abduscht.
Zufrieden schaut Felix zu und hofft, dass seine Flucht von der Lehrerin nicht bemerkt wird.

Text Base Changes
Die Klasse macht einen Ausflug in den Zoo und Felix freut sich besonders auf die Elefanten.
Zuerst schauen sie sich die Affen an und beobachten fasziniert, wie die Schimpansen sich lausen.
Schon bald wird Felix ungeduldig und fragt die Lehrerin, wann sie denn weitergehen.
Da der Rest der Klasse aber noch bei den Affen bleiben möchte, schleicht Felix sich heimlich davon.
Kurz darauf steht er vor dem lange erwarteten Elefantengehege, in dem der Wärter gerade einen Babyelefanten abduscht.
Zufrieden schaut Felix zu und hofft, dass die Lehrerin sein Fehlen nicht bemerkt.

Situation Changes
Die Klasse macht einen Ausflug in den Zoo und alle sind besonders auf die Elefanten gespannt.
Zuerst fahren sie zu den Affen und beobachten fasziniert, wie die Schimpansen sich lausen.
Schon bald wird Felix müde und fragt die Lehrerin, wann sie denn zu den Elefanten gehen.
Da der Rest der Klasse aber noch die Affen füttern möchte, schleicht Felix sich heimlich davon.
Nach langem Suchen steht er endlich vor dem Elefantengehege, in dem der Wärter gerade einen Babyelefanten abduscht.
Zufrieden schaut Felix zu und hofft, dass die Klassenkameraden seine Flucht nicht bemerken.

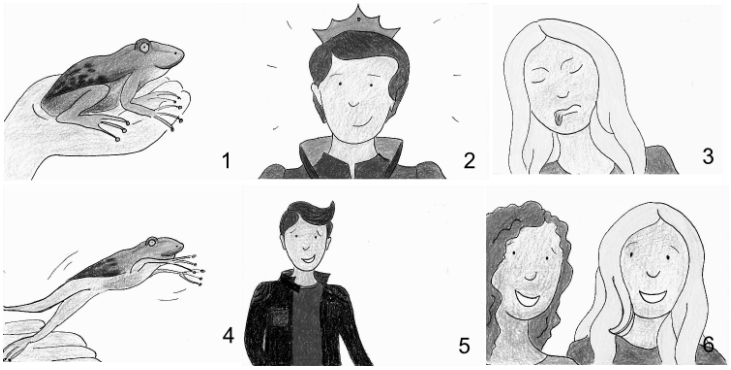
7. Reisevorbereitungen mit Hindernissen

Original
Morgen ganz früh soll es für Familie Schmidt endlich mit dem vollgepackten Auto in den Urlaub gehen.
Als Florian die letzten Koffer verlädt, entdeckt er überrascht, dass der rechte Hinterreifen des Autos ein Loch hat.
Er ist ganz traurig, weil er sich schon so arg auf den Urlaub gefreut hat.
Schnell holt Florian seinen Vater, damit der den Reifen reparieren kann.
Der Vater hebt mit seinem Wagenheber das Auto hoch und wechselt den Reifen.
Damit kann morgen der Urlaub losgehen und Florian ist sehr glücklich.

Text Surface Changes
Am nächsten Tag ganz zeitig soll es für Familie Schmidt mit dem vollgepackten Auto endlich in den Urlaub gehen.
Als Florian die letzten Koffer verlädt, entdeckt er verblüfft, dass der rechte Hinterreifen des Autos ein Loch hat.
Er ist ganz traurig, weil er sich auf den Urlaub bereits so sehr gefreut hat.
Schnell holt Florian seinen Vater, damit der den Reifen in Ordnung bringen kann.
Der Vater hebt mit seinem Wagenheber das Auto hoch und tauscht den Reifen aus.
Damit kann der Urlaub morgen beginnen und Florian ist sehr glücklich.

Text Base Changes
Morgen ganz früh soll es für Familie Schmidt endlich mit dem vollgepackten Auto auf die Reise gehen.
Als Florian die letzten Koffer verlädt, entdeckt er überrascht, dass der rechte Hinterreifen des Autos keine Luft hat.
Er ist ganz traurig, weil er sich schon so arg auf die Reise gefreut hat.
Schnell holt Florian seinen Vater, damit der das Auto reparieren kann.
Der Vater hebt mit seinem Wagenheber das Auto hoch und behebt den Schaden.
Damit kann morgen die Familie losfahren und Florian ist sehr glücklich.

Situation Changes
Übermorgen ganz früh soll es für Familie Schmidt endlich mit dem vollgepackten Auto in den Urlaub gehen.
Als Florian die letzten Koffer verlädt, entdeckt er überrascht, dass der rechte Hinterreifen des Autos weg ist.
Er ist ganz traurig, weil er sich schon so arg auf die Freunde gefreut hat.
Schnell holt Florian seinen Vater, damit der einen Reifen kaufen kann.
Der Vater hebt mit seinem Wagenheber das Auto hoch und wechselt die Bremsen.
Damit kann morgen die Klassenfahrt losgehen und Florian ist sehr glücklich.



8. Das Froschabenteuer

Original
Leonie und Sophia sind im Garten und haben am Teich einen Frosch gefangen.
Leonie hält Sophia den Frosch entgegen und erklärt, dass Frösche verwunschene Prinzen sind und man sie küssen muss, damit sie sich zurückverwandeln.
Angeekelt schließt Sophia die Augen, kommt dem Frosch näher und küsst ihn kurz.
Aber nichts passiert, nur der arme Frosch hüpft zurück in den Teich.
Da raschelt es auf einmal in der Hecke und ein hübscher Junge tritt hervor.
Es ist Moritz, der Sohn der Nachbarn, und die Mädchen lächeln, denn der Kuss war nicht umsonst.

Text Surface Changes
Leonie und Sophia befinden sich im Garten und haben am Teich einen Frosch gefangen.
Leonie hält Sophia den Frosch entgegen und erklärt, dass Frösche verwunschene Prinzen sind, denen man einen Kuss geben muss, damit sie sich zurückverwandeln.
Angeekelt schließt Sophia die Augen, nähert sich dem Frosch und küsst ihn kurz.
Aber nichts geschieht, nur der arme Frosch hüpft zurück in den Teich.
Da raschelt es auf einmal im Gebüsch und ein hübscher Junge tritt hervor.
Es ist Moritz, der Nachbarssohn, und die Mädchen lächeln, denn der Kuss war nicht umsonst.

Text Base Changes
Leonie und Sophia spielen im Garten und haben am Teich einen Frosch gefangen.
Leonie hält Sophia den Frosch entgegen und erklärt, dass Frösche verwunschene Prinzen sind und man sie küssen muss, damit sie wieder Prinzen werden.
Angeekelt schließt Sophia die Augen, kommt dem Frosch näher und küsst ihn hastig.
Aber kein Prinz erscheint, nur der arme Frosch hüpft zurück in den Teich.
Da bewegt sich auf einmal etwas in der Hecke und ein hübscher Junge tritt hervor.
Es ist Moritz, der nebenan wohnt, und die Mädchen lächeln, denn der Kuss war nicht umsonst.

Situation Changes
Leonie und Sophia sind im Garten und haben am Bach einen Frosch gefangen.
Leonie hält Sophia den Frosch entgegen und erklärt, dass Frösche verwunschene Prinzen sind und man sie streicheln muss, damit sie sich zurückverwandeln.
Angeekelt schließt Sophia die Augen, kommt dem Frosch näher und streichelt ihn kurz.
Aber nichts passiert, nur der arme Frosch hüpft zurück in den Bach.
Da pfeift es auf einmal in der Hecke und ein hübscher Junge tritt hervor.
Es ist Moritz, der Enkel der Nachbarn, und die Mädchen lächeln, denn der Kuss war nicht umsonst.

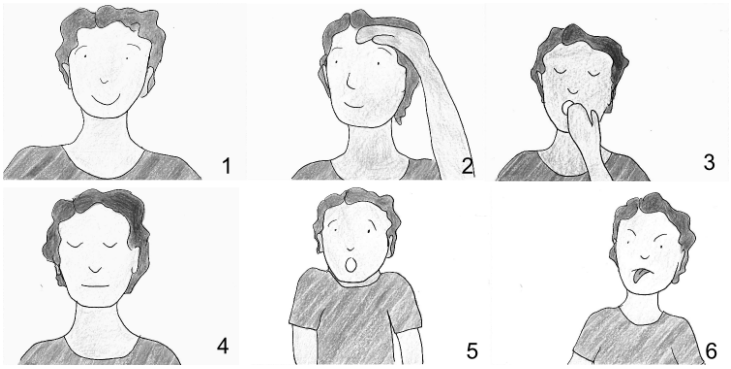
9. Hannah geht Einkaufen

Original
Auf dem Weg zum Supermarkt bemerkt Hannah, dass sie ihren Geldbeutel verloren hat.
Sie hat Angst, dass ihre Mutter deswegen sehr böse sein wird.
So kehrt Hannah um, sucht den Gehweg ab und fängt sogar an zu weinen.
Sie schaut auch unter die Büsche, aber außer viel Müll kann sie nichts finden.
Als sie schon aufgeben will, sieht sie auf einmal ihre Mutter mit dem Geldbeutel auf sie zu rennen.
Lächelnd erzählt ihr die Mutter, dass der Geldbeutel zu Hause auf Hannahs Tisch lag.

Text Surface Changes
Auf dem Weg zum Supermarkt bemerkt Hannah, dass sie ihre Geldbörse verloren hat.
Sie hat Angst, dass ihre Mutter darum sehr wütend sein wird.
So dreht Hannah um, sucht den Bürgersteig ab und beginnt sogar zu weinen.
Sie schaut auch unter die Büsche, aber außer viel Abfall kann sie nichts entdecken.
Als sie bereits aufgeben möchte, sieht sie auf einmal ihre Mutter mit dem Geldbeutel auf sie zu rennen.
Lächelnd erzählt ihr die Mutter, dass der Geldbeutel daheim auf Hannahs Tisch lag.

Text Base Changes
Auf dem Weg zum Supermarkt bemerkt Hannah, dass ihr Geldbeutel nicht mehr da ist.
Sie hat Angst, dass ihre Mutter deswegen sehr schimpfen wird.
So kehrt Hannah um, schaut genau auf den Gehweg und fängt sogar an zu weinen.
Sie schaut auch unter die Büsche, aber den Geldbeutel kann sie nicht finden.
Als sie schon nicht mehr weitersuchen will, sieht sie auf einmal ihre Mutter mit dem Geldbeutel auf sie zu rennen.
Lächelnd erzählt ihr die Mutter, dass der Geldbeutel zu Hause in Hannahs Zimmer lag.

Situation Changes
Auf dem Weg zum Supermarkt bemerkt Hannah, dass ihr Geldbeutel geklaut wurde.
Sie hat Angst, dass ihre Mutter deswegen sehr traurig sein wird.
So kehrt Hannah um, sucht den Parkplatz ab und fängt sogar an zu weinen.
Sie schaut auch unter die Büsche, aber außer viel Löwenzahn kann sie nichts finden.
Als sie schon ausrasten will, sieht sie auf einmal ihre Mutter mit dem Geldbeutel auf sie zu rennen.
Lächelnd erzählt ihr die Mutter, dass der Geldbeutel zu Hause unter Hannahs Tisch lag.



10. Das Versteckspiel

Original
Julius steht ganz still hinter einem großen Busch und lächelt.
Schon lange beobachtet er seinen Bruder, der durch den Garten läuft und ihn sucht.
Langsam wird er in seinem Versteck richtig müde und so legt er sich ins Gras.
Die Sonne scheint ihm ins Gesicht und allmählich fallen ihm die Augen zu.
Auf einmal ertönt neben ihm ein ganz lautes Pfeifen und er schreckt hoch.
Grinsend steht sein gemeiner Bruder Benny mit seiner Trillerpfeife da und Julius streckt ihm ärgerlich die Zunge raus.

Text Surface Changes
Julius steht ganz ruhig hinter einem großen Busch und lächelt.
Schon seit langem beobachtet er seinen Bruder, der durch den Garten läuft und nach ihm sucht.
Langsam wird er in seinem Versteck richtig müde und so legt er sich im Gras nieder.
Die Sonne strahlt auf sein Gesicht und allmählich fallen ihm die Augen zu.
Auf einmal hört er neben sich ein ganz lautes Pfeifen und er schreckt hoch.
Grinsend steht sein fieser Bruder Benny mit der Trillerpfeife da und Julius streckt ihm ärgerlich die Zunge raus.

Text Base Changes
Julius versteckt sich hinter einem großen Busch und lächelt.
Schon lange beobachtet er seinen Bruder, der über den Rasen läuft und ihn sucht.
Langsam wird er in seinem Versteck richtig müde und so legt er sich auf den Boden.
Die Sonne wärmt ihm das Gesicht und allmählich fallen ihm die Augen zu.
Auf einmal ertönt neben ihm ein ganz unangenehmes Pfeifen und er schreckt hoch.
Schadenfroh steht sein gemeiner Bruder Benny mit seiner Trillerpfeife da und Julius streckt ihm ärgerlich die Zunge raus.

Situation Changes
Julius steht ganz still hinter einem großen Holzstoß und lächelt.
Schon lange beobachtet er seine Schwester, die durch den Garten läuft und ihn sucht.
Langsam wird er in seinem Versteck richtig müde und so setzt er sich ins Gras.
Die Sonne scheint ihm auf den Rücken und allmählich fallen ihm die Augen zu.
Auf einmal ertönt neben ihm ein ganz lautes Gebrüll und er schreckt hoch.
Grinsend steht sein gemeiner Bruder Benny mit seiner Trompete da und Julius streckt ihm ärgerlich die Zunge raus.

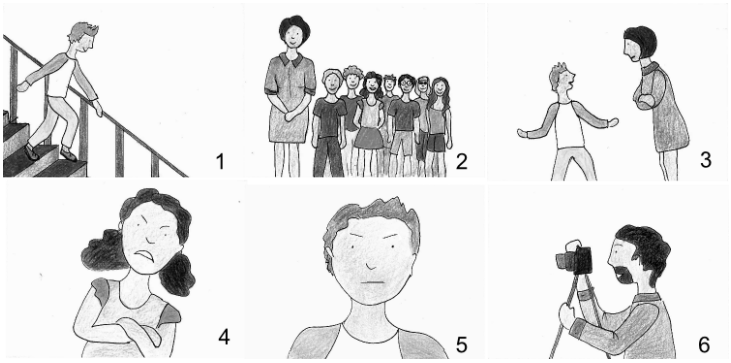
11. Auf dem Bauernhof

Original
Kathrin und ihre Eltern fahren in den Ferien zu ihrer Lieblingstante Inge auf den Bauernhof.
Das Schwein Bertha hat sechs kleine Ferkel bekommen und Kathrin ist schon sehr gespannt, sie zu sehen.
Als sie am Hof ankommen, springt sie aus dem Auto und will schnell in den Stall.
Ihre Tante führt sie zu den Ferkeln und legt ihr eins in den Arm.
Die Tante sagt, dass Kathrin einen Namen für das Ferkel aussuchen darf und sie denkt ganz genau nach.
Grinsend erklärt sie dann, dass das Ferkel ein bisschen ihrem Bruder Fred ähnlich sieht und deshalb genauso heißen soll.

Text Surface Changes
In den Ferien fahren Kathrin und ihre Eltern zu ihrer liebsten Tante Inge auf den Bauernhof.
Das Schwein Bertha hat sechs winzige Ferkel bekommen und Kathrin ist schon total gespannt, sie zu sehen.
Als sie den Hof erreichen, springt sie aus dem Auto und will rasch in den Stall.
Ihre Tante bringt sie zu den Ferkeln und gibt ihr eins in den Arm.
Die Tante sagt, dass Kathrin einen Namen für das Ferkel auswählen darf und sie überlegt ganz genau.
Grinsend erklärt sie dann, dass das Ferkel ein bisschen wie ihr Bruder Fred aussieht und es deshalb auch so heißen soll.

