

Wissenserwerb durch Bild-Text-Kohärenz im Fremdsprachenunterricht

Şerife Çelikkaya , Eskişehir – Gönül Karasu , Eskişehir

 <https://doi.org/10.37583/diyalog.1824366>

Abstract (Deutsch)

Ziel der vorliegenden Studie ist es, die Bild-Text-Kohärenz im Lehrwerk *Mit Deutsch* (Niveau A1.1) systematisch zu evaluieren. Die Untersuchung basiert auf einem qualitativen Forschungsdesign, das sich der Methode der Lehrwerkanalyse bedient. Aus der Perspektive der Lehrwerkanalyse wurden Bild-Text-Beziehungen nach dem dreistufigen Modell von Ballstaedt (2005) analysiert. Dieses Modell differenziert die Relationen zwischen Bild und Text in drei Kategorien: kongruent, komplementär und elaborativ. Zur Strukturierung der Analyse wurde ein Kategoriensystem entwickelt, das die Zuordnung der Bild-Text-Beziehungen zu den genannten Kategorien ermöglichte.

Die Ergebnisse der Untersuchung verdeutlichen, dass im Lehrbuch *Mit Deutsch* A1.1 überwiegend kongruente Bild-Text-Kombinationen vorliegen. Das bedeutet, dass Bild und Text zumeist dieselben Informationen transportieren und so eine redundante Verstärkung des Lerninhalts erfolgt. Weniger stark vertreten sind komplementäre und insbesondere elaborative Relationen, die potenziell dazu beitragen könnten, Lernende über die reine Textaussage hinaus zum vertieften Verständnis und zu neuen Deutungen anzuregen. Die Befunde legen nahe, dass zukünftige Lehrwerke eine stärkere Berücksichtigung elaborativer Bild-Text-Beziehungen in die Gestaltung einbeziehen sollten. Auf diese Weise kann nicht nur die Textkohärenz erweitert, sondern auch die Förderung der Sprach- und Interpretationskompetenz von Lernenden gezielt unterstützt werden.

Schlüsselwörter: *Deutsch als Fremdsprache, Lehrwerkanalyse, Bild, Text, Kohärenz.*

Abstract (English)

Knowledge Acquisition through Picture–Text Coherence in Foreign Language Education

The aim of the present study is to systematically evaluate picture-text coherence in the textbook *Mit Deutsch* (Level A1.1). The research is based on a qualitative design using textbook analysis as its methodological foundation. From the perspective of textbook analysis, picture-text relationships were examined according to Ballstaedt's (2005) three-level model. This model distinguishes between three types of relations: congruent, complementary, and elaborative. A categorization system was developed to assign the observed picture-text connections to these categories in a structured manner.

The findings reveal that the textbook *Mit Deutsch* A1.1 predominantly employs congruent picture-text relations. In most cases, picture and text convey identical information, thus reinforcing the learning content redundantly. Complementary and especially elaborative relations are less frequent. However, these types of relations hold the potential to stimulate learners beyond the surface meaning of the text,

encouraging deeper understanding, critical thinking, and new interpretive perspectives. The results suggest that the design of future textbooks should place greater emphasis on elaborative picture-text coherence. Such an approach would not only enrich textual coherence but also foster learners' language competence and their ability to interpret multimodal learning materials more effectively.

Keywords: *German as a foreign language, textbook analysis, image, text, coherence.*

EXTENDED ABSTRACT

In recent years, the integration of multimodal elements in foreign language teaching has received increasing attention, especially as the digitalization of education and new media literacy redefine how learners engage with content. Among these multimodal elements, visual materials—such as illustrations, photographs, and infographics—play a crucial role in enriching the communicative and cognitive functions of instructional texts. This extended abstract presents the findings of a qualitative study that explores the coherence between visual and textual elements in the officially approved German textbook *Mit Deutsch* (Level A1.1), which has been implemented in secondary schools in Türkiye since the 2023/2024 academic year. The central aim of the study is to evaluate how image-text relationships contribute to the development of linguistic competence and comprehension skills among learners of German as a Foreign Language.

The research employs a qualitative methodology based on textbook analysis, focusing on selected units within the textbook. A specifically developed three-stage analytical model was used to examine the visual-verbal relationships in the textbook. The model consists of: First Impression, which assesses the visual appeal, clarity, and attention-guiding properties of images; Activation of Prior Knowledge, which considers how images evoke learners' existing knowledge structures and prepare them cognitively for the accompanying text; and Image-Text Connection, which examines the degree to which visuals and texts are conceptually aligned to reinforce or expand meaning.

In addition to this three-level model, a categorization system was applied to classify the image-text relationships into three types: Congruent: Images that convey the same information as the text and reinforce its content, thereby aiding comprehension through redundancy. Complementary: Images that introduce new but thematically related information, enriching the content without repeating it. Elaborative: Images that encourage deeper thinking by prompting interpretation or inference, going beyond the literal text to activate higher cognitive processes.

The study's findings reveal that the majority of image-text relationships in the *Mit Deutsch* textbook fall into the congruent category. This suggests that most illustrations serve a confirmatory function, visually reinforcing what the text already communicates. Such visual support can facilitate basic comprehension, particularly for beginner-level learners. However, complementary relationships appeared less frequently, and most notably, elaborative relationships were completely absent. This absence indicates a missed opportunity for fostering higher-order cognitive skills such as inference-making, synthesis, and critical engagement.

Another notable finding is the inconsistency in the structural design of some text segments. Several texts lacked headings, which are essential for orienting readers and framing textual information. In some cases, the pairing between text and image did not seem to serve a functional purpose in terms of meaning-making or learner support. This lack of alignment may reduce the instructional value of multimodal presentation, especially in a foreign language learning context where learners often rely on visual cues for scaffolding understanding.

From a pedagogical standpoint, the study highlights both strengths and limitations in the textbook's design. On the one hand, the visual elements were found to be age-appropriate, aesthetically appealing, and generally well-integrated in terms of layout and formatting. These features can positively influence learner motivation and attentional focus. On the other hand, the underutilization of elaborative image-text coherence points to an important area for instructional improvement. At early stages of language acquisition, where students are developing basic literacy in the target language, the inclusion of more cognitively demanding visual-text relationships could foster deeper learning and engagement.

In light of these findings, the study emphasizes the importance of rethinking textbook design. Future instructional materials should aim to incorporate a more balanced mix of congruent, complementary, and elaborative image-text pairings. In particular, elaborative visuals can serve as cognitive tools that stimulate interpretation, critical reflection, and contextual learning. Such enhancements would not only support vocabulary acquisition and grammar understanding but also promote communicative competence, visual literacy, and independent learning strategies.

In conclusion, while the *Mit Deutsch* textbook offers a solid foundation for beginner-level language instruction through its visually supportive and clearly structured format, the lack of more complex image-text relationships signals a need for pedagogical advancement. Instructional designers, educators, and textbook publishers should consider integrating more elaborative and meaningful visual content in future editions to enrich learners' cognitive experiences and promote deeper linguistic and intercultural understanding.