Text Base Changes
Kathrin und ihre Eltern verbringen die Ferien bei ihrer Lieblingstante Inge auf dem Bauernhof.
Das Schwein Bertha hat sechs kleine Ferkel bekommen und Kathrin kann es kaum erwarten, sie zu sehen.
Als sie am Hof ankommen, springt sie aus dem Auto und will schnell zu den Ferkeln.
Ihre Tante zeigt ihr die Ferkel und legt ihr eins in den Arm.
Die Tante sagt, dass Kathrin einen Namen für das Ferkel aussuchen darf und sie denkt lange nach.
Grinsend erklärt sie dann, dass das Ferkel sie ein bisschen an ihren Bruder Fred erinnert und deshalb genauso heißen soll.

Situation Changes
Kathrin und ihre Eltern fahren übers Wochenende zu ihrer Lieblingstante Inge auf den Bauernhof.
Das Schwein Bertha hat sechs kleine Ferkel bekommen und Kathrin ist schon etwas ängstlich sie zu sehen.
Als sie am Hof ankommen, springt sie aus dem Auto und will schnell ins Haus.
Ihre Tante ist schon bei den Ferkeln und legt ihr eins in den Arm.
Die Tante sagt, dass Kathrin sich ein Ferkel aussuchen darf und sie denkt ganz genau nach.
Grinsend erklärt sie dann, dass das Ferkel ein bisschen ihrem Vater Fred ähnlich sieht und deshalb genauso heißen soll.



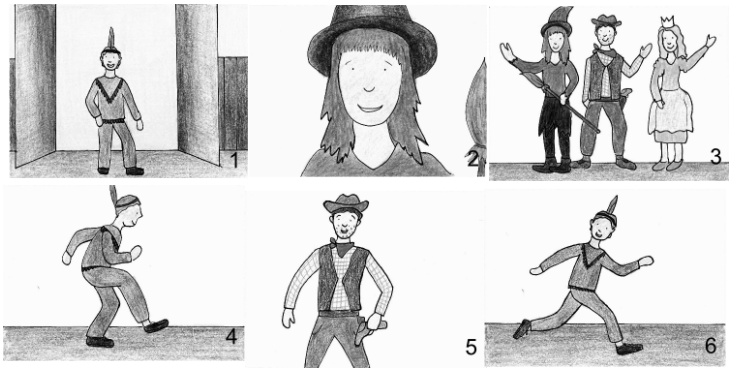
12. Simon ist zu spät

Original
Heute wird ein Klassenfoto gemacht und so hüpft Simon die Stufen hinunter und rennt auf den Schulhof.
Da steht nämlich schon seine ganze Klasse bereit und wartet nur auf ihn.
Simon entschuldigt sich bei der Lehrerin und stellt sich schnell neben Martin in die erste Reihe.
Doch da beschwert sich die Nervensäge Sarah, weil sie hinter Simon gar nicht zu sehen ist.
Wie immer muss Simon also in die hinterste Reihe und ist deshalb sehr sauer.
Er ist darum auch der einzige, der nicht lächelt, als der Fotograf dann das Foto macht.

Text Surface Changes
Heute wird die Klasse fotografiert und so hüpft Simon die Stufen hinunter und rennt auf den Schulhof.
Da steht nämlich schon seine ganze Klasse bereit und wartet bloß noch auf ihn.
Simon entschuldigt sich bei der Lehrerin und stellt sich schnell an Martins Seite in die erste Reihe.
Doch da beschwert sich die Nervensäge Sarah, weil man sie hinter Simon gar nicht sehen kann.
Wie jedes Mal muss Simon also in die letzte Reihe und ist deshalb sehr sauer.
Er ist darum auch die einzige Person, die nicht lächelt, als der Fotograf dann das Foto macht.

Text Base Changes
Heute wird ein Gruppenfoto gemacht und so hüpft Simon die Stufen hinunter und rennt auf den Schulhof.
Da steht nämlich schon seine ganze Klasse bereit und wartet ungeduldig auf ihn.
Simon entschuldigt sich bei der Lehrerin und stellt sich schnell neben Martin nach vorne.
Doch da beschwert sich die Nervensäge Sarah, weil Simon zu groß ist.
Wie schon letztes Mal muss Simon also in die hinterste Reihe und ist deshalb sehr sauer.
Er ist darum auch der einzige, der nicht freundlich schaut, als der Fotograf dann das Foto macht.

Situation Changes
Heute wird ein Sportfest gemacht und so hüpft Simon die Stufen hinunter und rennt auf den Schulhof.
Da steht nämlich schon seine ganze Klasse bereit und wartet auf den Fotografen.
Simon entschuldigt sich bei der Lehrerin und stellt sich schnell vor Martin in die erste Reihe.
Doch da beschwert sich die Nervensäge Sarah, weil sie hinter Simon gar nicht zu hören ist.
Zum ersten Mal muss Simon also in die hinterste Reihe und ist deshalb sehr sauer.
Er ist darum auch der einzige, der nicht in die Kamera schaut, als der Fotograf dann das Foto macht.



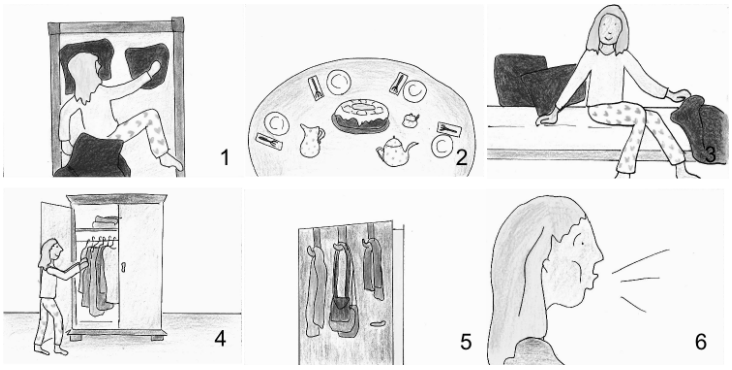
13. Kinderfasching

Original
Heute feiern alle im Kindergarten Fasching und der Indianer Lukas betritt gespannt die geschmückte Turnhalle.
Als erstes sieht er eine gruselige Hexe und gleich daneben unterhalten sich ein Cowboy und eine Prinzessin.
Er bemerkt, dass es seine drei besten Freunde sind, als sie ihm zuwinken.
Schnell pirscht Lukas sich an und bewirft die drei mit einer Ladung Konfetti.
Da zieht der Cowboy seine Wasserpistole und spritzt den verdutzten Lukas nass.
Lukas ergreift die Flucht und so jagen die beiden durch die Menge.

Text Surface Changes
An diesem Tag feiern alle im Kindergarten Fasching und der Indianer Lukas betritt gespannt die dekorierte Turnhalle.
Zuerst sieht er eine gruselige Hexe und gleich daneben plaudern ein Cowboy und eine Prinzessin.
Er erkennt, dass es seine drei engsten Freunde sind, als sie ihm zuwinken.
Schnell pirscht Lukas sich an und beschmeißt die drei mit einer Ladung Konfetti.
Da zieht der Cowboy seine Wasserpistole und macht den überraschten Lukas nass.
Lukas ergreift die Flucht und so jagen die zwei durch die Menge.

Text Base Changes
Heute feiern alle im Kindergarten Fasching und der Indianer Lukas betritt gespannt die kunterbunte Turnhalle.
Als erstes sieht er eine gruselige Hexe und gleich daneben sind ein Cowboy und eine Prinzessin.
Er bemerkt, dass er die drei bestens kennt, als sie ihm zuwinken.
Schnell pirscht Lukas sich an und ärgert die drei mit einer Ladung Konfetti.
Da zieht der Cowboy seine Wasserpistole und schon ist der verdutzte Lukas nass.
Lukas ergreift die Flucht und so jagen die beiden durch die Turnhalle.

Situation Changes
Heute feiern alle in der Schule Fasching und der Indianer Lukas betritt gespannt die geschmückte Turnhalle.
Als letztes sieht er eine gruselige Hexe und gleich daneben unterhalten sich ein Cowboy und eine Prinzessin.
Er bemerkt nicht, dass es seine drei besten Freunde sind, als sie ihm zuwinken.
Schnell pirscht Lukas sich an und bewirft die drei mit einer Ladung Bonbons.
Da zieht der Cowboy seine Wasserpistole und spritzt den verschmutzten Lukas nass.
Lukas ergreift die Flucht und so jagen die vier durch die Menge.



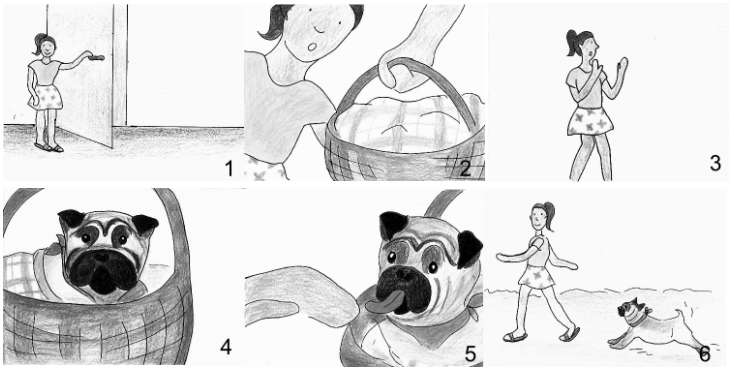
14. Maries Geburtstag

Original
Schlaflos wälzt sich Marie in ihrem Bett, denn morgen wird sie sieben Jahre alt.
Sie denkt an die vielen Geschenke und die lieben Freunde, die zu Kakao und Kuchen kommen.
Um sechs hält sie es in ihrem Bett nicht mehr aus und sie steht auf.
Sie zieht gerade ihr Lieblingskleid aus dem Schrank, als sie auf dem Flur leise Stimmen hört.
Da öffnet sich ihre Zimmertür und ihre Eltern kommen mit einer sehr großen Torte herein.
Marie bläst alle Kerzen auf Anhieb aus und wünscht sich ganz fest, dass sie jeden Tag Geburtstag hat.

Text Surface Changes
Schlaflos wälzt sich Marie in ihrem Bett, denn morgen wird sie sieben.
Sie denkt über die zahlreichen Geschenke und die netten Freunde nach, die zu Kakao und Kuchen kommen.
Um sechs erträgt sie es nicht mehr in ihrem Bett und sie steht auf.
Sie zieht gerade ihr Lieblingskleid aus dem Schrank, als sie leise Stimmen auf dem Gang hört.
Da öffnet sich ihre Zimmertür und ihre Eltern kommen mit einer riesigen Torte herein.
Marie bläst alle Kerzen auf Anhieb aus und wünscht sich ganz stark, dass jeder Tag ihr Geburtstag ist.

Text Base Changes
Schlaflos wälzt sich Marie in ihrem Bett, denn morgen wird sie ein Jahr älter.
Sie denkt an die vielen Geschenke und die lieben Gäste, die zu Kakao und Kuchen kommen.
Morgens hält sie es in ihrem Bett nicht mehr aus und sie steht auf.
Sie zieht gerade ihr Lieblingskleid aus dem Schrank, als sie auf dem Flur leise Geräusche hört.
Da öffnet sich ihre Zimmertür und ihre Eltern überraschen sie mit einer sehr großen Torte.
Marie bläst alle Kerzen auf Anhieb aus und wünscht sich ganz fest, dass sie jeden Tag Geschenke bekommt.

Situation Changes
Schlaflos wälzt sich Marie in ihrem Bett, denn gestern wurde sie sieben Jahre alt.
Sie denkt an die vielen Spiele und die lieben Freunde, die zu Kakao und Kuchen kommen.
Um zehn hält sie es in ihrem Bett nicht mehr aus und sie steht auf.
Sie zieht gerade ihr Lieblingskleid aus dem Schrank, als sie vor dem Haus leise Stimmen hört.
Da öffnet sich ihre Zimmertür und ihre Eltern kommen mit einem sehr großen Geschenk herein.
Marie bläst alle Kerzen auf Anhieb aus und wünscht sich ganz fest, dass sie jeden Tag schulfrei hat.



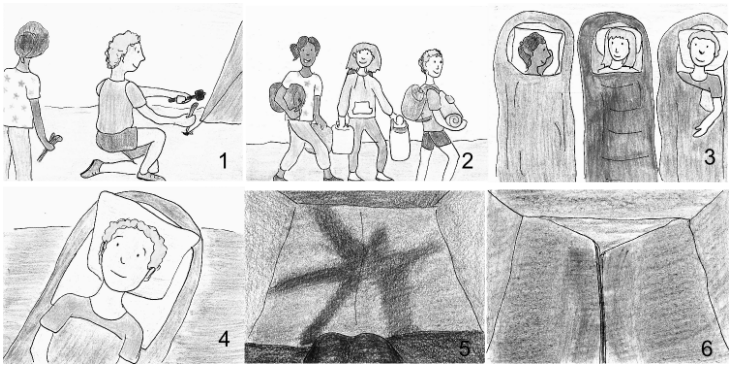
15. Die Tante kommt zu Besuch

Original
Sandra öffnet ihrer Tante die Haustür und die beiden umarmen sich freudig.
Die Tante trägt einen geheimnisvollen Korb unterm Arm und Sandra versucht hineinzuschauen.
Auf einmal bewegt sich der Korb und Sandra springt erschrocken zurück.
Da hört sie ein Winseln aus dem Korb und ein kleiner Hund streckt seinen Kopf heraus.
Sandra streckt ihm vorsichtig ihre Hand entgegen und der kleine Hund schleckt die Hand zutraulich ab.
Sie hebt ihn aus dem Korb und Sandra tollt gemeinsam mit dem Hund im Garten herum.

Text Surface Changes
Sandra öffnet ihrer Tante die Haustür und die beiden umarmen sich freudig.
Die Tante trägt einen geheimnisvollen Korb unterm Arm und Sandra versucht hineinzuschauen.
Auf einmal bewegt sich der Korb und Sandra springt erschrocken zurück.
Da hört sie ein Winseln aus dem Korb und ein kleiner Hund streckt seinen Kopf heraus.
Sandra streckt ihm vorsichtig ihre Hand entgegen und der kleine Hund schleckt die Hand zutraulich ab.
Sie hebt ihn aus dem Korb und Sandra tollt gemeinsam mit dem Hund im Garten herum.

Text Base Changes
Sandra öffnet ihrer Tante die Haustür und die beiden begrüßen sich freudig.
Die Tante bringt einen geheimnisvollen Korb mit und Sandra versucht hinzuschauen.
Auf einmal bewegt sich etwas in dem Korb und Sandra springt erschrocken zurück.
Da hört sie Laute aus dem Korb und ein kleiner Hund streckt seinen Kopf heraus.
Sandra streckt ihm langsam ihre Hand entgegen und der kleine Hund schleckt die Hand zutraulich ab.
Sie hilft ihm aus dem Korb und Sandra tollt gemeinsam mit dem Hund im Garten herum.

Situation Changes
Sandra öffnet ihrer Tante die Haustür und die beiden kichern freudig.
Die Tante versteckt einen geheimnisvollen Korb unterm Arm und Sandra versucht hineinzuschauen.
Auf einmal bewegt sich die Tante und Sandra springt erschrocken zurück.
Da hört sie ein Bellen aus dem Korb und ein kleiner Hund streckt seinen Kopf heraus.
Sandra streckt ihm stürmisch ihre Hand entgegen und der kleine Hund schleckt die Hand zutraulich ab.
Ihre Tante hebt ihn aus dem Korb und Sandra tollt gemeinsam mit dem Hund im Garten herum.