Einleitung

Heutzutage steigt das Bedürfnis der Menschen, miteinander zu kommunizieren, rapide an. Mit der Entwicklung der Technologie erweitern sich auch die Bereiche, in denen kommuniziert wird. Damit der Einzelne sich in einem gesunden Kommunikationsprozess befindet, sollte er in der Lage sein, seine Gefühle und Gedanken frei auszudrücken, das Gelesene und Gehörte richtig zu verstehen, und das Gesprochene und Geschriebene verständlich auszudrücken. Darüber hinaus reicht es nicht aus, dass die Menschen nur ihre eigene Sprache beherrschen, um sich auf breiter Ebene verständigen zu können. Es ist äußerst wichtig, dass die Menschen neben ihrer eigenen Sprache auch Fremdsprachen lernen, damit diese Kommunikation ein größeres Netz erreicht. Das Erlernen einer Fremdsprache ist heute zu einer Notwendigkeit geworden.

Lehrbücher sind die am häufigsten verwendeten Elemente beim Sprachenlernen. Lehrbücher sind eines der Lehrmaterialien, in vielen Fällen sogar das einzige Lehrmaterial, das die Quelle für die Lernerfahrungen der Schüler ist, um die Bildungsziele zu erreichen. Im Lern- und Lehrprozess werden sie als Quelle benutzt, die das, was die Schüler lernen und was die Lehrer lehren, maßgeblich beeinflusst. Aufgrund dieser Funktion ist die Qualität von Lehrbüchern einer der Faktoren, die den Bildungs- und Ausbildungsprozess bestimmen.

Die Lehrbücher spielen eine entscheidende Rolle im Lehr- und Lernprozess. Um die Effektivität von Lehrbüchern zu gewährleisten, sollten bestimmte Aspekte während der Vorbereitungsphase berücksichtigt werden. Die Lehrbücher sollten im Einklang mit dem Lehrplan erstellt werden, wobei Aspekte wie visuelles Layout, Design, physische Struktur, Inhalt, Sprache und Erzählung zu berücksichtigen sind (Mentiş Taş 2007). İşeri (2007) stellt außerdem fest, dass das Lehrbuch ein grafisches Designprodukt ist, das in Form und Inhalt mit Informationen angereichert ist, und weist auf die Bedeutung der Illustration der für die Bücher ausgewählten Texte und deren pädagogische Gestaltung hin.

Der Harmonie zwischen Text und Bildern wird in den Lehrbüchern für den Fremdsprachenunterricht eine grosse Bedeutung zugeschrieben. Visuelle Darstellungen, die mit dem Text harmonieren, helfen dem Schüler, den Text besser zu verstehen und Rückschlüsse auf den Text zu ziehen, indem sie die abstrakte Struktur des Textes konkretisieren. Da visuelle Darstellungen einen ästhetischen Wert haben, ziehen sie die Aufmerksamkeit auf sich, werden schneller assimiliert als sprachliche Texte und wecken die Neugier auf den Text. Dies führt zu einer höheren Aktivierung, einer Stimulierung des Langzeitgedächtnisses bei den Kindern und einem besseren Abruf von Bildern als von Texten (Makowska 2013: 119). Sowohl bei den Kindern als auch bei Jugendlichen sollte das Bildmaterial in Fremdsprachenlehrbüchern Aufmerksamkeit erregen können. Denn wenn Informationen mündlich oder schriftlich, aber mit visueller Unterstützung gegeben werden, werden die Informationen aus mündlichen und schriftlichen Texten in der linken Hemisphäre des Gehirns gespeichert, während die Informationen aus visuellem Material in der rechten Hemisphäre gespeichert werden (Levie / Levin 1982). Auch nach der Dual-Coding-Theorie von Paivio (1986) werden Informationen sowohl verbal als

auch visuell gespeichert. Diese doppelte Kodierung führt zu einer besseren Erinnerungsleistung und erleichtert das Verstehen von abstrakten Konzepten. Verschiedene Studien haben gezeigt, dass die Wiedererkennensrate von Bildern außergewöhnlich hoch ist und dass bildhafte Darstellungen beim Erinnern und Reproduzieren von Informationen sprachlichem Material überlegen sind (Engelkamp 1998). Auch eine Untersuchung von Dwyer (1978) ergab, dass Bilder mit einem niedrigen Realitätsgrad die Lernleistung am effektivsten fördern, wenn das Lerntempo vorgegeben ist. Hingegen zeigen Bilder mit einem mittleren bis hohen Realitätsgrad einen größeren positiven Effekt auf die Lernleistung, wenn die Lernenden ihr eigenes Tempo bestimmen können. Es wurde auch festgestellt, dass funktionale Bild-Text-Kombinationen die Assoziationsmöglichkeiten erweitern und das Textverständnis positiv beeinflussen (Schnotz 1994). Darüber hinaus fördert eine gezielte Bild-Text-Kohärenz die Motivation und das Interesse der Lernenden, da sie komplexe sprachliche Strukturen durch visuelle Unterstützung besser erfassen können. Besonders bei selbstgesteuertem Lernen tragen visuell angereicherte Materialien dazu bei, die Lerneffizienz zu steigern und den individuellen Lernerfolg zu maximieren.

Ein weiterer wichtiger Aspekt der Bild-Text-Kohärenz im Fremdsprachenunterricht ist die Förderung des inferentiellen Denkens. Lernende müssen oft zwischen den visuellen und textuellen Informationen Verknüpfungen herstellen, um den Gesamtzusammenhang zu verstehen. Dies unterstützt nicht nur das Sprachverständnis, sondern auch die Entwicklung kognitiver Fähigkeiten wie Analyse, Interpretation und Schlussfolgerung. Zusätzlich beeinflussen verschiedene Faktoren die Effektivität der Bild-Text-Kombination. Dazu gehören die Gestaltung der Bilder, ihre Positionierung im Verhältnis zum Text sowie der Grad der Abstraktion. Dwyer (1978) stellte fest, dass Bilder mit einer geringen Realitätsnähe in strukturierten Lernumgebungen besonders effektiv sind, während realistischere Darstellungen beim selbstgesteuerten Lernen zu besseren Ergebnissen führen.

Die Multimodale Lernforschung zeigt zudem, dass eine enge semantische Verknüpfung von Bild und Text essenziell für den Lernerfolg ist (Mayer 2009: 89; Rentel 2011: 56), welches diesbezüglich auf formaler und inhaltlicher Ebene u.a. ihr Forschungsbereich ist. Unter Multimodalität werden die vielseitigen Sinnreize der Medien (Siever 2014; Wildfeuer / Bateman / Hiippala 2020) untersucht, was auch im Interesse dieser Arbeit liegt. Wenn z. B. Bilder lediglich dekorativ sind oder nicht klar mit dem Inhalt des Textes übereinstimmen, kann dies zu einer kognitiven Überlastung führen und das Verständnis erschweren. Deshalb ist es entscheidend, dass Bilder gezielt eingesetzt werden sollten, um den Lernprozess zu unterstützen, anstatt ihn zu komplizieren.

Im Kontext des Fremdsprachenlernens ist es besonders vorteilhaft, wenn visuelle Elemente nicht nur als unterstützende Elemente eingesetzt, sondern aktiv in den Lernprozess integriert werden. So können beispielsweise visuelle Elemente in Aufgaben integriert oder für interaktive Übungen genutzt werden, um den Lernerfolg zu steigern. Es lässt sich sagen, dass die gezielte Nutzung von Bild-Text-Kohärenz im Deutsch-als-Fremdsprache-Unterricht das Lernen effizienter und nachhaltiger gestaltet. Durch die

richtige Auswahl und Integration von Bildmaterial kann das Textverständnis verbessert, die Merkfähigkeit gesteigert und die Motivation der Lernenden gefördert werden. Laut Studien zur Emotionsforschung beim Lernen, die sich mit den Emotionen mithilfe konzeptioneller und methodischer Instrumente befassen, können Bilder, die positive Emotionen hervorrufen, die Aufmerksamkeit und damit die Verarbeitungsintensität erhöhen (Pekrun 2006). Auch diese Arbeit untersucht die Bild-Text Beziehung um herausarbeiten zu können, inwieweit es den Lerner vielseitig (sprachlich, kulturell) motivieren kann.

Auch der Abstand zwischen Bild und Text ist wichtig für die Assoziation von Informationen. Aus diesem Grund wird darauf hingewiesen, dass auf die Anordnung geachtet werden sollte und die Platzierung von Bildern über, unter oder neben dem Text die Lesbarkeit und Verarbeitung erleichtert (Ballstaedt 2012).

In Kinderbüchern erfüllen Bilder eine zentrale didaktische und motivationale Funktion. Sie unterstützen das Textverständnis, indem sie Inhalte kongruent abbilden (z.B. eine Illustration, die die Handlung einer Figur zeigt) und damit das Leseverstehen junger Kinder erleichtern. Gleichzeitig wirken sie oft komplementär, indem sie Emotionen oder Details vermitteln, die im Text nicht explizit genannt sind. Bei elaborativen Darstellungen erweitern Bilder den narrativen Rahmen, indem sie zusätzliche Interpretationsräume eröffnen. Damit übernehmen Bilder nicht nur eine Stütze für das Textverständnis, sondern auch eine Brücke zur kindlichen Lebenswelt, indem sie Interesse und Fantasie anregen (vgl. Kümmerling-Meibauer 2018).

Auch in wissenschaftlichen Kontexten dienen Bilder (z. B. Diagramme, Abbildungen, Tabellen, Modelle) in erster Linie der Veranschaulichung komplexer Inhalte. Sie haben meist eine kongruente Funktion, indem sie Daten oder Theorien grafisch darstellen und den Text präzisieren. In vielen Fällen übernehmen sie eine komplementäre Rolle, indem sie zusätzliche Informationen liefern, die der Text allein nicht vermitteln könnte (z. B. räumliche Strukturen in naturwissenschaftlichen Abbildungen). Elaborative Bild-Text-Beziehungen sind hier besonders wertvoll da sie Interpretationsspielräume und Hypothesenbildung fördern, etwa durch die Visualisierung von Modellen, die über den reinen Text hinaus Denkanstöße geben (vgl. Weidenmann 1994b).

Besonders im Fremdsprachenunterricht haben Bilder eine didaktische Schlüsselrolle. Sie fungieren als Unterstützung für Wortschatz- und Grammatiklernen, indem sie neue Begriffe kongruent illustrieren. Komplementäre Bilder ermöglichen, dass Lernende situative Kontexte und kulturelle Bedeutungen besser verstehen (z.B. Alltagsszenen in Lehrwerken, die interkulturelle Aspekte verdeutlichen). Elaborative Bilder regen zu kommunikativen Aufgaben an, indem sie Interpretationsspielräume eröffnen. Auf diese Weise werden Bild-Text-Beziehungen zu einem zentralen Werkzeug der Sprachförderung, das sowohl kognitive als auch affektive Lernprozesse unterstützt (vgl. Eitel & Scheiter 2015).

Damit Bilder effektiv zum Lernerfolg beitragen, sollten sie diesen drei Prinzipien entsprechen:

Kohärenzprinzip: Bilder sollten inhaltlich mit dem Text übereinstimmen und relevante Informationen vermitteln, damit sie den Lerner beim Erschließen des Lernstoffs nicht ablenken (Mayer / Moreno 2003: 49).

Signalisierungsprinzip: Wesentliche Informationen in multimedialen Lernmaterialien sollten durch Hervorhebungen, Markierungen oder visuelle Hinweise kenntlich gemacht werden, um die Aufmerksamkeit der Lernenden gezielt zu lenken und die kognitive Belastung zu reduzieren (Van Gog 2014: 20; Mayer 2009).

Redundanzprinzip: Bilder sollten nicht mit überflüssigen Informationen überladen sein, um kognitive Überlastung zu vermeiden (Sweller 1994: 303, 310).

Es fordert, dass Lernmaterialien prägnant und aufeinander abgestimmt gestaltet werden, um die kognitiven Ressourcen der Lernenden optimal zu nutzen.

Ein weiterer Punkt, der in Bezug auf die Bild-Text-Beziehung zu erwähnen ist, sind die Texte in den Lehrbüchern zur Fremdsprache. Heyd (1991: 73) definiert einen Text als „größere, semantisch-syntaktisch kohärente sprachliche Äußerung über ein Thema oder eine bestimmte Situation“. Der Text steht im Mittelpunkt jeder sprachlichen Kommunikation; er stellt eine externe und objektive Verbindung zwischen Sender und Empfänger her, unabhängig davon, ob sie wechselseitig oder auf Distanz miteinander kommunizieren (GER 2020). Texte werden auch als wichtiges Material im Fremdsprachenunterricht verwendet, um Wortschatz und Sprachkenntnisse zu entwickeln und Kultur zu vermitteln. Darüber hinaus wird die Sprache auf sinnvolle Weise in ihrem Kontext gelehrt, indem Sprachstrukturen und Wörter in Sätzen verwendet werden.

Im Fremdsprachenunterricht nehmen Lesetexte einen wichtigen Platz ein, wenn es darum geht, Material für alle Sprachlernfähigkeiten bereitzustellen. Sprachlernende lernen nicht nur die Struktur der Sprache, sondern beginnen auch, den kulturellen Hintergrund der Gesellschaft zu verstehen, in der die Sprache gesprochen wird. Dies ermöglicht es ihnen, die Sprache sinnvoller und effektiver zu nutzen. Der Prozess des kulturellen Bewusstseins durch Texte vermittelt den Lernenden nicht nur die Strukturen der Sprache, sondern auch, wie und wann die Sprache verwendet werden sollte und wie man sich in bestimmten sozialen und kulturellen Situationen benimmt. Der Gemeinsame Europäische Referenzrahmen für Sprachen (GER 2020) betont, dass die Lernenden mit einer großen Vielfalt an Texten in Berührung kommen sollten, um ihre Fähigkeiten zu entwickeln. Diese Texte gehören zu den Elementen, die zur Verstärkung des Sprachenlernens im Bildungs- und Ausbildungsprozess der Lernenden beitragen.