16. Die Zeltnacht

Original
An einem schönen Sommerabend bauen Tom, Alexa und Maja ein Zelt im Garten auf.
Sie dürfen draußen übernachten und holen die Schlafsäcke und Isomatten aus dem Keller.
Als es dunkel wird, kuscheln sie sich in ihre Schlafsäcke und erzählen sich Gruselgeschichten.
Bald schnarchen Alexa und Maja laut, nur Tom liegt hellwach im Zelt.
Da sieht er einen großen Schatten an der Zeltwand und kriecht tiefer in seinen Schlafsack.
Langsam geht der Reißverschluss des Zeltes auf und da hört Tom erleichtert die flüsternde Stimme seiner Mutter.

Text Surface Changes
An einem schönen Sommerabend bauen Tom, Alexa und Maja ein Zelt im Garten auf.
Sie dürfen draußen übernachten und holen die Schlafsäcke und Isomatten aus dem Keller.
Als es dunkel wird, kuscheln sie sich in ihre Schlafsäcke und erzählen sich Gruselgeschichten.
Bald schnarchen Alexa und Maja laut, nur Tom liegt hellwach im Zelt.
Da sieht er einen großen Schatten an der Zeltwand und kriecht tiefer in seinen Schlafsack.
Langsam geht der Reißverschluss des Zeltes auf und da hört Tom erleichtert die flüsternde Stimme seiner Mutter.

Text Base Changes
An einem warmen Sommerabend bauen Tom, Alexa und Maja ein Zelt im Garten auf.
Sie dürfen im Zelt übernachten und holen die Schlafsäcke und Isomatten aus dem Keller.
Als es kühler wird, kuscheln sie sich in ihre Schlafsäcke und erzählen sich Gruselgeschichten.
Bald schlafen Alexa und Maja tief, nur Tom liegt hellwach im Zelt.
Da sieht er einen großen Schatten an der Zeltwand und versteckt sich in seinem Schlafsack.
Langsam geht der Reißverschluss des Zeltes auf und da hört Tom erleichtert die vertraute Stimme seiner Mutter.

Situation Changes
An einem kühlen Sommerabend bauen Tom, Alexa und Maja ein Zelt im Garten auf.
Sie müssen draußen übernachten und holen die Schlafsäcke und Isomatten aus dem Keller.
Als es dunkel wird, kuscheln sie sich in ihre Schlafsäcke und erzählen sich Witze.
Bald streiten Alexa und Maja laut und Tom liegt hellwach im Zelt.
Da sieht er einen großen Schatten an der Zeltwand und schlüpft aus seinem Schlafsack.
Langsam geht der Reißverschluss des Zeltes auf und da hört Tom erleichtert die laute Stimme seiner Mutter.

The stories *Auf dem Bauernhof*, *In der Stadt*, *Reisevorbereitungen mit Hindernissen*, and *Die Fahrradtour* were only used in study II.

B Materials Study III

Practice Story 1

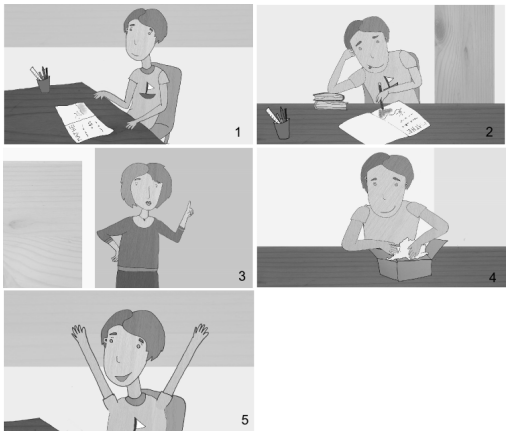
Die Sonne scheint vom Himmel und Patrick trifft sich mit seinen Freunden im Freibad.
Er will heute zum ersten Mal vom 5-Meter-Turm springen und steigt nun mutig die Leiter hoch.
Oben angekommen, geht er Schritt für Schritt nach vorne und kuckt nach unten.
Da zuckt er zusammen und schleicht vorsichtig zurück zur Leiter.
Aber leider kommen schon seine Freunde die Leiter hoch und feuern ihn an.
Jetzt kann er auf keinen Fall mehr zurück und so nimmt er Anlauf und springt.
Query word: Leiter

Practice Story 2

Weil Kathi und Annika ihr Zimmer streichen dürfen, fahren sie mit ihrem Vater Farbe kaufen.
Als sie ins Geschäft gehen, sagt der Vater, dass sie sich eine Farbe aussuchen sollen.
Schon bald streiten die beiden heftig, weil Kathi lieber ein grünes und Annika lieber ein rotes Zimmer haben.
Bald wird der Vater wütend und droht ihnen, ohne Farbe zu gehen, wenn sie nicht gleich still sind.
Plötzlich hat Kathi eine Idee und schnappt sich einfach zwei Farbeimer.
Wieder zu Hause, streichen sie die eine Wand grün und die andere rot und beide Mädchen sind zufrieden.
Query word: Witz

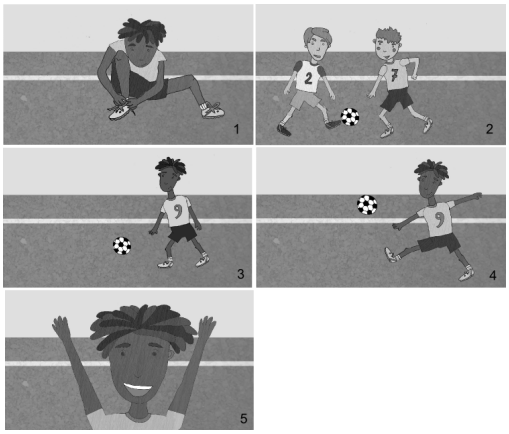
Note: Colored versions of the following pictures have been used in the experiments.

Experimental stories



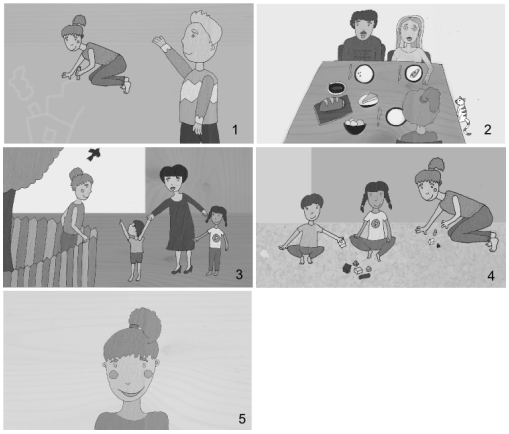
Benedikt sitzt am Schreibtisch und schaut ins Leere.
Er kritzelt in sein Matheheft und stellt sich wieder mal vor, wie es wäre Fotograf zu werden.
Als Fotograf könnte er viele berühmte Leute treffen und um die Welt reisen.
Da kommt seine Mutter herein und ermahnt ihn, seine Hausaufgaben weiter zu machen.
Auf einmal klingelt der Postbote und bringt ein großes Paket für Benedikt.
Er reißt das Paket auf und es liegt eine tolle Kamera darin.
Seine Oma hat ihm doch tatsächlich eine Kamera geschickt.
Benedikt ist überglücklich.

Global target word: Fotograf
Local target word: Kamera



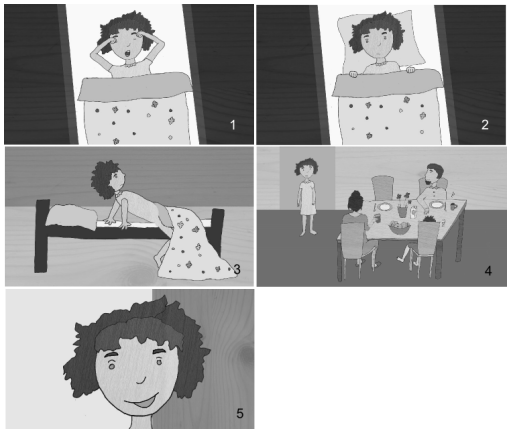
Heute ist das letzte Spiel der Saison und Christoph schnürt seine Fußballschuhe fest zu.
Er will sein Bestes geben, denn in diesem Jahr darf die Siegermannschaft sogar einen tollen Pokal mit nach Hause nehmen.
Während des Spiels kämpfen alle Spieler verbissen um den Ball, denn alle wollen unbedingt den Pokal holen.
Kurz vor Spielende sind beide Mannschaften gleich auf und es steht 1:1.
Christoph bekommt den Ball und rennt über das Spielfeld.
Kurz vor dem Tor holt er kräftig aus und schießt.
Der Ball geht ins Tor und sofort danach pfeift der Schiedsrichter das Spiel ab.
Christoph strahlt über das ganze Gesicht.

Global target word: Pokal
Local target word: Tor
Question: Welche Mannschaft gewinnt?



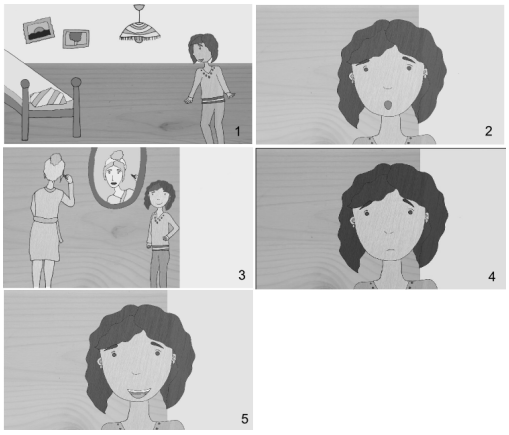
Linda sitzt auf der Straße und malt mit bunter Kreide, als ihr bester Freund Nico für einen kurzen Besuch vorbeikommt.
Der hat von seinen Eltern eine tolle, neue Uhr bekommen und Linda will unbedingt auch so eine haben.
Beim Abendessen erklären ihr ihre Eltern aber, dass sie ihr keine Uhr kaufen und sagen, dass sie sich selber darum kümmern muss.
Warum kann Linda nicht auch so tolle Eltern haben wie Nico?
Am nächsten Tag trifft sie die Nachbarin mit ihren zwei Kindern und die beiden plaudern ein bisschen.
Die Nachbarin fragt, ob Linda auf die Kinder aufpassen kann und dafür auch Geld bekommen soll.
So spielt Linda jeden Nachmittag mit den Kindern und schafft es schnell, eine Menge Geld zu verdienen.
Linda freut sich sehr.

Global target word: Uhr
Local target word: Geld



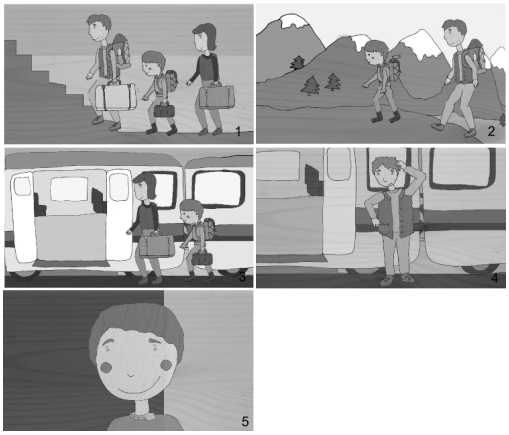
Ida ist gerade aufgewacht.
Sie zieht sich ihre Decke wieder über die Augen, denn sie will viel lieber im Bett bleiben und nicht in die Schule gehen.
In der Schule ist es immer so langweilig und außerdem mag Ida ihre neue Lehrerin nicht.
Aber auf einmal bemerkt sie, dass sie großen Hunger hat und so steht sie wohl oder übel auf.
Sie läuft runter in die Küche und reißt verwundert die Augen auf.
Da sitzt nämlich schon ihre ganze Familie um den Tisch und plötzlich fällt es ihr ein: Heute ist ja Sonntag!
Und am Sonntag frühstückt die ganze Familie immer lange und ausgiebig.
Ida ist begeistert.

Global target word: Schule
Local target word: Sonntag
Question: Wo ist Idas Zimmer?



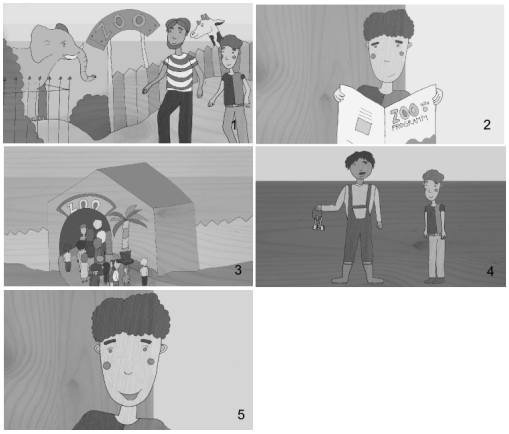
Maria läuft schon ganz aufgeregt in ihrem Zimmer umher.
Denn heute Abend darf sie mit ihren Eltern auf einen Ball gehen.
Schon so lange hat sie sich vorgestellt, wie toll es sein muss auf einen Ball zu gehen.
Am frühen Abend sieht sie ihrer Mutter zu, wie diese sich im Badezimmer schminkt und in ein hübsches Kleid schlüpft.
Sie darf ihrer Mutter sogar den feinen Schmuck aus der Kommode bringen.
Als sie selbst ihr feines Kleid anziehen will, bemerkt sie entsetzt, dass es viel zu klein geworden ist.
Doch da zieht ihre Mutter ein wunderschönes, glitzerndes Kleid für Maria aus einer Einkaufstüte.
Maria ist überglücklich.

Global target word: Ball
Local target word: Kleid
Question: Wo liegt der Schmuck, den Maria ihrer Mutter bringt?



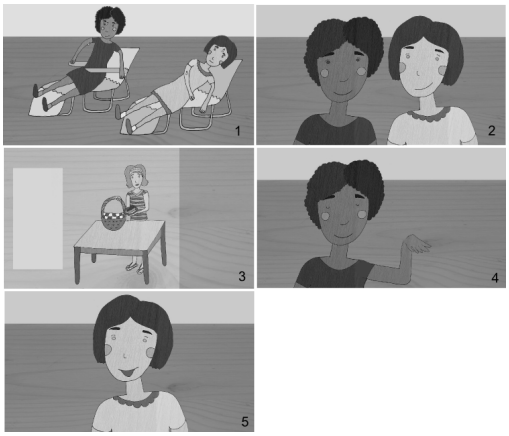
Familie Meier schleppt ihre vollgepackten Koffer die Treppen zum Bahngleis hoch.
Heute wollen sie in den Urlaub in die Berge fahren und Karl kann es kaum erwarten.
Im letzten Urlaub hat er tolle Wanderungen mit seinem Vater gemacht.
Da fährt auch schon der ICE mit quietschenden Bremsen in den Bahnhof ein.
Die Mutter und Karl steigen sofort in den Zug.
Nur der Vater steht noch draußen und kramt nervös in seiner Jackentasche nach den Tickets.
Doch dann streckt die Mama den Kopf durch das Zugfenster und wedelt mit den Tickets in ihrer Hand.
Karl grinst und freut sich.

Global target word: Urlaub
Local target word: Tickets



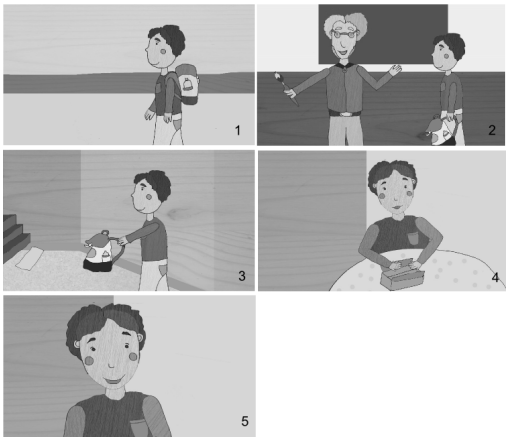
Heute geht Simon mit seinem Papa in den Zoo.
Simon freut sich besonders auf die Löwen.
Im Zooprogrammheft liest er, dass die Löwen um 10 Uhr gefüttert werden.
Am Eingang des Zoos steht schon eine lange Schlange.
Simon hasst lange Schlangen und er versteht einfach nicht, warum das alles so ewig dauert.
Da sieht er auf einmal, dass ein Mitarbeiter herbeieilt und eine zweite Kasse öffnet.
Er springt aus der Schlange und stellt sich schnell an der zweiten Kasse an.
Simon freut sich und strahlt seinen Vater an.

Global target word: Löwen
Local target word: Kasse



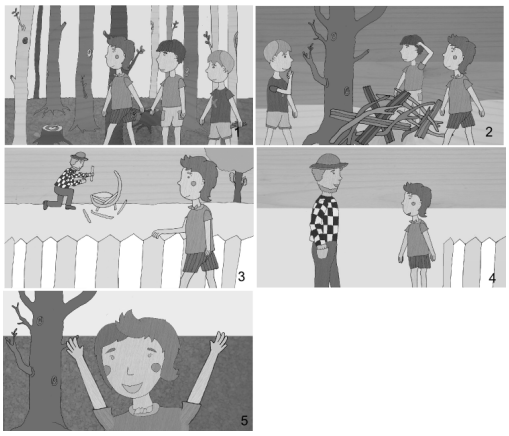
Emma ist bei Nele zu Besuch und die beiden sitzen an diesem heißen Tag schwitzend im Garten.
Da es einfach viel zu heiß ist, wollen sie im nahen Baggersee schwimmen gehen.
Sie grinsen sich an, denn in den kühlen Baggersee zu springen ist heute genau das Richtige.
Neles Mutter richtet den beiden einen großen Picknickkorb mit leckeren Broten.
Mhm lecker, sogar zwei Stücke Kuchen packt die Mutter in den Korb.
Als sie ihre Sachen zusammen packen, bemerkt Emma entsetzt, dass sie ja gar keinen Bikini dabei hat.
Doch Nele winkt beruhigend ab und zeigt ihr einen superschönen pinken Bikini, den sie ihr schenkt.
Emma strahlt übers ganze Gesicht und freut sich.

Global target word: Baggersee
Local target word: Bikini
Question: Zu welcher Jahreszeit spielt die Gechichte?



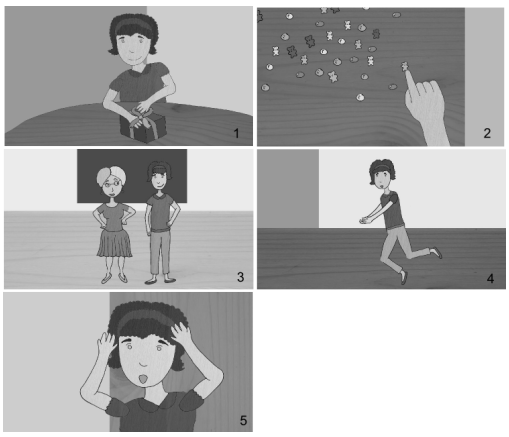
Global target word: Bild
Local target word: Stifte
Question: Warum ist das Päckchen mit den Stiften verstaubt?

Fabian läuft von der Schule nach Hause.
Er fängt sogar an zu rennen, denn er hat im Kunstunterricht ein Bild angefangen und will es jetzt noch fertig malen.
Der Lehrer fand das Bild schon ganz gut, aber hat ihm noch ein paar Tipps gegeben.
Zu Hause angekommen, schmeißt er seinen Ranzen in die Ecke und rennt schnell hoch in sein Zimmer.
Da sieht er, dass seine Mutter ihm ein verstaubtes braunes Päckchen auf seinen Schreibtisch gelegt hat.
Neugierig öffnet er es und sieht erfreut, dass es seine besten Stifte sind, die er letztes Jahr zu Weihnachten bekommen hat.
Diese Stifte hat er schon ewig gesucht und seine Mutter muss sie wohl beim Putzen gefunden haben.
Zufrieden grinst er vor sich hin.



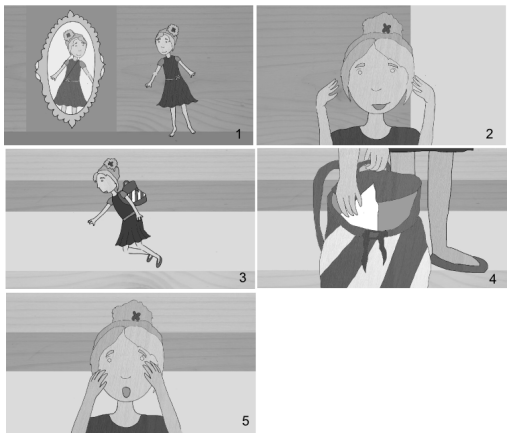
Global target word: Baumhaus
Local target word: Hammer

Jonas und seine zwei Freunde sind im Wald hinterm Haus.
Sie wollen ein Baumhaus bauen und Jonas hat auch schon den perfekten Platz gefunden.
Sie haben schon genug Holz für das Baumhaus zusammengetragen, aber sie wissen noch nicht, wie sie das alles zusammenbauen sollen.
Aber jetzt ist erstmal Zeit für eine kurze Pause.
Gerade als sie ins Haus reingehen, sieht Jonas seinen Nachbarn im Garten einen alten Schaukelstuhl reparieren.
Da kommt ihm eine Idee und er fragt den Nachbarn, ob er einen Hammer ausleihen könnte.
Der Nachbar lächelt und sagt ihm, dass er noch einen alten Hammer übrig hat und den einfach behalten darf.
Jonas lächelt und freut sich.



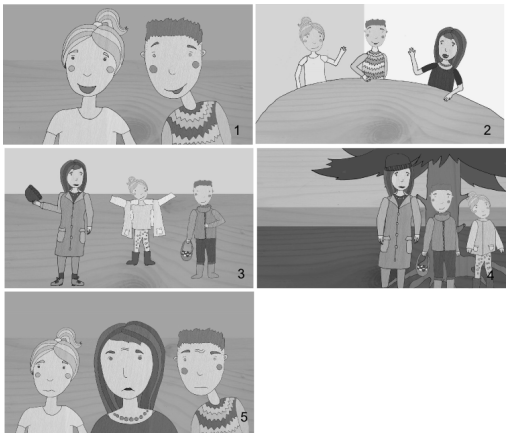
Pia verpackt ein Geschenk für ihre beste Freundin Katharina, die morgen Geburtstag hat.
Danach will sie für Katharina noch einen leckeren Kuchen backen.
Der Kuchen soll mit vielen Smarties und Gummibärchen verziert werden.
Katharina wird morgen schon acht Jahre alt und ist damit die Ältere der beiden.
Aber Pia ist die größere der beiden, sie ist sogar größer als ihre Lehrerin Frau Schubert.
Gerade hat Pia die Eier aus dem Kühlschrank genommen, da passiert es.
Pia stolpert und alle Eier purzeln auf die Fliesen und gehen kaputt.
Pia ist verzweifelt.

Global target word: Kuchen
Local target word: Eier
Question: Wie alt wird Katharina?



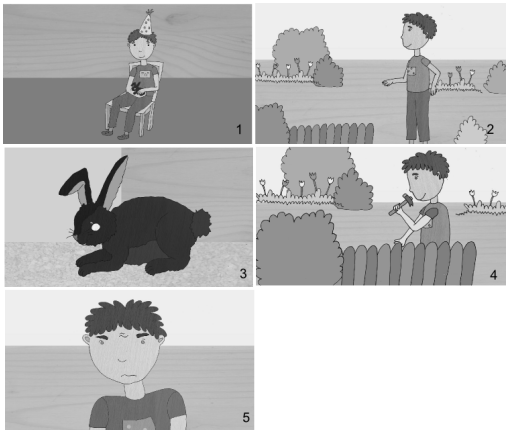
Hannah schaut sich nochmal im Spiegel an.
Heute wird nämlich ein Foto von der gesamten Klasse gemacht und sie fährt sich nochmal mit dem Kamm durch die Haare.
Hannah ist schon sehr aufgeregt, denn das Foto soll sogar in der Schülerzeitung abgedruckt werden.
Nach dem Frühstück schnappt sie sich ihren Ranzen und läuft los.
Sie nimmt die gewohnte Abkürzung über das Feld von Bauer Franz, doch auf einmal stolpert sie über einen großen Stein und fällt hin.
Als sie sich wieder aufrichtet, sieht sie, dass ihr schönes Kleid von oben bis unten mit Schlamm bespritzt ist.
Sie holt ein Taschentuch aus dem Ranzen und versucht vergeblich den Schlamm wegzuwischen.
Hannah ist entsetzt.

Global target word: Foto
Local target word: Schlamm



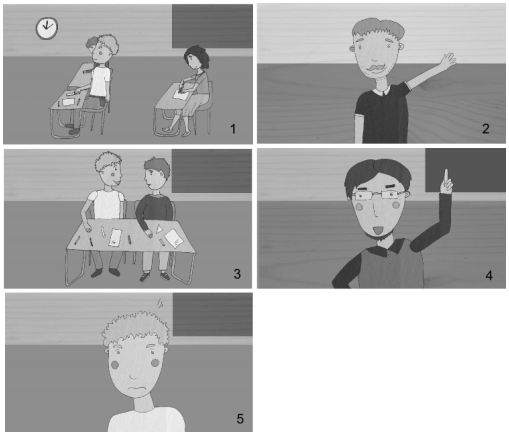
Global target word: Pilze
Local target word: Regen

Lena und ihr großer Bruder Finn freuen sich, denn die Oma kommt heute zu Besuch.
Mittags überlegen sich Lena, Finn und die Mutter, dass es zum Abendessen Pilze geben könnte.
Lena und Finn schlagen vor, die Pilze im Wald hinter ihrem Haus zu sammeln.
So ziehen sie sich ihre Jacken an und laufen frohen Mutes los.
Lange laufen sie durch raschelndes Laub, als auf einmal dunkle Wolken aufziehen.
Schnell stellen sie sich unter eine große Tanne und prompt setzt starker Regen ein, der einfach nicht aufhören will.
So nimmt die Mutter Lena an die Hand und die drei rennen so schnell sie können durch den Regen zurück nach Hause.
Im Haus angekommen sind alle drei ziemlich schlecht gelaunt.



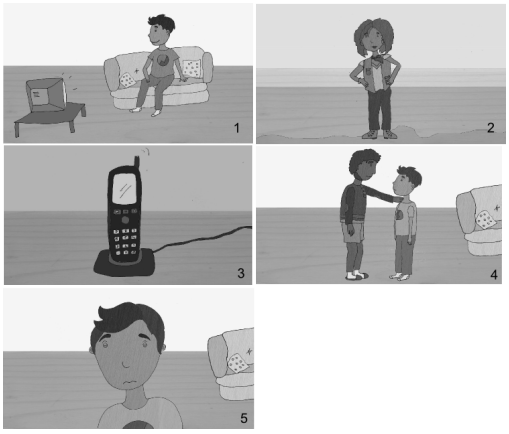
Global target word: Käfig
Local target word: Holzbrett
Question: Wie heißt der Hase von Luis?

Luis hat gestern zu seinem Geburtstag einen kleinen Hasen namens Lutz bekommen.
Er geht in den Garten, denn er will für den Hasen einen schönen Käfig bauen.
Im Moment wohnt der Hase in einem Schuhkarton, da ist so ein richtiger Käfig schon viel besser.
Der Hase ist rabenschwarz und hat nur einen einzigen weißen Punkt unter dem linken Auge.
Gerade knabbert er mit seinen kleinen Zähnchen an einem riesigen Blatt Löwenzahn.
Luis nimmt das letzte Holzbrett von einem Stapel neben dem Haus und will mit Schwung den ersten Nagel reinhauen.
Doch - oh nein - das Holzbrett ist schon ganz morsch und bricht gleich beim ersten Schlag in zwei Teile.
Luis ärgert sich sehr.



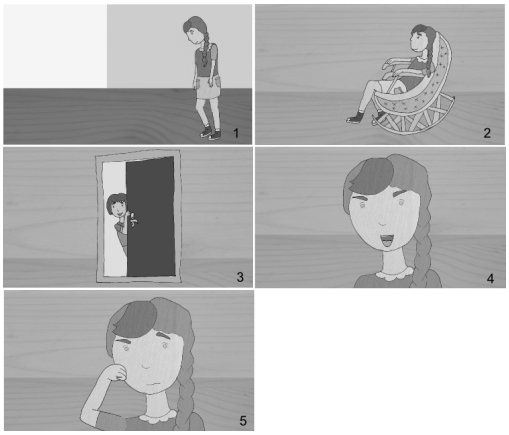
Global target word: Minigolf
Local target word: Nachsitzen
Question: Welches Unterrichtsfach hat Julian in der Geschichte?

Immer wieder schaut Julian auf die Uhr über der Tafel und kann es kaum erwarten, dass die Schule endlich vorbei ist. Denn am Nachmittag will er mit seinem Vater zum Minigolf gehen.
Er hat noch nie Minigolf gespielt und ist schon sehr gespannt.
Der Matheunterricht ist mal wieder total langweilig und so quatscht Julian lieber mit seinem Banknachbarn.
Der strenge Lehrer ermahnt die beiden, dass sie jetzt aber aufpassen sollen.
Und weil Julian aber das Quatschen einfach nicht lassen kann, verdonnert der gemeine Lehrer ihn zum Nachsitzen, gleich heute nach dem Unterricht.
Der Lehrer macht einen Klassenbucheintrag und erklärt ihm, dass er beim Nachsitzen auch gleich ein paar extra Aufgaben rechnen kann.
Julian ist sehr wütend.



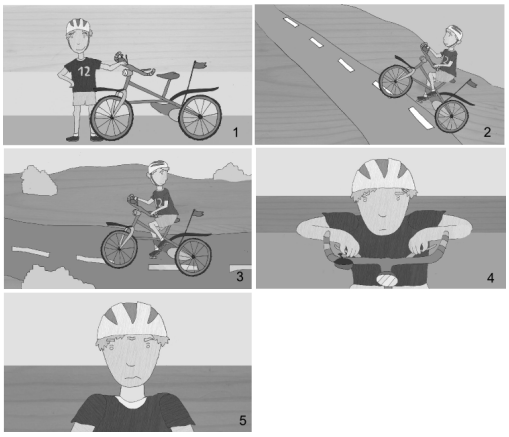
Global target word: Ausflug
Local target word: Verletzt

Abends sitzt David vor dem Fernseher und denkt freudig an morgen.
Morgen kommt nämlich seine coole Tante Mary vorbei und holt ihn ab für einen Ausflug.
Sie wollen mit Marys Motorrad einen Ausflug ans Meer machen.
Auf einmal klingelt das Telefon.
Er hört, wie sein Vater ans Telefon geht.
Kurz darauf kommt der Vater ins Wohnzimmer und erklärt ihm, dass die Tante sich bei einem Motorradunfall das Bein verletzt hat.
Sie hat sich sogar so schwer verletzt, dass sie für sechs Wochen einen Gips tragen muss.
David ist sehr traurig.



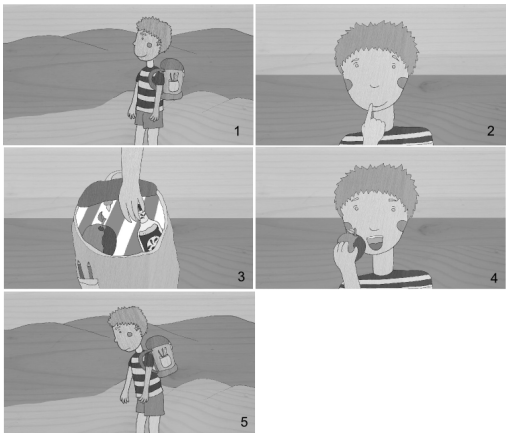
Global target word: schlafen
Local target word: Geige
Question: Wo möchte Frederike mit Christine spielen?