Im Fremdsprachenunterricht werden authentische Texte verwendet. In der Fachliteratur werden authentische Texte als sprachliche Materialien definiert, die aus realen Kommunikationskontexten der Zielsprache stammen und eine Vielzahl von Textsorten umfassen können. Dazu zählen unter anderem Lieder, Gedichte, Rezepte, Einkaufslisten und Audioguides sowie digitale Kommunikationsformen wie E-Mails oder traditionelle Medien wie Postkarten. Diese Texte stehen im Gegensatz zu didaktisch konzipierten Lehr- und Lernmaterialien, die speziell für pädagogische Zwecke erstellt

wurden, um sprachliche Strukturen gezielt zu vermitteln. Die authentischen Texte sind für Muttersprachler produzierte Texte, die den Lernenden der Fremdsprache helfen, ihre Kommunikationsfähigkeiten in der Zielsprache zu entwickeln, da sie natürliches Sprachmaterial enthalten. Berardo (2006: 65) nennt die Vorteile der Verwendung authentischer Texte im Lesematerial wie folgt:

- Authentische Texte spiegeln die reale Sprache wider und zeigen die Vielfalt der Sprache auf.
- Die Schülerinnen und Schüler werden über das aktuelle Geschehen in der Welt informiert.
- Authentische Texte enthalten keine Themen, die sehr schnell aus der Mode kommen, und neigen nicht dazu, Dinge zu enthalten, die sprachlich nicht angemessen sind.
- Authentische Texte sind ideal für das Lehren und Üben kleinerer Fertigkeiten.
- Sie enthalten ein breites Spektrum an Textarten und Sprachstilen, die in herkömmlichen Unterrichtsmaterialien nicht vorkommen.
- Authentische Texte regen die Schüler zum Lesen an, weil sie interessante Themen enthalten.

Das Verhältnis zwischen den beiden Hauptelementen eines Lehrbuchs, dem Bild und dem Text, wirkt sich unmittelbar auf die Leistungen der Lernenden und das Erreichen der Lehrplanziele aus. Eine ansprechende Gestaltung von Lehrbüchern, mitunter durch Unterstützung computerpräsentierter Texte, trägt zum Verständnis und zur Motivation bei, wenn sie mit dem Text verbunden ist (Drewniak 1992). Ein Bild ist kein Text, kann aber wie ein Text gelesen werden und semantisch textähnliche Funktionen übernehmen (Ehlich 2005). Dies kann zu einer genaueren Betrachtung des Bildes, einem gesteigerten Interesse am Inhalt und zu einer intensiveren Beschäftigung mit dem Text führen. Aus diesem Grund heißt es in der Verordnung über Lehrbücher und Lehrmittel des Ministeriums für nationale Bildung (Millî Eğitim Bakanlığı 2024), dass die visuelle Gestaltung und die inhaltliche Gestaltung von Lehrbüchern das Lernen unterstützen und die Entwicklungsmerkmale der Lernenden berücksichtigen sollten. Damit die Lernenden auf höchstem Niveau von den Bildern im Lehrbuch profitieren können, ist es sehr wichtig, dass die zu bestimmenden Bilder zusammen mit dem entsprechenden Text in das Lehrbuch aufgenommen werden, dass ihre Farben angemessen sind und dass sie mit dem Thema in Verbindung stehen. Beides zusammen ist wichtig, um andere Fähigkeiten zu unterstützen, insbesondere die kommunikative Funktion und die Wahrnehmung von Informationen (Ballstaedt 2012). Textbezogenes Bildmaterial hilft, das Thema in ein Szenario einzuordnen, regt Ideen an und unterstützt die Erstellung eines mentalen Modells eines komplexen Themas (Weidenmann 1994a).

Texte	Bilder
*nennen und beschreiben	*zeigen
*sind diskursiv, d.h. nicht den Sinnen gegeben, sondern werden durch innere Vorstellungen über Begriffe verarbeitet.	*sind präsentativ d.h. unmittelbar den Sinnen gegeben, z.B. durch Farben, Formen und räumliche Darstellungen
*können mündlich oder schriftlich sein	*bieten hohen Reizwert

*zeichnen sich durch Kohäsion und Kohärenz aus	*sind Momentaufnahmen
*übermitteln Emotionen und Informationen	*fördern, wenn unbekannt, das Spekulieren und aktivieren Wissen
*ermöglichen abstrahierende, reflektierende, analytische Operationen	*sind grundsätzlich offener als Texte
*dienen der Überlieferung	* haben einen hohen Identifikationswert
*dienender Speicherung sprachlichen Handelns	*haben eine im Vergleich zu Texten hohe Informationsdichte
*sind sukzessive Darstellungen, d.h. Elemente werden nacheinander dargeboten	*sind simultane Darstellungen, d.h. Elemente werden gleichzeitig dargeboten

Tab. 1: Wichtige Eigenschaften und Aufgaben von Text und Bild (Weidenmann 1994a)

Tabelle 1 zeigt in Stichpunkten die wichtigen Eigenschaften und Aufgaben der Texte und Bilder. Diese sollten als Leitlinien bei der Wahl der Texte und Bilder in Acht genommen werden. Dementsprechend sollte man bei der Erstellung der Text-Bild Kombination auf deren Beziehung achten.

Kategorie	Beschreibung
kongruent	Bild und Text transportieren inhaltlich dieselbe Information ohne Ergänzungen.
komplementär	Bild und Text ergänzen sich gegenseitig, indem zusätzliche Informationen vermittelt werden.
elaborativ	Bild oder Text vertiefen bzw. erweitern die Information der jeweils anderen Modalität.

Tab. 2: Bild und Text Beziehungen (Ballstaedt 2005: 66)

Die Beziehungen zwischen Bild und Text wie in Tabelle 2 (siehe: Ballstaedt 2005: 66) werden zwischen drei Bezügen unterschieden:

Kongruenz: Sie wiederholen den Textinhalt und erleichtern das Verständnis, indem sie Informationen visuell verstärken (Schnotz / Bannert 2003: 146).

Komplementarität Sie ergänzen den Text, indem sie zusätzliche Informationen liefern, die im Text nicht enthalten sind (Weidenmann 1994 a: 60ff.).

Elaboration: Diese Art von Bildern geht über den Text hinaus und fordert die Lernenden dazu auf, Zusammenhänge selbst zu erschließen (Ballstaedt 2005: 66; Holzbrecher 2006).

Bei der Untersuchung der Fachliteratur stellt sich heraus, dass es Studien zu den Lehrbüchern u.a. für Deutsch als Fremdsprache bezüglich Bild -Text Beziehung gibt. Verschiedene (Ballstaedt 2012, Karatekin / Seymen 2022, Pekrun 2006, Ünver / Genç 2013, Weidenmann 1994a) haben gezeigt, dass die Wahl und Gestaltung der Bilder in

den Lehrbüchern sorgfältig geplant werden muss, um den Lernprozess optimal zu unterstützen. Ünver und Genç (2013) kamen in ihrer Studie zu dem Schluss, dass Lehrer die visuellen Elemente in den Lehrbüchern als ausreichend empfinden. Allerdings bewerteten sie deren Eignung für die Schüler sowie deren Einfluss auf die Motivation der Schüler, die Entwicklung der Sprachkompetenz und die Vermittlung von Landeskunde/ Kulturunterricht als negativ. In der Studie von Karatekin und Seymen (2022) wurden die im Schuljahr 2020-2021 vom türkischen Bildungsministerium verpflichtend verwendeten deutschen Lehrbücher anhand von Prinzipien des visuellen Designs analysiert. In der Studie wurden drei für den Fremdsprachenunterricht Deutsch in der Sekundarstufe vorgeschriebene Lehrbücher durch eine Umfrage unter Deutschlehrkräften hinsichtlich ihrer visuellen Gestaltung bewertet. Dabei wurde festgestellt, dass das Lehrbuch Deutsch für Gymnasien A1.1 visuell am ansprechendsten ist, das Lehrbuch „Wie Bitte! A1.1“ in Bezug auf die visuelle Gestaltung weniger beeindruckend ist und das Lehrbuch „Deutsch macht Spaß A1.2“ in einigen visuellen Elementen effektiv ist.