Christine kommt total geschafft von der Schule nach Hause.
Sie möchte sich nur in ihren Schaukelstuhl legen und einfach eine Runde schlafen.
Sie will zumindest noch kurz schlafen, denn heute Abend ist sie noch bei einer Party eingeladen.
Da streckt ihre kleine Schwester Frederike den Kopf zur Tür und will mit Christine draußen spielen.
Christine schüttelt aber den Kopf und sagt, dass Frederike alleine spielen soll.
Ein paar Minuten später hört Christine entsetzt, wie Frederike anfängt, auf ihrer Geige zu spielen.
Christine hält sich die Ohren zu, denn immer wenn Frederike Geige übt, klingt das so fürchterlich.
Christine ist genervt.



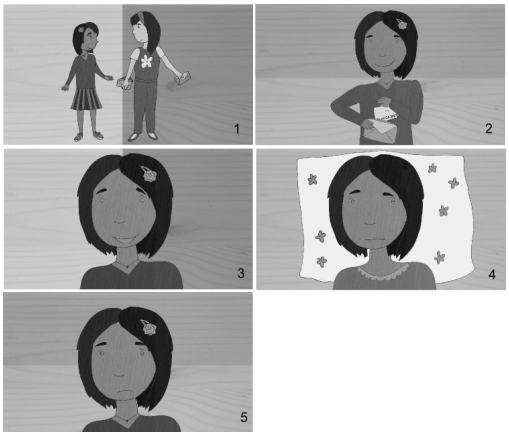
Global target word: Gewinnen
Local target word: Reifen
Question: Wann fährt Lukas den anderen davon?

Lukas fährt für sein Leben gerne Rennrad.
Heute ist in der Stadt ein wichtiger Wettkampf und er will unbedingt gewinnen.
Damit er gewinnen kann, hat er viel trainiert und fühlt sich heute Morgen topfit.
Als der Startschuss ertönt, fährt er den anderen gleich davon.
Am letzten Anstieg dreht er sich um und sieht, wie sich ein anderer Junge herankämpft und er tritt nochmal kräftig in die Pedale.
Auf der Zielgeraden bemerkt Lukas auf einmal, dass sein vorderer Reifen langsam die Luft verliert.
Mit plattem Reifen kann er nicht mehr weiterfahren und er bremst notgedrungen ab.
Lukas ärgert sich sehr.



Global target word: Hunger
Local target word: Schimmel
Question: Warum holt Jakob seine Trinkflasche aus dem Rucksack?

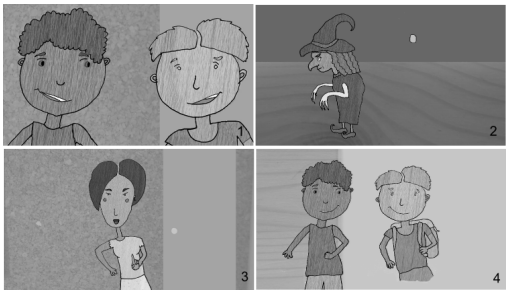
Jakob hat seine Hausaufgaben schon fertig und so schlendert er ein bisschen durch die Wiesen.
Nach einer ganzen Weile bemerkt er, dass er riesigen Hunger hat.
Er überlegt sich, wo er jetzt etwas zum Essen herbekommen könnte, um seinen Hunger zu stillen.
Inzwischen brennt auch die Sonne ganz schön heiß vom Himmel.
Er holt seine Trinkflasche aus dem Rucksack und entdeckt dabei einen leckeren Apfel, den ihm wohl seine Mutter in den Rucksack getan haben muss.
Er will gerade in den Apfel beißen, da sieht er, dass der Apfel voller Schimmel ist.
Vor Schreck schmeißt er den Apfel mit dem ekligen Schimmel weg.
Missmutig läuft er nach Hause.



Global target word: Party
Local target word: Krank

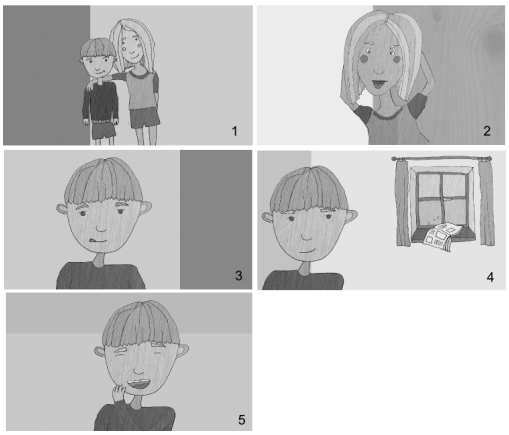
Gestern hat Luisas beste Freundin Elisabeth in der Schule bunte Briefe verteilt.
Luisa öffnet den Brief und zieht eine Einladungskarte für eine große Party, die schon morgen bei Elisabeth steigen soll, aus dem Umschlag.
Die Party ist sogar mit Übernachtung und nur die Mädchen haben eine Einladung bekommen.
Luisa freut sich darüber, denn die Jungs in ihrer Klasse sind einfach nur nervig.
Als Luisa am Abend ihren Schlafsack aus dem Schrank holt, fühlt sie sich gar nicht wohl und geht früh ins Bett.
Am nächsten Morgen wacht sie auf und ihr ist ganz heiß, so als ob sie krank werden würde.
Ihre Mutter eilt herbei und auch sie sieht auf den ersten Blick, dass Luisa krank ist.
Luisa ist sehr traurig.

Fillerstories with 8 sentences



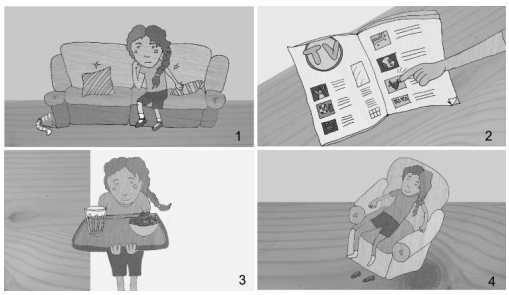
Query word: Rekord

Jan und Moritz treffen sich heute nach der Schule zum Spielen.
Sie wollen das neue Computerspiel von Jan ausprobieren.
In dem Computerspiel muss man eine böse Hexe mit faulen Eiern bewerfen.
Zuerst darf Jan ran und stellt gleich einen Rekord auf.
Moritz hat sich gerade den Joystick geschnappt, da kommt die Mutter herein.
Ziemlich sauer sagt sie, dass die beiden bei dem schönen Wetter draußen spielen sollen.
Sie wollen gerade schon missmutig zur Tür hinaus, da sehen sie, wie das Wetter schlecht wird und Regenwolken aufziehen.
Moritz und Jan grinsen sich an und freuen sich.



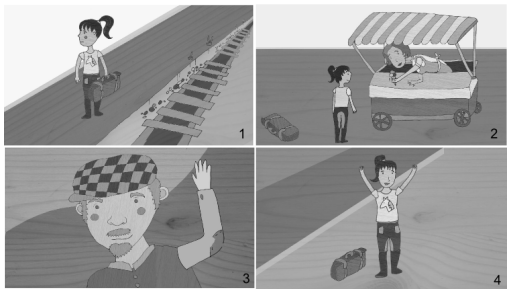
Query word: Regenwurm
Question: Warum kann Mia Tobias nicht dabei erwischen, wie er das Pupskissen unter ihren Schreibtischstuhl legt?

Eigentlich versteht sich Tobias mit seiner großen Schwester Mia richtig gut.
Aber heute will er ihr mal wieder einen richtigen Streich spielen.
Tobias mag es, Mia einen Streich zu spielen, denn sie sieht immer so lustig aus, wenn sie sich ärgert.
Solange Mia noch in der Schule ist, läuft Tobias durch das Wohnzimmer und überlegt.
Er könnte Mia einen Regenwurm ins Bett legen, aber das kommt ihm doch ein bisschen zu gemein vor.
Da sieht er auf einmal in seiner Kinderzeitschrift auf der Fensterbank das Extra der Woche: Ein Pupskissen.
Er schleicht in Mias Zimmer und platziert das Pupskissen unauffällig auf ihrem Schreibtischstuhl.
Tobias lacht sich ins Fäustchen.



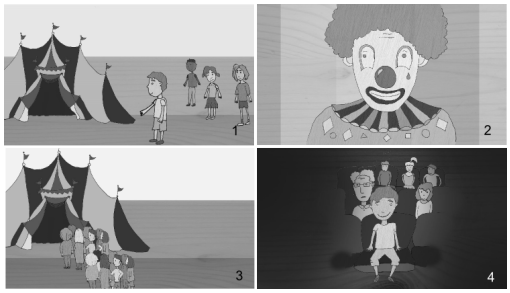
Query word: Flugzeug
Question: Wo lag der Stapel Zeitschriften?

Jule sitzt allein auf dem Sofa, denn heute hat niemand Zeit mit ihr zu spielen.
Aber zum Glück sieht sie in der Fernsehzeitung, dass ihr Lieblingsfilm läuft.
In fünf Minuten soll ihr Lieblingsfilm beginnen.
Sie rennt noch kurz in die Küche, um sich ein Glas Saft und ein paar Salzbrezeln zu holen.
Sie trägt alles ins Wohnzimmer und zieht sich den bequemen Fernsehsessel von Papa heran.
Jetzt sucht sie nur noch auf dem Couchtisch nach der Fernbedienung.
Hektisch schmeißt sie einen Stapel Zeitschriften auf den Boden und tatsächlich: Darunter liegt die Fernbedienung.
Jule lehnt sich zufrieden zurück.



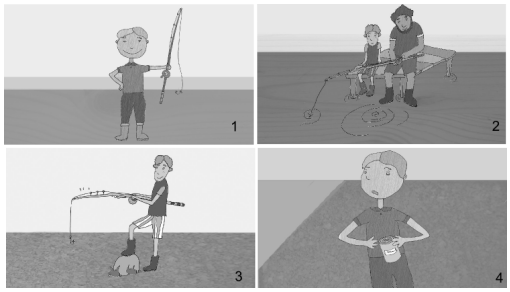
Query word: Krümel
Question: Wie lange steht Carolin am Bussteig und wartet auf den Bus?

Carolin hat ihre Koffer aus dem Zug gehievt und schaut sich suchend um.
In zehn Minuten soll der Bus kommen, der sie zum Reiterhof bringen soll.
Es ist das erste Mal, dass Carolin einen Urlaub auf einem Reiterhof machen.
Am Bahnsteig geht sie als erstes zu einem Kiosk und kauft sich noch schnell ein leckeres Eis.
Dann steht sie eine halbe Stunde auf dem leeren Bussteig herum, schleckt ihr Eis, aber der Bus kommt und kommt nicht.
Auf einmal biegt eine große Pferdekutsche um die Ecke und ein netter Mann sitzt auf dem Kutschbock.
Der Mann winkt ihr zu und erklärt ihr, dass er sie mit der Pferdekutsche zum Reiterhof bringt.
Carolin steigt begeistert auf und freut sich.



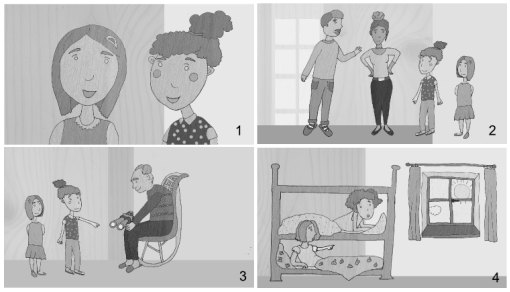
Query word: Kaffee
Question: Was findet Marlene am besten?

Markus geht heute mit seiner ganzen Klasse in den Zirkus.
Er will vor allem den Clown sehen, am besten ganz aus der Nähe.
Denn der Clown ist das Beste am ganzen Zirkus, zumindest findet Markus das.
Seine Klassenkameradin Marlene sieht das ganz anders.
Sie findet die Trapezkünstler mit den glitzernden Kleidern am besten.
Kurz vor der Vorstellung strömen die Zuschauen ins Zelt und alle wollen natürlich in die erste Reihe.
Aber Markus schafft es sich an den meisten Zuschauern vorbei zu drängeln und ergattert einen Platz in der ersten Reihe.
Markus strahlt über das ganze Gesicht.



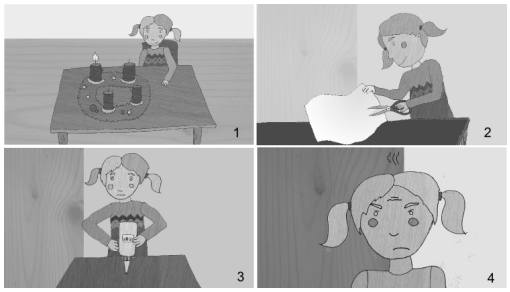
Query word: Nadel

Freds Freund Max hat zu seinem Geburtstag eine Angel bekommen.
Und heute will Fred mit ihm im nahen Bach Karpfen fangen gehen.
Fred war schon oft mit seinem Vater Karpfen angeln und kennt sich aus.
Er hat versprochen Max das Angeln beizubringen.
Fred bringt seine Profiangel und eine kleine Blechbüchse mit Ködern zum Angeln mit.
Max öffnet die Blechbüchse neugierig und darin winden sich hunderte kleine Maden.
Igitt, vor Schreck lässt er die Blechbüchse fallen und die vielen Maden verteilen sich auf dem Boden.
Fred wird ganz rot vor Wut.



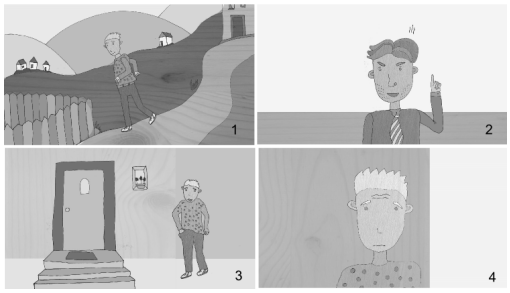
Global target word:
Local target word:

Carla und Sophie sind schon richtig aufgeregt.
Dieses Jahr haben die beiden sich nämlich fest vorgenommen, den Osterhasen beim Verstecken der Eier zu beobachten.
Ihre Eltern haben ihnen aber gesagt, dass der Osterhase sehr scheu ist und deshalb schon früh unterwegs sein wird.
Am Tag zuvor haben sie sich schon extra ein Versteck im Garten gesucht.
Außerdem haben sie sich noch das alte Fernglas von Carlas Opa geliehen.
Bevor sie schlafen gehen, stellen sie noch extra den Wecker, damit sie morgens früh wach werden.
Doch - oh nein - als die beiden am nächsten Morgen ihre Augen aufschlagen, bemerken sie, dass der blöde Wecker nicht geklingelt hat und es schon heller Tag ist.
Beide sind ganz verärgert.



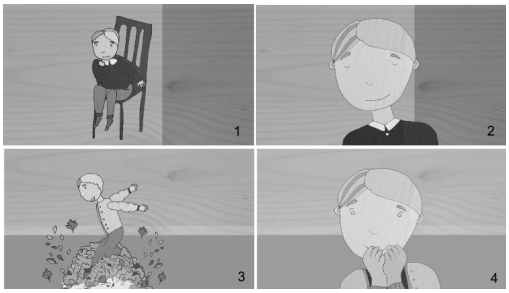
Query word: Kabel
Question: Mit wem hat Lisa die Pappe gekauft?

Heute ist der erste Advent und Lisa sitzt am großen Wohnzimmertisch.
Sie möchte viele tolle Fensterbilder basteln.
Mit den Fensterbildern soll das ganze Haus festlich geschmückt werden.
Mit ihrer Mutter war sie gestern im Bastelgeschäft und hat ganz viel bunte Pappe gekauft.
Fleißig schneidet Lisa die Pappe zurecht.
Doch als sie anfangen möchte die Motive zusammenzukleben, merkt sie, dass der Kleber alle ist.
So ein Mist – ohne Kleber kommt sie nicht weiter.
Lisa ärgert sich sehr.



Query word: Planet
Was wird der Hund jetzt wohl machen?

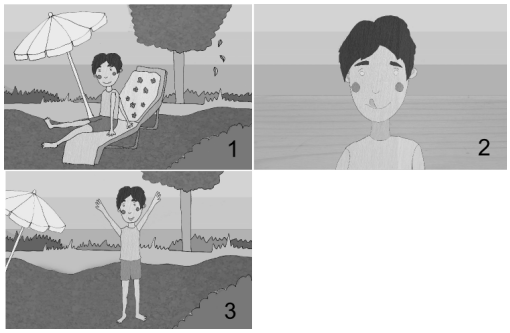
Hannes beeilt sich nach Hause zu kommen.
Er muss dann nämlich sofort mit seinem Hund Gassi gehen.
Wenn der Hund jetzt nicht sofort Gassi geht, wird er wieder auf den Wohnzimmerteppich pinkeln.
Uih, da hat der Vater das letzte Mal aber geschimpft!
Er steht schon vor der Haustür und hört seinen Hund von drinnen laut bellen.
Er kramt in seiner Tasche nach dem Hausschlüssel und da fällt es ihm wieder ein:
Der Hausschlüssel liegt oben in seinem Zimmer!
Hannes ist verzweifelt.



Query word: Dorf
Question: Warum raschelt es im Laub?

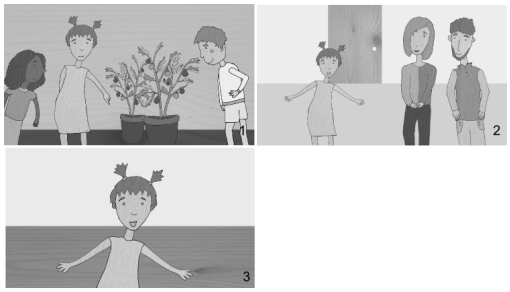
Wie so oft sitzt Paul alleine in seinem Zimmer und langweilt sich.
Er schließt die Augen und wünscht sich nichts sehnlicher als ein Haustier, mit dem er immer spielen kann.
Aber seine Eltern wollen ihm einfach kein Haustier erlauben.
Später geht er in den Garten und springt in die großen Laubhaufen, die er letzte Woche mit seinem Vater zusammengerecht hat.
Da raschelt es auf einmal im Laub direkt vor ihm.
Vorsichtig schaut Paul hin und sieht einen süßen kleinen Igel, der seinen Kopf aus dem Laub streckt.
Ganz zutraulich krabbelt er auf Paul zu und der nimmt den Igel ganz vorsichtig in die Hände.
Paul freut sich.

Fillerstories with 7 sentences



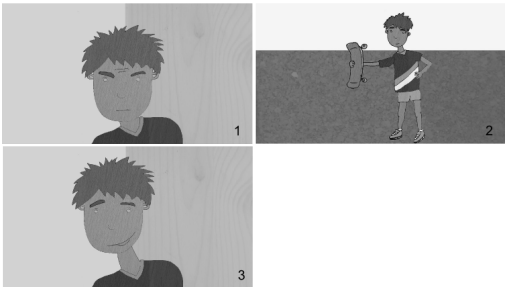
Es wird langsam Abend und Alex ist draußen im Garten. (1)
Er will heute mit seinem Vater ein Lagerfeuer machen.
Das Lagerfeuer soll sehr groß werden.
Außerdem wollen sie leckere Würstchen und Stockbrot machen und Alex läuft schon das Wasser im Mund zusammen. (2)
Der Vater sagt, dass sie viel Holz brauchen.
Und Alex findet hinter dem Gartenhaus einen riesigen Stapel mit Holz.
Alex grinst und freut sich. (3)

Query word: Karotte
Question: Was wollen Alex und sein Vater essen?



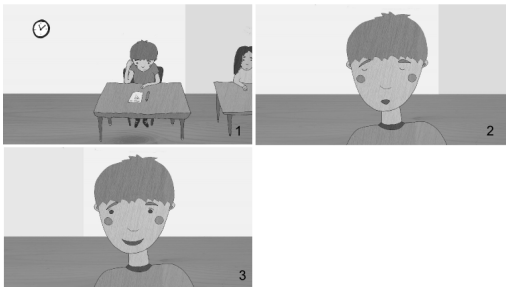
Kerstin hat heute in der Schule etwas über Pflanzen gelernt.
Die Klasse war im Schulgarten und die Kinder haben sich die Tomaten angeschaut.
Sie durften die Tomaten gießen und Unkraut jäten.
Das hat Kerstin so viel Spaß gemacht, dass sie es gleich zu Hause ihren Eltern erzählt.
Der Vater sagt lächelnd, dass er eine Überraschung für sie hat und führt sie hinters Haus.
Da sieht Kerstin, dass ihr Vater ihr ein kleines Beet gebaut hat - was für eine tolle Überraschung!
Kerstin ist überglücklich.

Query word: summen
Question: Wo hat Kerstin heute etwas über Pflanzen gelernt?



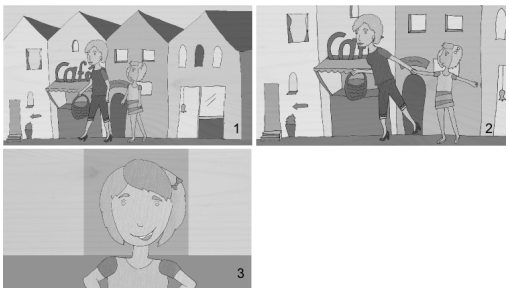
Query word: Kino

Tom ist in seinem Zimmer und steht wütend am Fenster.
Er hat Hausarrest bekommen und dabei will er heute doch unbedingt auf den Spielplatz.
Tom hat ein neues Skateboard bekommen und das wollte er auf dem Spielplatz unbedingt ausprobieren.
Tom ist sauer.
Da bemerkt er auf einmal den dicken Ast direkt vor seinem Fenster.
Über diesen Ast könnte er nach unten klettern, ohne, dass seine Eltern es merken.
Tom lächelt in sich hinein.



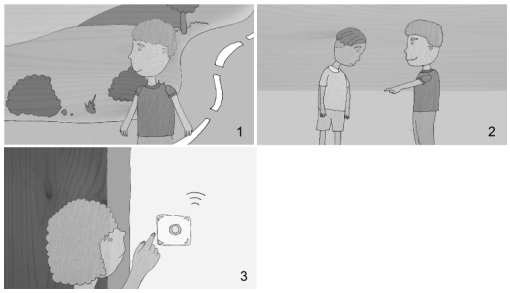
Query word: Kuh

Leon sitzt ganz konzentriert auf seinem Platz im Klassenzimmer.
Die Lehrerin teilt gerade die Aufgabenblätter für die Mathearbeit aus.
Leon will unbedingt eine gute Note in der Mathearbeit schreiben.
Denn dann bekommt er von seiner Mutter ein neues Computerspiel.
Er liest sich die Aufgaben durch und seufzt erleichtert.
Die Hälfte der Aufgaben hat er gestern mit seiner Mutter geübt.
Leon grinst und fängt sofort an zu rechnen.



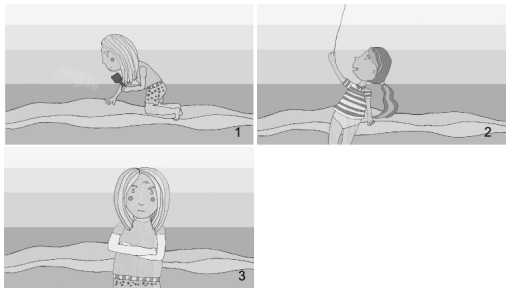
Query word: Soße
Question: Was wird sich Lilly mit dem Euro wohl kaufen?

Es ist ein heißer Sommertag und Lilly schlendert mit ihrer Mutter durch die Stadt.
Sie kommen an einem Café vorbei und Lilly will unbedingt ein Eis.
So ein Eis mit Himbeergeschmack wäre jetzt genau das Richtige!
Doch die Mutter möchte schnell weiter und noch Lebensmittel einkaufen.
Vor dem Supermarkt schaut Lilly nach unten und findet auf einmal einen Euro auf dem Gehweg.
Mit diesem Euro kann sie sich kaufen, was sie möchte.
Sie schaut zu ihrer Mutter und lächelt ihr zu.



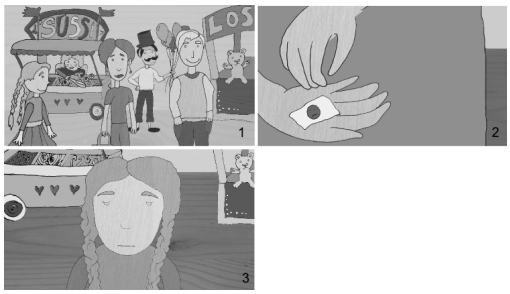
Daniel läuft langsam die Straße entlang.
Er will zu seinem Freund Tim, um sich bei ihm zu entschuldigen.
Er muss sich entschuldigen, denn er hat heute in der Schule etwas ganz Gemeines zu Tim gesagt.
Tim hatte sogar Tränen in den Augen und das tut Daniel sehr leid.
Er klingelt an Tims Tür und wartet ungeduldig.
Doch niemand öffnet ihm die Tür.
Enttäuscht geht Daniel zurück nach Hause.

Query word: Tränen



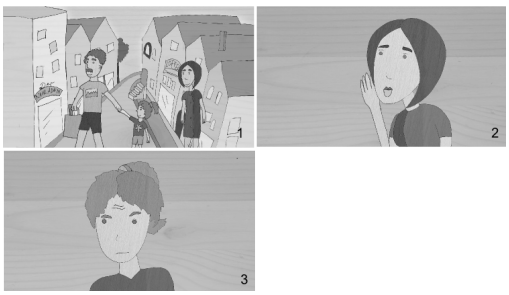
Judith fährt heute mit ihren Eltern und ihrer Schwester ans Meer.
Am Strand beginnt sie gleich damit, eine riesige Burg bauen.
Aber die großen Wellen spülen ihre Burg immer wieder weg.
Deshalb geht Judith ein bisschen weiter weg vom Wasser.
Ihre Schwester rennt durch den Sand und lässt ihren Drachen durch die Luft fliegen.
Plötzlich fällt der Drachen herunter und landet genau auf Judiths Kopf.
Heute hat Judith einfach kein Glück.

Query word: Sand



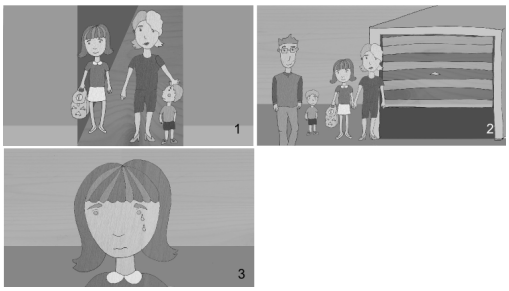
Leonie ist mit ihren Eltern auf dem Jahrmarkt.
An der Losbude sieht sie einen riesigen Teddy, den man gewinnen kann.
So einen Teddy wollte sie schon immer haben.
Sie quengelt so lange, bis ihre Eltern ihr schließlich genervt ein Los kaufen.
Ganz gespannt öffnet sie es, aber es ist eine Niete.
Mist, mit einer Niete gewinnt man nichts.
Leonie ist traurig.

Query word: Eiswürfel



Query word: Vogel

Marie war gestern mit ihren Eltern in der Stadt.
Sie hat neue Inlineskates bekommen und heute will sie diese unbedingt ausprobieren.
Sie steht früh auf und schlüpft in ihre Inlineskates, die sie sich direkt neben das Bett gestellt hat.
Sie will gerade zur Tür hinaus, da ruft die Mutter nach ihr.
Sie fragt, ob Marie vergessen hat, dass sie heute Tante Berta besuchen wollen.
Oh nein, bei Tante Berta ist es immer so langweilig.
Marie ärgert sich.



Query word: Nudel
Question: Warum ist Sonjas Familie spät dran?

Sonjas ganze Familie steht im Flur und ist schon sehr aufgeregt.
Denn heute ist Sonjas Einschulung.
Für die Einschulung hat sie einen tollen Schulranzen mit blauen Delfinen bekommen.
Da Papa mal wieder verschlafen hat, sind sie schon recht spät dran und stürzen aus dem Haus.
Als das Garagentor aufgeht, bemerken sie entsetzt, dass Auto einen Platten hat.
Oh nein, mit so einem Platten kann man nicht mit dem Auto fahren.
Sonja fängt an zu weinen.

C Instruction Study I

Gleich hörst du eine Stimme aus dem Computer. Das ist Katharina, die dir viele kurze Geschichten vorliest. Die Geschichten handeln von Kindern, die ganz unterschiedliche Sachen erleben. Katharina liest dir immer vier Geschichten vor und dann kommt deine Aufgabe. Du hörst die Geschichten nochmal einzeln Satz für Satz und nach jedem Satz sollst du entscheiden, ob er beim zweiten Mal irgendwie verändert wurde, also falsch vorgelesen wurde. Z.B. liest sie in der Geschichte vor: Der Junge isst ein Brot. Beim zweiten Mal liest sie dann vor: Der Junge isst ein Brötchen. Das wäre dann ein Fehler.
Immer wenn du merkst, dass Katharina beim zweiten Vorlesen einen Fehler gemacht hat, dann drückst du auf das Kreuz (auf Taste unter Kreuz zeigen). Und wo drückst du, wenn Katharina beim zweiten Mal alles so wie beim ersten Mal vorliest? Genau, auf das Häckchen. Wenn falsch: Bist du dir sicher?

Also nochmal: Wenn alles richtig ist Häkchen drücken, wenn etwas anders vorgelesen wird Kreuz drücken.
Damit es nicht zu leicht für dich ist, sind die Sätze in den Geschichten beim zweiten Mal durcheinander. D.h. der Beginn einer Geschichte kann dann auch mal am Ende vorgelesen werden. Das zählt dann aber nicht als Fehler.

Wenn Bedingung A:

Schau mal: Immer wenn du die Geschichten zum ersten Mal hörst, also gut zuhören sollst, siehst du diesen Punkt hier (Punkt zeigen). Immer wenn du Fehler suchen sollst, siehst du dieses Männchen.(Männchen zeigen) Später machen wir es dann noch ein bisschen anders, da hörst du die Geschichten dann nicht nur, sondern siehst auch Bilder dazu oder du kannst die Geschichten selber lesen. Aber das erklär ich dir dann.

Wenn Bedingung B:

Immer wenn du die Geschichten zum ersten Mal hörst, siehst du verschiedene Bilder zu den einzelnen Sätzen. Immer wenn du Fehler suchen sollst, erscheint auf dem Bildschirm dieses Bild (Männchen zeigen). Später machen wir es dann ein bisschen anders, da hörst du dann die Geschichten nur und siehst nur diesen Punkt (Punkt zeigen) oder kannst die Geschichten dann auch selber lesen. Aber das erklär ich dir dann auch nochmal.

Wenn Bedingung C:

Wir machen jetzt gleich eine Übung, bei der du die Geschichte hörst und diesen Punkt (Punkt zeigen) dazu siehst. Immer wenn du den Punkt siehst, sollst du gut zuhören. Immer wenn du Fehler suchen sollst, erscheint auf dem Bildschirm dieses Männchen hier (Männchen zeigen). Wenn wir dann richtig anfangen, ist es dann aber am Anfang so, dass du die Geschichten nicht hörst, sondern selber auf dem Bildschirm lesen kannst. Aber das erklär ich dir auch gleich nochmal.