In multimedialen Lernumgebungen wird Wissen häufig aus verschiedenen Quellen abgeleitet, die unterschiedliche Modalitäten ansprechen. Hierbei lassen sich insbesondere geschriebene bzw. gedruckte Lehrbuchtexte und abgebildete Illustrationen von digitalen oder dynamischen Medienformaten unterscheiden.

Es lässt sich also festhalten, dass Wissen im multimedialen Bereich aus verschiedenen Quellen abgeleitet wird: Einerseits aus gedruckten Texten und statischen Bildern, andererseits aus digitalen und dynamischen Darstellungen. Diese unterscheiden sich grundlegend in ihrer Rezeptionsweise und erfordern unterschiedliche Gestaltungsprinzipien, um Lernprozesse effektiv zu unterstützen.

Die Lücke an Studien über die Kohärenz zwischen Text und Bild in Fremdsprachenlehrbüchern hat diese Studie zu Forschungen in diesem Bereich veranlasst.

Ziel dieser Untersuchung

Im Rahmen dieser Studie wurde das Schülerbuch „Mit Deutsch“ auf dem Niveau A1.1 analysiert, das ab dem Schuljahr 2023/ 2024 vom türkischen Bildungsministerium verpflichtend zugelassen wurde. Ziel der Untersuchung ist es, die Kohärenz zwischen Bildern und Text im Lehrbuch zu bewerten und aus der Sicht der Lehrwerkanalyse zu ermitteln, inwiefern die bildliche Gestaltung den Text unterstützt und somit den Wissenserwerb erleichtert.

Dazu wird die Beziehung zwischen Bild und Text auf drei Ebenen betrachtet: (1) Ersteindruck, (2) Aktivierung von Vorwissen und (3) die konkrete Bild-Text-Verknüpfung. Diese Ebenen werden im Hinblick auf drei zentrale Beziehungsarten – Kongruenz, Komplementarität und Elaboration – untersucht. Die Untersuchung folgt einem qualitativen Forschungsansatz und wurde mittels Lehrwerkanalyse durchgeführt.

Methode

In diesem Abschnitt werden Informationen zur Untersuchung, zum Modell, der Datenerhebung und deren Analyse gegeben.

Forschungsmodell

In der Untersuchung wurde ausgehend von einem didaktischen Aspekt eine Lehrwerkanalyse verwendet, die bezüglich der Inhaltsanalyse eine der qualitativen Forschungsmethoden darstellt. Die Lehrwerkanalyse im allgemeinen Sinn umfasst die Analyse des Aufbaus des Lehrwerks, des Layouts, der Inhalte oder auch der Textsorten etc. (Kast / Neuner 1994).

Die Grenzen dieser Arbeit bilden die Begriffe Kongruenz, Komplementarität und Elaborativität der Texte. Die Auswahl dieser drei Kategorien begründet sich darin, dass sie in der multimodalen Lernforschung (vgl. Weidenmann 1990) als zentrale Dimensionen der Bild-Text-Beziehungen betrachtet werden und eine systematische Analyse der Interaktion zwischen visuellen und verbalen Elementen ermöglichen.

Im Rahmen des Forschungsdesigns wurde daher ein qualitativ-deskriptiver Ansatz gewählt, der sich auf die Analyse didaktischer Materialien stützt. Durch die Fokussierung auf diese drei Kategorien soll eine klare Abgrenzung geschaffen werden, die es ermöglicht, die Bild-Text-Beziehungen nicht nur oberflächlich zu beschreiben, sondern auch hinsichtlich ihrer funktionalen Rolle im Lernprozess zu untersuchen.

Die Begrenzung auf diese Kategorien dient gleichzeitig der Reduktion der Komplexität, worin Bild-Text-Relationen sehr vielfältig sein können. Damit wird eine analytische Tiefe gewährleistet, die eine Vergleichbarkeit zwischen verschiedenen Unterrichtsmaterialien zulässt. Das Forschungsdesign sieht daher vor, die Kategorien systematisch auf ausgewählte Texte anzuwenden, um Muster, didaktische Funktionen und mögliche Lernwirksamkeiten herauszuarbeiten.

Stichprobe

Die Stichprobe der Studie besteht aus den Texten und Abbildungen des Schülerbuchs „Mit Deutsch“ A1.1 Grundstufe für Deutsch als Fremdsprache, das seit dem Schuljahr 2023-2024 vom türkischen Bildungsministerium verpflichtend unterrichtet wird.

Datenanalyse

In der Studie wurde die deskriptive Analysetechnik zur Datenanalyse verwendet. In der Phase der Datenanalyse wurden die gewonnenen Daten unabhängig voneinander von zwei Forschern zu unterschiedlichen Zeitpunkten kodiert, um die Zuverlässigkeit zu gewährleisten. Auf diese Weise wurde versucht, die „Coder-Reliabilität“ (Miles / Huberman 1995) zu gewährleisten. Anschließend diskutierten die Forscher die Kodierung und trafen gemeinsam die endgültige Entscheidung, worin sie sich einig waren. In der

letzten Phase der Analysen wurden Anpassungen vorgenommen und die Kodierung wurde abgeschlossen.

Befunde

Im Folgenden werden die Ergebnisse über die Beziehung zwischen den Bildern und Texten dargestellt.

Bild	Erst eindruck	Aktivierung Vorwissen	Bild-Text Verbindu ng	Kongruent	Kompleme ntär	Elaborativ
1 s. 14	Alltagsszene, entspannter sozialer Austausch	Personenbes chreibung, Strukturen (Alter, Aussehen)	Text beschreibt detailliert drei Personen	Ja – Bild zeigt exakt, was im Text beschrieben wird	Ja – Bild vermittelt Atmosphär e, Mimik, Gestik	Nein – kein Zusatzwissen, alles im Text
2 s. 15	Geburtstagsfei er, weckt sofort Assoziationen	Geburtstagsr ituale	Bild entspricht dem Textinhalt	Ja – Torte entspricht Textinhalt	Ja – Bild verstärkt emotionale Wirkung	Nein – keine neuen Interpretatione n
3 s. 16	Teilweise stimmig, wirken wie Freunde, nicht wie Fremde	Alltagssituat ion, authentische Figuren	Gesprächss ituation im Bild, im Text. Alter, Adresse, Telefonnu mmer	Teilweise kongruent (Gesprächssitu ation) / Nicht kongruent (Alter, Adresse, Telefonnumm er)	Schwach kompleme ntär: Nähe im Bild → Dialog impliziert	Nein – keine zusätzlichen Informationen
4 s. 17	Positiv, sympathisch, Mia als aktive Figur	Sport, Freizeit, Jahreszeiten, Freundschaft , kulturelle Aktivitäten	Im Bild keine Aktivitäten aus dem Text erkennbar	Teilweise kongruent – Person (Mia) stimmt, Aktivitäten fehlen	Ja – Bild ergänzt Mias Charakter, emotionale Ausstrahlu ng	Nein – keine neuen Informationen oder interpretative Tiefe
5 s. 27	Schulischer Kontext, typische Szene im Unterricht.	Das Bild aktiviert schulische Alltagssituat ionen	Bild illustriert perfekt den Text (Mathe, rechnen).	Ja, Bild zeigt Azra beim Rechnen	Ja, Text ergänzt Info zur Lieblingsfa rbe, die im Bild nur angedeutet ist.	Nein- Kein zusätzlicher Informationsg ehalt durch das Bild.