Hast du bis jetzt Fragen?
Ok, dann üben wir das jetzt erst mal noch zusammen und du lernst Katharina kennen.

Enter. Katharina: Hallo ich heiße Katharina und ich freue mich, dass du heute mit dabei bist. Dir wurde ja schon ein bisschen erklärt, was du heute machen sollst und deshalb können wir auch gleich loslegen. Zuerst lese ich dir eine kurze Übungsgeschichte vor, also pass gut auf. Musik (Immer wenn die Musik kommt, fängt eine Geschichte an). Übungsgeschichte Katharina: Anne-Marie freut sich, denn ihre große Schwester spielt mit ihr. Zuerst verkleiden sie sich mit den Kleidern ihrer Mutter und schminken sich vor dem großen Spiegel. Danach gehen sie raus in den Garten und schaukeln ganz wild auf der Schaukel. Musik. So das war auch schon die Geschichte von Anne-Marie. Und jetzt bist du mit der Fehlersuche dran. Los geht´s.

Enter. Männchen erscheint.

Katharina: Anne-Marie freut sich, denn ihre Cousine spielt heute mit ihr.
Und, habe ich den Satz richtig vorgelesen? Dann drück das Häckchen, wenn du einen Fehler entdeckt hast, drück das Kreuz.

Und wo drückst du?

Wenn Kreuz gedrückt: *Super, genau, was war denn der Fehler?*

Wenn richtiger Fehler. *Genau das war die Schwester und nicht die Cousine.*
Wenn falscher Fehler: *Nein, das hat nicht ganz gestimmt. Hör mal nochmal genau zu: Beim ersten Mal hat Katharina gesagt: Anne-Marie freut sich, denn ihre große Schwester spielt mit ihr. Beim zweiten Mal hat sie dann aber vorgelesen: Anne-Marie freut sich, denn ihre große Cousine spielt mit ihr. Und ist dir jetzt was aufgefallen?*

Wenn Haken gedrückt:
Mhm, Nein, das hat nicht ganz gestimmt. Hör mal nochmal genau zu: Beim ersten Mal hat Katharina gesagt: Anne-Marie freut sich, denn ihre große Schwester spielt mit ihr. Beim zweiten Mal hat sie dann aber vorgelesen: Anne-Marie freut sich, denn ihre große Cousine spielt mit ihr. Und ist dir jetzt was aufgefallen? Genau die Cousine hat nicht gestimmt. Hören wir mal weiter.

Enter. Katharina: Achtung jetzt kommt schon der nächste Satz. Danach gehen sie raus in den Garten und schaukeln ganz wild auf der Schaukel. Und was drückst du jetzt?

Und wo drückst du? Kind Taste drücken lassen.

Wenn Haken gedrückt: *Super, genau, da hat sie alles richtig vorgelesen.*

Wenn Kreuz gedrückt: *Nein, das hat nicht ganz gestimmt. Hör mal nochmal genau zu: Beim ersten Mal hat Katharina gesagt: Danach gehen sie raus in den Garten und schaukeln ganz wild auf der Schaukel.. Beim zweiten Mal hat sie dann vorgelesen: Danach gehen sie raus*

in den Garten und schaukeln ganz wild auf der Schaukel. Und ist dir da ein Fehler aufgefallen? Nein, da hat alles gestimmt oder? Hören wir mal weiter.

Enter. Katharina: Jetzt kommt auch schon der letzte Übungssatz. Zuerst verkleiden sie sich mit den Klamotten ihrer Mutter und schminken sich vor dem großen Spiegel. Und welche Taste jetzt?

Kind Taste drücken lassen.

Wenn Kreuz gedrückt: *Super, genau, was war denn der Fehler?*

Wenn richtiger Fehler. *Genau beim ersten Mal hat Katharina Kleider vorgelesen und beim zweiten Mal dann Klamotten. Das ist ganz schon gemein gewesen. Sehr gut.*

Wenn falscher Fehler: *Nein, das hat nicht ganz gestimmt. Hör mal nochmal genau zu: Beim ersten Mal hat Katharina gesagt: Zuerst verkleiden sie sich mit den Kleidern ihrer Mutter und schminken sich vor dem großen Spiegel Beim zweiten Mal hat sie dann aber vorgelesen Zuerst verkleiden sie sich mit den Klamotten ihrer Mutter und schminken sich vor dem großen Spiegel. Und hast du den Fehler jetzt gehört? Das war aber auch schon echt ganz schön schwer.*

Nein, das hat nicht ganz gestimmt. Hör mal nochmal genau zu: Beim ersten Mal hat Katharina gesagt: Zuerst verkleiden sie sich mit den Kleidern ihrer Mutter und schminken sich vor dem großen Spiegel Beim zweiten Mal hat sie dann aber vorgelesen Zuerst verkleiden sie sich mit den Klamotten ihrer Mutter und schminken sich vor dem großen Spiegel. Und hast du den Fehler jetzt gehört? Das war aber auch schon echt ganz schön schwer.

Zusätzlich!!! bei Bedingung C:
Später hörst du die Geschichten dann genauso, wie du es gerade erlebt hast. Jetzt gleich machen wir es dann aber ein bisschen anders. Und zwar kannst du dir die Geschichten selber durchlesen. D.h. die Geschichten erscheinen jetzt immer Satz für Satz auf dem Bildschirm und immer wenn du einen Satz gelesen hast, drückst du hier auf Enter. Damit verschwindet der Satz und der nächste Satz erscheint dann. Wichtig ist, dass du dir die Sätze nur einmal anschauen kannst, wenn du einen Satz weggedrückt hast, dann kannst du nicht wieder zurückgehen. Auch wenn die Fehlersuche kommt, werden dir die Sätze nicht vorgelesen, sondern du kannst sie selber lesen. Okay, lass es uns mal ausprobieren.
Wenn du dir mal nicht ganz so sicher bist, dann kannst du auch raten. Hast du noch Fragen? Gut, dann fangen wir an. **Enter.** Versuch beginnt.

Nach vier Geschichten: Pausezeichen.

Bei Bedingung A:

Okay. Jetzt geht es ein bisschen anders weiter. Jetzt hörst du die Geschichten wieder, siehst aber dazu auch noch verschiedene Bilder. Alles andere bleibt gleich. Weiter geht´s. **Enter**

Bei Bedingung B:

Okay. Jetzt geht es ein bisschen anders weiter. Jetzt hörst du die Geschichten nicht mehr, sondern du kannst sie selber lesen. D.h. die Geschichten erscheinen jetzt immer Satz für Satz auf dem Bildschirm und immer wenn du einen Satz gelesen hast, drückst du hier auf Enter. Damit verschwindet der Satz und der nächste Satz erscheint dann. Wichtig ist, dass du dir die Sätze nur einmal anschauen kannst, wenn du einen Satz weggedrückt hast, dann kannst du nicht wieder zurückgehen. Auch wenn die Fehlersuche kommt, werden dir die Sätze nicht vorgelesen, sondern du kannst sie selber lesen. Okay, lass es uns mal ausprobieren. Weiter geht´s. **Enter**

Bei Bedingung C:

Okay, jetzt machen wir es so, wie es vorhin auch schon in der Übung war, d.h. Katharina liest dir die Geschichten vor und währenddessen siehst du einen Punkt. Bei der Fehlersuche liest sie dir die Geschichten nochmal vor und du versucht dann Fehler herauszuhören. Weiter geht´s.

Nach vier Geschichten: Pausezeichen

Bei Bedingung A:

Okay. Jetzt geht es ein bisschen anders weiter. Jetzt hörst du die Geschichten nicht mehr, sondern du kannst sie selber lesen. D.h. die Geschichten erscheinen jetzt immer Satz für Satz auf dem Bildschirm und immer wenn du einen Satz gelesen hast, drückst du hier auf Enter. Damit verschwindet der Satz und der nächste Satz erscheint dann. Wichtig ist, dass du dir die Sätze nur einmal anschauen kannst, wenn du einen Satz weggedrückt hast, dann kannst du nicht wieder zurückgehen. Auch wenn die Fehlersuche kommt, werden dir die Sätze nicht vorgelesen, sondern du kannst sie selber lesen. Okay, lass es uns mal ausprobieren. Weiter geht´s. **Enter**

Bei Bedingung B:

Okay, jetzt machen wir es so, wie es vorhin auch schon in der Übung war, d.h. Katharina liest dir die Geschichten vor und währenddessen siehst du einen Punkt. Bei der Fehlersuche liest sie dir die Geschichten nochmal vor und du versucht dann Fehler herauszuhören. Weiter geht´s. **Enter**

Bei Bedingung C:

Okay, nun geht es fast genauso weiter wie gerade nur siehst du jetzt bei den Geschichten, wenn du sie zum ersten Mal hörst keinen Punkt, sondern verschiedene Bilder. Die Fehlersuche bleibt gleich. Los geht´s. **Enter**

D Instruction Study II, 7-year-olds

Wir machen jetzt zusammen ein Spiel. Gleich hörst du eine Stimme aus dem Computer. Das ist Katharina, die dir viele kurze Geschichten vorliest. In den Geschichten kommen immer Kinder vor, die ungefähr genauso alt sind wie du und ganz unterschiedliche Sachen erleben. Katharina liest dir immer zwei Geschichten vor und danach kommt deine Aufgabe. Da sollst du nämlich Fehler suchen. Du hörst die Geschichte nochmal einzeln Satz für Satz und nach jedem Satz sollst du entscheiden, ob Katharina das zweite Mal den Satz irgendwie anders vorliest wie beim ersten Mal, also einen Fehler macht. Z.B. liest sie in der Geschichte vor: Der Junge isst ein Brot. Beim zweiten Mal liest sie dann vor: Der Junge isst ein Brötchen. Das wäre dann ein Fehler.

Immer wenn du merkst, dass Katharina beim zweiten Vorlesen einen Fehler gemacht hat, dann drückst du auf das Kreuz **(auf Taste unter Kreuz zeigen)** *Drück mal da auf die Taste* (**Kind drücken lassen**). *Und wo drückst du, wenn Katharina beim zweiten Mal alles so wie beim ersten Mal vorliest? Genau, auf das Häkchen* Wenn falsch: *Bist du dir sicher?*

Also nochmal: Wo würdest du drücken, wenn Katharina beim zweiten Mal den Satz richtig vorliest? Genau, auf das Häkchen. Und wo drückst du, wenn Katharina beim zweiten Mal einen Fehler macht? Richtig auf das Kreuz, bzw. die Taste unter dem Kreuz.

Damit es nicht zu leicht für dich ist, sind die Sätze in den Geschichten bei der Fehlersuche durcheinandergemischt. D.h. der Beginn einer Geschichte kann dann auch mal am Ende vorgelesen werden. Das zählt dann aber nicht als Fehler.

Wenn Bedingung A oder C:

Schau mal: Immer wenn du die Geschichten zum ersten Mal hörst, siehst du ganz verschiedene Bilder zu den Geschichten. Immer wenn du Fehler suchen sollst, erscheint auf dem Bildschirm dieses Bild (**Bild zeigen**). *Das ist ein Männchen, was sich am Kopf kratzt und sich fragt, ob Katharina auch alles richtig vorliest. Dabei musst du ihm helfen. Nach einer Weile wird es dann noch ein bisschen anders. Da siehst du dann, wenn du die Geschichten zum ersten Mal hörst und zuhören sollst, immer diesen Punkt hier* (**Punkt zeigen**) *Aber da sagt dir auch noch die Katharina Bescheid.*

Wenn Bedingung B oder D:

Schau mal: Immer wenn du die Geschichten zum ersten Mal hörst, also gut zuhören sollst, siehst du diesen Punkt hier (**Punkt zeigen**). *Immer wenn du Fehler suchen sollst, erscheint auf dem Bildschirm dieses Bild* (**Bild zeigen**). *Das ist ein Männchen, was sich am Kopf kratzt und sich fragt, ob Katharina auch alles richtig vorliest. Dabei musst du ihm helfen Nach einer Weile machen wir es dann ein bisschen anders. Da siehst du dann, wenn du die Geschichten zum ersten Mal hörst, nicht den Punkt, sondern immer Bilder zu den Geschichten. Aber da sagt dir auch noch die Katharina Bescheid.*

Hast du bis jetzt Fragen? Ok, dann üben wir das jetzt erst mal noch zusammen und du lernst Katharina kennen.

Enter. Katharina: Hallo ich heiße Katharina und ich freue mich, dass du heute mit dabei bist. Dir wurde ja schon ein bisschen erklärt, was du heute machen sollst und deshalb können wir auch gleich loslegen. Zuerst lese ich dir eine kurze Übungsgeschichte vor, also pass gut auf. **Musik**

Übungsgeschichte Katharina: Anne-Marie freut sich, denn ihre große Schwester spielt mit ihr. Zuerst verkleiden sie sich mit den Kleidern ihrer Mutter und schminken sich vor dem großen Spiegel. Danach gehen sie raus in den Garten und schaukeln ganz wild auf der Schaukel. Musik. So das war auch schon die Geschichte von Anne-Marie. Und jetzt bist du mit der Fehlersuche dran. Los geht´s.

Enter. Katharina: Anne-Marie freut sich, denn ihre Cousine spielt heute mit ihr. Und, habe ich den Satz richtig vorgelesen? Dann drück das Häkchen, wenn du einen Fehler entdeckt hast, drück das Kreuz.

Und wo drückst du? (**Taste drücken lassen**).

Wenn Kreuz gedrückt: *Super, genau, was war denn der Fehler?*
Wenn richtiger Fehler. *Genau das war die Schwester und nicht die Cousine.*

Wenn falscher Fehler: *Nein, das hat nicht ganz gestimmt. Hör mal nochmal genau zu: Beim ersten Mal hat Katharina gesagt: Anne-Marie freut sich, denn ihre große Schwester spielt mit ihr. Beim zweiten Mal hat sie dann aber vorgelesen: Anne-Marie freut sich, denn ihre große Cousine spielt mit ihr. Und ist dir jetzt was aufgefallen?*
Wenn Haken gedrückt: *Nein, das hat nicht ganz gestimmt. Hör mal nochmal genau zu: Beim ersten Mal hat Katharina gesagt: Anne-Marie freut sich, denn ihre große Schwester spielt mit ihr. Beim zweiten Mal hat sie dann aber vorgelesen: Anne-Marie freut sich, denn ihre große Cousine spielt mit ihr. Und ist dir jetzt was aufgefallen? Genau die Cousine hat nicht gestimmt. Hören wir mal weiter.*

Enter. Katharina: Achtung jetzt kommt schon der nächste Satz. Danach gehen sie raus in den Garten und schaukeln ganz wild auf der Schaukel. Und was drückst du jetzt?

Und wo drückst du? **Kind Taste drücken lassen.**

Wenn Haken gedrückt: *Super, genau, da hat sie alles richtig vorgelesen.*

Wenn Kreuz gedrückt: *Nein, das hat nicht ganz gestimmt. Hör mal nochmal genau zu: Beim ersten Mal hat Katharina gesagt: Danach gehen sie raus in den Garten und schaukeln ganz wild auf der Schaukel.. Beim zweiten Mal hat sie dann vorgelesen: Danach gehen sie raus in den Garten und schaukeln ganz wild auf der Schaukel. Und ist dir da ein Fehler aufgefallen?*

Nein, da hat alles gestimmt oder? Hören wir mal weiter

Enter. Katharina: Jetzt kommt auch schon der letzte Übungssatz. Zuerst verkleiden sie sich mit den Klamotten ihrer Mutter und schminken sich vor dem großen Spiegel. Und welche Taste jetzt?

Kind Taste drücken lassen.