6 S. 31	Zwei Jugendliche in einer Gesprächssituation, freundlich, ruhig, schulisch.	Neue Mitschüler kennenlernen, über Herkunft, Alter, Lieblingsfächer sprechen	Der Text ist ein Dialog mit vielen W-Fragen	Ja, eine Gesprächssituation	Nein, keine direkten Inhalte des Textes wie „Musik“ oder „Mathe“	Nein- Keine neuen Informationen im Bild, die über den Text hinausgehen.
7 S. 34	Freundliches Ehepaar, Glückliche Kleinfamilie	Typische Familiensituationen	Bild entspricht dem Textinhalt	Fast alle Bilder stimmen mit den Textinformationen überein	Ja, Die Bilder verstärken die emotionale Dimension	Es werden keine zusätzlichen inhaltlichen Informationen geliefert
8 S.37	authentische Alltagsszene, entspannter sozialer Austausch	Typische Personenbeschreibung; Bekannte Strukturen (Alter, Aussehen etc.).	Der Text beschreibt detailliert drei Personen.	Bild zeigt exakt, was im Text beschrieben wird (Kopftuch, Haarfarben etc.).	Ja, Bild vermittelt soziale Atmosphäre, Mimik, Gestik	Nein- Kein Zusatzwissen im Bild – alles Wesentliche steht im Text.
9 S.41	Familiäre Situation, Eltern mit Kind, häusliche Aktivität	fröhliche Familienszenen	Text beschreibt Tante Annika, ihren Mann Felix, Cousine Sandra und Cousin Tim → passt nicht vollständig zum Bild	Teilweise kongruent (Familie wird dargestellt, aber nicht die im Text genannte)	Geringe Ergänzung, Bild zeigt positive Familienatmosphäre	Keine zusätzlichen inhaltlichen Infos. Text liefert alle Details.
10 S.41	Drei unterschiedliche Personen werden vorgestellt (Ebru, Oliver, Ming).	Die Bilder zeigen Jugendliche aus verschiedenen Herkunftsländern (Deutschland, China), was kulturelle Vielfalt anspricht.	Die Porträtbilder passen zu den Namen und zum Inhalt der Texte.	Die Bilder bestätigen den Inhalt der Texte visuell, ohne neue Informationen hinzuzufügen.	Die Bilder tragen zur besseren Identifikation mit den Personen bei, liefern aber keine zusätzlichen Informationen über die	Nein- keine neuen Inhalte, die im Text nicht genannt werden.

					Essgewohnheiten.	
11 S. 64	Positiv und ansprechend. Die bunten Farben und die klare Struktur lenken die Aufmerksamkeit auf den Baum.	weckt durch die Darstellung des Baumes mit den verschiedenen Jahreszeiten-Symbolen Vorwissen über typische Merkmale der Jahreszeiten	Die Texte beschreiben in Gedichtform jede Jahreszeit mit typischen Eigenschaften	Der Baum visualisiert exakt das, was die Texte beschreiben	Die Texte liefern zusätzliche Informationen, die im Bild nicht sichtbar sind.	Ja, das Bild erweitert das Textverständnis, indem es den Inhalt visuell aufbereitet.
12 S. 66	Das Bild weckt sofort Assoziationen mit einer Geburtstagsfeier.	Das Bild aktiviert das Vorwissen über Geburtstagsrituale	Das Bild ist eng mit dem Inhalt des Textes verbunden und verdeutlicht visuell, was im Text beschrieben wird.	Ja, das Bild ist inhaltlich deckungsgleich mit dem Text, da beide sich auf eine Geburtstagstorte und eine Geburtstagsfeier beziehen.	Teilweise, das Bild ergänzt den Text, indem es eine konkrete visuelle Vorstellung der Torte liefert, die im Text erwähnt wird	Weniger, das Bild gibt keine zusätzlichen Informationen über den Kontext hinaus
13 S.69	Eine authentische Alltagssituation im Schulkontext.	Freundliche, interaktive Lernsituation.	Die Bild-Text-Kombination aktiviert themenbezogenes Vorwissen und schafft einen emotionalen Bezug.	Text und Bild zeigen dieselbe Information, Text stützt das Bild direkt.	Bild und Text ergänzen sich, um die Situation zu erklären.	Der Text liefert zusätzliche Information, die im Bild nicht visuell gezeigt wird.
14 S.72	Aktivitäten und Hobbys	Alltag, Kochen, Brettspiele, Musik...	Text und Bild sind eindeutig verknüpft	Handlung stimmt exakt mit Text überein	Ja, Bilder zeigen konkrete Handlungen	Keine zusätzliche interpretative Ebene

15 s. 74	Freundliche, sympathische Schüler	Jugendliche Aktivitäten	Text nicht direkt im Bild sichtbar	Alle drei Bilder sind kongruent, weil sie Alter, Geschlecht und Rolle (Schüler/ in) passend zum Text darstellen	Komplementär wirken sie, weil die Gesichtsausdrücke (freundlich, offen, jugendlich) die im Text genannten Interessen und Eigenschaften unterstreichen.	Elaborativ sind sie nicht, da die Bilder keine neuen Informationen (z. B. tatsächliche Darstellung von Hobbys wie Tennis, Kochen, Gitarre) hinzufügen
16 s. 78	Sammelobjekte bzw. Hobbys	Comics, Briefmarken und Freundschaftsbänder sammeln	Text beschreibt Sammeln und Lesen von Comics, sammeln von Briefmarken und Verschenken der Bänder	Bilder zeigen genau das, was Text beschreibt	Ja, Bild unterstützt emotionale Wirkung, zeigt die Tätigkeit praktisch und farbenfrohe Darstellung	Keine neuen Informationen über Text hinaus, nur visuelle Bestätigung
17 s.81	Positiv, sympathisch, lebensfrohe Persönlichkeit	Freizeit, Sportarten, Jahreszeiten, Kulturelle Aktivitäten	Das Bild zeigt Mia als Person, fröhlich, dynamisch – das passt zur Textaussage. Aber: Keine der genannten Aktivitäten ist im Bild konkret dargestellt	Teilweise – Person (Mia) stimmt, Aktivitäten nicht sichtbar	Ja – Bild ergänzt Mias Charakter, emotionale Ausstrahlung	Nein – keine neue Information oder interpretative Tiefe

18 S. 82	Lächelnde Jugendliche vermitteln Interesse an Mode	Mode als persönlicher Stil	Kleidung und Körperhaltung	Alle drei Bild-Text-Paare zeigen eine hohe Kongruenz, also direkte Übereinstimmung von Bild und Textinhalt.	Die kompletären Aspekte wie Mimik, Haltung und Accessoires tragen zusätzlich zur Charakterisierung bei	Elaborative Tiefe wird nicht erreicht – die Bilder erweitern den Text nicht, sie visualisieren ihn lediglich passend.
19 S. 84	Ersteindruck ist funktional und eindeutig: Bild und Text stehen in direkter, unterstützender Beziehung.	Sehr gute Aktivierung des Vorwissens – sowohl sprachlich als auch lebensweltlich.	Die Verbindung ist bei allen 16 genannten Kleidungsstücken klar und exakt.	Absolut kongruent: Jedes Kleidungsstück (z. B. Hose, T-Shirt, Gürtel, Handschuhe, Mütze...) wird explizit im Text genannt und ist bildlich eindeutig darstellbar	Bilder ergänzen den Text durch visuelle Anschaulichkeit.	Keine vertiefenden Bedeutungen oder neuen Informationen werden vermittelt.
20 S. 88	Typischer Verkaufsraum	Starke Aktivierung von Vorwissen und eindeutige Verknüpfung von Wortschatz mit anschaulichem Bildmaterial.	Text erwähnt das, was Bild zeigt	Sehr hoch, fast alle Bilder decken sich exakt mit den Textinformationen.	Vor allem atmosphärisch	Keine neuen inhaltlichen Informationen, Bilder bleiben rein unterstützend, nicht interpretativ
21 S. 91	Die Bilder sind klar und modern; sie zeigen junge Menschen mit deutlicher Kleidung	Aktivieren Vorwissen zu Kleidungsstücken	Die Texte beschreiben jeweils genau die Person im Bild. Kleidung, Farben und Accessoires	Text und Bild stimmen weitgehend überein	Einige zusätzliche Informationen sind nur im Text enthalten, nicht im Bild erkennbar	Es gibt elaborative Elemente, die den Lernenden helfen, über das reine „Sehen“ hinaus zusätzliche sprachliche

						und inhaltliche Informationen zu erschließen.
22 S.92	Wirkt historisch und offiziell, Symbol für Staatsgründer.	Vorwissen über Atatürk als Gründer und Präsident der Türkei.	Text und Bild sind eindeutig verknüpft	Alle Bilder sind kongruent mit den Textinformationen	Bilder ergänzen den Text sinnvoll	Durch die Bilder wird der Text nicht nur bestätigt, sondern auch vertieft

Tab 3: Tabellarische Darstellung der Bild-Text-Beziehungen im Schülerbuch „Mit Deutsch“ A1.1

Die Tabelle 3 enthält bereits die zweiundzwanzig Bild-Text-Beziehungen im Schülerbuch „Mit Deutsch“ A1.1 nach den Dimensionen Ersteindruck, Aktivierung von Vorwissen, Bild-Text-Verbindung, Kongruenz, Komplementarität und Elaboration. Die vorliegende Tabelle zeigt differenziert, wie Bilder und Texte im untersuchten Schülerbuch „Mit Deutsch“ A1.1 zueinander in Beziehung gesetzt werden. Dabei wird deutlich, dass die Bild-Text-Kohärenz auf verschiedenen Ebenen unterschiedlich stark ausgeprägt ist:

Der Ersteindruck: Der Ersteindruck der Bilder in der Tabelle zeigt, dass die Lehrbuchgestaltung stark auf Alltagsnähe, Authentizität und emotionale Ansprache gesetzt ist. Dies unterstützt sowohl die Motivation als auch die thematische Orientierung der Lernenden. Kritisch ist jedoch anzumerken, dass nicht alle Bilder gleichermaßen einen starken affektiven oder narrativen Ersteindruck auslösen – manche bleiben auf einer illustrativen Ebene stehen. Für eine optimale Wirkung sollten Bilder so gestaltet sein, dass sie unmittelbar Assoziationen und Emotionen aktivieren und damit den Spracherwerb nachhaltig unterstützen.

Aktivierung des Vorwissens: Viele Bild-Text-Paare aktivieren erfolgreich das Alltagswissen (z.B. Einkaufsstrassen, Familien- und Freundschaftssituationen). Dies unterstützt die anschlussfähige Verarbeitung bei den Lernenden, wobei bekannte Kontexte als Brücke zum Verstehen dienen. Allerdings variiert der Grad dieser Aktivierung: Während Alltagsszenen stark anschlussfähig sind, erfordern kulturell spezifische Inhalte (z. B. Rituale, Symbole) eine genauere didaktische Rahmung.

Bild- Text Verbindung: Die Tabelle zeigt, dass das Lehrbuch insgesamt auf eine hohe Bild-Text-Kohärenz setzt. Bilder illustrieren die Texte meist direkt oder ergänzen diese um emotionale bzw. kontextuelle Aspekte. Besonders wertvoll sind die Fälle, in denen Bild und Text komplementär wirken, wodurch sich eine mehrdimensionale Bedeutung entwickelt. Kritisch zu bewerten sind dagegen Stellen, an denen Bild und Text nur lose gekoppelt sind, da dies den didaktischen Mehrwert einschränkt.

Kongruenz: In einem Großteil der Beispiele besteht eine hohe inhaltliche Übereinstimmung zwischen Bild und Text. Dies bedeutet, dass die Bilder im Wesentlichen dieselben Informationen transportieren wie die Texte (z. B. Personenbeschreibung, Alltagsszenen, konkrete Handlungen). Diese Kongruenz erleichtert nach Mayer (2009) die kognitive Verarbeitung, welche dieselbe Information in mehreren Modalitäten präsentiert und somit das Textverständnis stützt.

Komplementarität: Neben den kongruenten Beziehungen zeigt sich auch eine Reihe komplementärer Ergänzungen: Bilder erweitern die Textinformationen, indem sie etwa Atmosphäre, Mimik, Gestik oder emotionale Dimensionen hinzufügen. Diese Ebene ist didaktisch besonders wertvoll. Dadurch ermöglichen sie den Lernenden eine tiefere Kontextualisierung, die über die rein inhaltliche Redundanz hinausgeht.

Elaboration: Auf der elaborativen Ebene fällt jedoch auf, dass nur wenige Beispiele zusätzliche Wissensdimensionen liefern. Meist bleibt das Bild auf die Funktion der Visualisierung des Textinhalts oder der emotionalen Ergänzung beschränkt, ohne neue inhaltliche Informationen oder interpretative Tiefe zu schaffen. Damit bleibt das Potenzial, durch Bilder erweiterte Lernräume zu eröffnen, ungenutzt.