Wenn Kreuz gedrückt: *Super, genau, was war denn der Fehler?*

Wenn richtiger Fehler. *Genau beim ersten Mal hat Katharina Kleider vorgelesen und beim zweiten Mal dann Klamotten. Das ist ganz schon gemein gewesen. Sehr gut.*

Wenn falscher Fehler: *Nein, das hat nicht ganz gestimmt. Hör mal nochmal genau zu: Beim ersten Mal hat Katharina gesagt: Zuerst verkleiden sie sich mit den Kleidern ihrer Mutter und schminken sich vor dem großen Spiegel Beim zweiten Mal hat sie dann aber vorgelesen Zuerst verkleiden sie sich mit den Klamotten ihrer Mutter und schminken sich vor dem großen Spiegel. Und hast du den Fehler jetzt gehört? Das war aber auch schon echt ganz schön schwer.*

Wenn Haken gedrückt: *Nein, das hat nicht ganz gestimmt. Hör mal nochmal genau zu: Beim ersten Mal hat Katharina gesagt: Zuerst verkleiden sie sich mit den Kleidern ihrer Mutter und schminken sich vor dem großen Spiegel Beim zweiten Mal hat sie dann aber vorgelesen Zuerst verkleiden sie sich mit den Klamotten ihrer Mutter und schminken sich vor dem großen Spiegel. Und hast du den Fehler jetzt gehört? Das war aber auch schon echt ganz schön schwer.*

Das hat doch schon sehr gut geklappt. Wenn du dir mal nicht ganz so sicher bist, dann kannst du auch raten. Auch wenn du mal einen Fehler machst, ist das gar nicht schlimm. Du probierst es einfach so gut wie möglich zu machen, okay? Hast du noch Fragen? Gut, dann fangen wir an.
Enter. Versuch beginnt.

E Instruction Study II, 9- and 11-year-olds

Gleich hörst du eine Stimme aus dem Computer. Das ist Katharina, die dir viele kurze Geschichten vorliest. Die Geschichten handeln von Kindern, die ganz unterschiedliche Sachen erleben. Katharina liest dir immer vier Geschichten vor und dann kommt deine Aufgabe. Du hörst die Geschichten nochmal einzeln Satz für Satz und nach jedem Satz sollst du entscheiden, ob er beim zweiten Mal irgendwie verändert wurde, also falsch vorgelesen wurde. Z.B. liest sie in der Geschichte vor: Der Junge isst ein Brot. Beim zweiten Mal liest sie dann vor: Der Junge isst ein Brötchen. Das wäre dann ein Fehler.

Immer wenn du merkst, dass Katharina beim zweiten Vorlesen einen Fehler gemacht hat, dann drückst du auf das Kreuz (**auf Taste unter Kreuz zeigen**). *Und wo drückst du, wenn Katharina beim zweiten Mal alles so wie beim ersten Mal vorliest? Genau, auf das Häckchen* wenn falsch: *Bist du dir sicher?*

Also nochmal: Wenn alles richtig ist: Häkchen drücken, wenn etwas anders vorgelesen wird: Kreuz drücken.

Damit es nicht zu leicht für dich ist, sind die Sätze in den Geschichten beim zweiten Mal durcheinander. D.h. der Beginn einer Geschichte kann dann auch mal am Ende vorgelesen werden. Das zählt dann aber nicht als Fehler.

Wenn Bedingung A oder C:

Immer wenn du die Geschichten zum ersten Mal hörst, siehst du verschiedene Bilder zu den Geschichten. Immer wenn du Fehler suchen sollst, erscheint auf dem Bildschirm dieses Bild (**Bild zeigen**). *Nach einer Weile wird es dann noch ein bisschen anders. Da siehst du dann, wenn du die Geschichten zum ersten Mal hörst und zuhören sollst, immer diesen Punkt hier* (**Punkt zeigen**) *Aber da bekommst du auch nochmal Bescheid.*

Wenn Bedingung B oder D:

Schau mal: Immer wenn du die Geschichten zum ersten Mal hörst, also gut zuhören sollst, siehst du diesen Punkt hier Punkt zeigen. Immer wenn du Fehler suchen sollst, erscheint auf dem Bildschirm dieses Bild (**Bild zeigen**). *Nach einer Weile machen wir es dann ein bisschen anders. Da siehst du dann, wenn du die Geschichten zum ersten Mal hörst, nicht den Punkt, sondern immer Bilder zu den Geschichten. Aber da sagt dir auch noch die Katharina Bescheid.*

Hast du bis jetzt Fragen? Ok, dann üben wir das jetzt erst mal noch zusammen und du lernst Katharina kennen.

Enter. Katharina: Hallo ich heiße Katharina und ich freue mich, dass du heute mit dabei bist. Dir wurde ja schon ein bisschen erklärt, was du heute machen sollst und deshalb können wir auch gleich loslegen. Zuerst lese ich dir eine kurze Übungsgeschichte vor, also pass gut auf. **Musik.**
Übungsgeschichte Katharina: Anne-Marie freut sich, denn ihre große Schwester spielt

mit ihr. Zuerst verkleiden sie sich mit den Kleidern ihrer Mutter und schminken sich vor dem großen Spiegel. Danach gehen sie raus in den Garten und schaukeln ganz wild auf der Schaukel. Musik. So das war auch schon die Geschichte von Anne-Marie. Und jetzt bist du mit der Fehlersuche dran. Los geht´s.

Enter. Katharina: Anne-Marie freut sich, denn ihre Cousine spielt heute mit ihr. Und, habe ich den Satz richtig vorgelesen? Dann drück das Häckchen, wenn du einen Fehler entdeckt hast, drück das Kreuz.

Und wo drückst du? (**Taste drücken lassen**).

Wenn Kreuz gedrückt: *Super, genau, was war denn der Fehler?*

Wenn richtiger Fehler. *Genau das war die Schwester und nicht die Cousine.*

Wenn falscher Fehler: *Nein, das hat nicht ganz gestimmt. Hör mal nochmal genau zu: Beim ersten Mal hat Katharina gesagt: Anne-Marie freut sich, denn ihre große Schwester spielt mit ihr. Beim zweiten Mal hat sie dann aber vorgelesen: Anne-Marie freut sich, denn ihre große Cousine spielt mit ihr. Und ist dir jetzt was aufgefallen?*

Wenn Haken gedrückt: *Nein, das hat nicht ganz gestimmt. Hör mal nochmal genau zu: Beim ersten Mal hat Katharina gesagt: Anne-Marie freut sich, denn ihre große Schwester spielt mit ihr. Beim zweiten Mal hat sie dann aber vorgelesen: Anne-Marie freut sich, denn ihre große Cousine spielt mit ihr. Und ist dir jetzt was aufgefallen? Genau die Cousine hat nicht gestimmt. Hören wir mal weiter*

Enter. Katharina: Achtung jetzt kommt schon der nächste Satz. Danach gehen sie raus in den Garten und schaukeln ganz wild auf der Schaukel. Und was drückst du jetzt?

Und wo drückst du? **Kind Taste drücken lassen.**

Wenn Haken gedrückt: *Super, genau, da hat sie alles richtig vorgelesen.*

Wenn Kreuz gedrückt: *Nein, das hat nicht ganz gestimmt. Hör mal nochmal genau zu: Beim ersten Mal hat Katharina gesagt: Danach gehen sie raus in den Garten und schaukeln ganz wild auf der Schaukel.. Beim zweiten Mal hat sie dann vorgelesen: Danach gehen sie raus in den Garten und schaukeln ganz wild auf der Schaukel. Und ist dir da ein Fehler aufgefallen? Nein, da hat alles gestimmt oder? Hören wir mal weiter*

Enter. Katharina: Jetzt kommt auch schon der letzte Übungssatz. Zuerst verkleiden sie sich mit den Klamotten ihrer Mutter und schminken sich vor dem großen Spiegel. Und welche Taste jetzt?

Kind Taste drücken lassen.

Wenn Kreuz gedrückt: *Super, genau, was war denn der Fehler?*

Wenn richtiger Fehler. *Genau beim ersten Mal hat Katharina Kleider vorgelesen und beim zweiten Mal dann Klamotten. Das ist ganz schon gemein gewesen. Sehr gut.*

Wenn falscher Fehler: *Nein, das hat nicht ganz gestimmt. Hör mal nochmal genau zu: Beim ersten Mal hat Katharina gesagt: Zuerst verkleiden sie sich mit den Kleidern ihrer Mutter und schminken sich vor dem großen Spiegel Beim zweiten Mal hat sie dann aber vorgelesen Zuerst verkleiden sie sich mit den Klamotten ihrer Mutter und schminken sich vor dem großen Spiegel. Und hast du den Fehler jetzt gehört? Das war aber auch schon echt ganz schön schwer.*

Wenn Haken gedrückt: *Nein, das hat nicht ganz gestimmt. Hör mal nochmal genau zu: Beim ersten Mal hat Katharina gesagt: Zuerst verkleiden sie sich mit den Kleidern ihrer Mutter und schminken sich vor dem großen Spiegel Beim zweiten Mal hat sie dann aber vorgelesen Zuerst verkleiden sie sich mit den Klamotten ihrer Mutter und schminken sich vor dem großen Spiegel. Und hast du den Fehler jetzt gehört? Das war aber auch schon echt ganz schön schwer.*

Das hat doch schon sehr gut geklappt. Wenn du dir mal nicht ganz so sicher bist, dann kannst du auch raten. Hast du noch Fragen? Gut, dann fangen wir an.

Enter. Versuch beginnt.

F Instruction Study III

Written condition

Lege als erstes deine Finger auf diese Tasten.

Jetzt erkläre ich dir, was du gleich machen wirst:
Bei dieser Aufgabe liest du gleich viele kurze Geschichten, in denen Kinder etwas erleben.
Zuerst siehst du auf dem Bildschirm kurz einen blauen Punkt – der bedeutet immer, dass eine neue Geschichte anfängt. Danach erscheint dann der erste Satz der Geschichte. Lies dir den Satz gut durch. Mit der Pfeiltaste gelangst du zum nächsten Satz. Du kannst jeden Satz solange lesen, wie du musst, um den Satz zu verstehen. Du kannst aber nicht wieder zurück, wenn du einmal die Pfeiltaste gedrückt hast. Wenn die Geschichte zu Ende ist, erscheint kurz dieses Kreuz auf dem Bildschirm (**Bild von Kreuz zeigen**). Dann musst du deine Finger auf jeden Fall auf der grünen und der roten Taste haben (aber das Beste ist, wenn du deine Finger die ganze Zeit so auf den Tasten lässt, wie ich es am Anfang gezeigt habe). Du siehst danach ein Wort auf dem Bildschirm.

Deine Aufgabe ist es dann, zu entscheiden, ob das Wort in der Geschichte vorgekommen ist oder nicht. Wenn das Wort in der Geschichte vorgekommen ist, drückst du auf die grüne Taste. Das bedeutet "Ja" (noch mal Taste zeigen). Wenn das Wort nicht in der Geschichte vorgekommen ist, drückst du auf die rote Taste. Das bedeutet "Nein" (noch mal Taste zeigen).
Du musst dich so **schnell** wie möglich entscheiden, ob das Wort in der Geschichte vorgekommen ist. Du sollst so schnell drücken, wie du kannst. Aber du sollst auch **keine Fehler** machen (mit Nachdruck sagen) – drücke also **so schnell wie möglich** aber auch **richtig**.

Nach manchen Geschichten wird dir außerdem noch eine Frage zu der Geschichte gestellt. Diese Fragen hörst du dann aus dem Computer und gleichzeitig steht die Frage auch in roter Schrift auf dem Bildschirm. Du sagst mir die Antwort (Info für VL: Nach jeder Frage geht es mit Pfeiltaste weiter).

Aber jetzt kannst du das erst mal mit zwei Geschichten üben! (*Mit Pfeiltaste startet die Übung; Wenn Vp falsch antwortet, steht auf dem Bildschirm "Oh nein, das war leider falsch" Wenn VP richtig antwortet, steht dort "Prima, das war richtig und auch ziemlich schnell!" Auch selbst noch loben, wenn richtig und wenn falsch "Das ist aber nicht schlimm, denn es war ja nur Übung und ich bin sicher, dass das gleich ganz super machen wirst!"*).

Wenn fertig mit Übung fertig, erscheint folgender Text auf dem Bildschirm: "Das hast du schon prima gemacht! Jetzt geht genauso weiter, sobald du die Pfeiltaste drückst."

Selbst noch sagen:
Das hast du prima gemacht. Dann kannst du es ja jetzt alleine machen. Du weißt ja wie es geht. Es kommt ein blauer Punkt und dann kommt die kurze Geschichte, Satz

für Satz. Mit der Pfeiltaste gehst du immer zum nächsten Satz. Nach jeder Geschichte erscheint ein Kreuz. Das ist das Signal, dass du bei dem Wort, das dann auf dem Bildschirm steht, so schnell du kannst entscheiden musst, ob das Wort in der Geschichte vorkam. Bei "Ja, das Wort kam vor", drückst du die grüne Taste und bei "Nein, das Wort kam nicht vor", drückst du die rote Taste.
Alles klar? Dann geht's jetzt los (*Start mit Pfeiltaste*)

Auditory and audiovisual condition

Lege als erstes deine Finger auf diese Tasten.

Jetzt erkläre ich dir, was du gleich machen wirst:
Bei dieser Aufgabe hörst du gleich viele kurze Geschichten, in denen Kinder etwas erleben.
Zuerst siehst du auf dem Bildschirm kurz einen blauen Punkt – der bedeutet immer, dass gleich eine neue Geschichte anfängt. Danach hörst du dann eine Geschichte.
Wenn die Geschichte zu Ende ist, erscheint kurz ein Kreuz auf dem Bildschirm. Dann musst du deine Finger auf jeden Fall auf der grünen und der roten Taste haben. (Aber das Beste ist, wenn du deine Finger die ganze Zeit so auf den Tasten lässt, wie ich es am Anfang gezeigt habe.) Danach hörst du nämlich ein Wort. **Deine Aufgabe ist es dann, zu entscheiden, ob das Wort in der Geschichte vorgekommen ist oder nicht. Wenn das Wort in der Geschichte vorgekommen ist, drückst du auf die grüne Taste. Das bedeutet "Ja". (noch mal grüne Taste zeigen). Wenn das Wort nicht in der Geschichte vorgekommen ist, drückst du auf die rote Taste. Das bedeutet "Nein". (noch mal Taste zeigen).**
Du musst dich so schnell wie möglich entscheiden, ob das Wort in der Geschichte vorgekommen ist. Du sollst so schnell drücken, wie du kannst. Aber du sollst auch keine Fehler machen (mit Nachdruck sagen) – Drücke also so schnell wie möglich aber auch richtig.
Aber jetzt kannst du das erstmal mit zwei Geschichten üben!

Mit Pfeiltaste startet die Übung. Wenn VP falsch antwortet, erscheint auf dem Bildschirm ein rotes Kreuz, wenn VP richtig antwortet, erscheint ein grünes Quadrat. Auch selbst noch loben, wenn richtig und wenn falsch "Das ist aber nicht schlimm, denn das war ja nur zum Üben und ich bin sicher, dass das gleich ganz super machen wirst!") Wenn die Übung fertig ist, erscheint folgender Text auf dem Bildschirm: „Weiter mit der Pfeiltaste!"

Selbst noch sagen:
Das hast du prima gemacht. Dann kannst du ja jetzt alleine weitermachen. Du weißt ja, wie es geht: Es kommt ein blauer Punkt und dann hörst du eine kurze Geschichte. Nach jeder Geschichte kommt ein Kreuz. Das zeigt dir, dass du gleich ein Wort hörst und du dann entscheiden sollst, ob das Wort in der Geschichte vorkam oder nicht. Bei Ja, drückst du die grüne Taste, bei Nein die rote Taste. Danach kommt wieder der Punkt. Dann geht's jetzt los!

Start mit Pfeiltaste.