Die Tabelle 3 zeigt somit, dass die Bild-Text-Kombinationen überwiegend kohärent und lernunterstützend gestaltet sind, wobei die Stärken in der Kongruenz und Komplementarität liegen. Die Schwäche ist jedoch die seltene Nutzung elaborativer Bild-Text-Beziehungen, die das Potenzial hätten, tiefere Verstehens- und Lernprozesse anzuregen. Insgesamt wird somit bestätigt, dass Bild-Text-Kohärenz ein tragendes Prinzip der Lehrbuchgestaltung darstellt, das jedoch durch eine bewusstere Integration elaborativer Dimensionen noch optimiert werden könnte.

Aufgrund der großen Anzahl von Texten mit Bildmaterial im Buch wurden für diese Arbeit drei Beispiele nach den Dimensionen der Text-Bild-Beziehung für die Untersuchung ausgewählt.



Abb. 1: Beispiel aus dem Schülerbuch "Mit Deutsch" S.37.

In Abbildung 1 sind drei Erwachsene in einem freundlichen Gespräch zu sehen. Eine Frau trägt ein Kopftuch, neben ihr sitzt eine Frau mit Brille und hat helle Haare, der Mann wirkt entspannt und hört aufmerksam zu. Der Text beschreibt drei Personen aus der Ich-Perspektive: Mutter (Funda), Frau Müller (Sandra) und Herr Müller (Alex).

Ersteindruck: authentische Alltagsszene, entspannter sozialer Austausch, womöglich Nachbarn oder Freunde.

Aktivierung des Vorwissens: Lernende können ihr Vorwissen zu Personenbeschreibungen, äußerem Erscheinungsbild und Familienverhältnissen aktivieren:

-Wer ist wer?; -Wie beschreibt man Alter, Haarfarbe, Augenfarbe?; -Wie sind die Verhältnisse zwischen den Personen?

Bekannte grammatische Strukturen („ist...“) und typische Wortfelder („groß“, „schlank“, „braun“) werden benutzt – das hilft beim Wortschatzerwerb und beim Lesen mit Bildunterstützung.

Text-Bild-Verbindung: Der Text beschreibt detailliert drei Personen. Das Bild zeigt diese – zumindest laut Aufgabenstellung – als Illustration der Beschreibungen.

Kongruente Beziehung:

- Ja, das Bild bestätigt direkt Teile des Textes:
- Eine Frau trägt ein Kopftuch → das ist die Mutter, Funda.
- Eine Frau ist blond mit Brille → das ist Sandra Müller.
- Ein Mann ist groß, mit kurzen braunen Haaren → das ist Alex Müller.
- Die äußeren Merkmale (Kopftuch, Haarfarben, Brille, Geschlecht) stimmen visuell mit der Textbeschreibung überein → klare kongruente Beziehung.

Komplementäre Beziehung:

- Das Bild ergänzt durch nonverbale Kommunikation: die Personen lächeln, wirken zugewandt, was im Text nicht explizit gesagt wird.
- Der soziale Kontext (freundschaftlich, kommunikativ) wird durch das Bild klarer unterstützt aber das Verständnis, ist nicht wörtlich im Text vorhanden.

Elaborative Beziehung:

- Es gibt keine zusätzlichen Informationen im Bild, die über den Text hinausgehen.

- Wir erfahren nichts Neues über das Alter, Beruf, Wohnsituation etc. – das bleibt ausschließlich textbasiert.

Ebene	Beschreibung
Ersteindruck	Alltagsszene, entspannter sozialer Austausch
Aktivierung/Vorwissen	Typische Personenbeschreibung; Bekannte Strukturen (Alter, Aussehen)
Text-Bildverbindung	Der Text beschreibt detailliert drei Personen.
Kongruent	✅ Bild zeigt exakt, was im Text beschrieben wird (Kopftuch, Haarfarbe)
Komplementarität	🗨️ Bild vermittelt soziale Atmosphäre, Mimik, Gestik → ergänzt den Text
Elaboration	❌ Kein Zusatzwissen im Bild – alles Wesentliche steht im Text.

Tab. 4: Tabellarische Darstellung der Textanalyse

Wie man aus der Tabelle 4 herauslesen kann, eignet sich Aufgabe 6a auf Seite 37 zur Sprachförderung im Bereich Personenbeschreibung. Die kongruente Bild-Text-Relation unterstützt das Textverständnis, während die komplementären Anteile (Körpersprache, soziales Miteinander) das Lernen in einem realitätsnahen Kontext verankern.

🗨️ Lies den Text! Wann hat Anja Geburtstag? Antwortet in der Klasse!

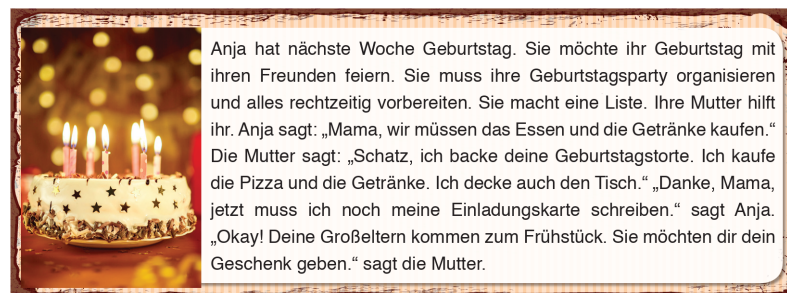


Abb. 2: Beispiel aus dem Schülerbuch „Mit Deutsch“ S.66.

Der Text in Abbildung 2 berichtet von Anjas bevorstehendem Geburtstag und den Vorbereitungen für die Feier, und das Bild zeigt eine festlich dekorierte Geburtstagsstorte mit Kerzen.

Ersteindruck: Das Bild weckt sofort Assoziationen mit einer Geburtstagsfeier. Es unterstützt den Text in seiner fröhlichen und festlichen Stimmung und vermittelt dem Betrachter unmittelbar das Thema „Geburtstag“.

Aktivierung des Vorwissens: Typische Aktivitäten wie Kuchenbacken, Einladungen schreiben, Tischdecken und Geschenke erhalten werden beschrieben. Die abgebildete Torte mit Kerzen ist ein starkes Symbol für Geburtstagsfeiern. Das Bild aktiviert das Vorwissen der Betrachter über Geburtstagsrituale – sie denken automatisch an Feiern, Kuchen, Freunde und Geschenke. Dadurch wird das Textverständnis gefördert.

Text-Bild-Verbindung: Die Mutter spricht konkret davon, Anjas Geburtstagsstorte zu backen. Das Bild zeigt eine Torte mit Kerzen, was genau zu dieser Textstelle passt. Das Bild ist eng mit dem Inhalt des Textes verbunden und verdeutlicht visuell, was im Text beschrieben wird.

Kongruente Beziehung: Ja. Das Bild ist inhaltlich deckungsgleich mit dem Text, da beide sich auf eine Geburtstagstorte und eine Geburtstagsfeier beziehen.

Komplementäre Beziehung: Teilweise. Das Bild ergänzt den Text, indem es eine konkrete visuelle Vorstellung der Torte liefert, die im Text erwähnt wird.

Elaborative Beziehung: Weniger. Das Bild gibt keine zusätzlichen Informationen über den Kontext hinaus, sondern illustriert nur einen Aspekt (die Torte), während der Text mehr Details zur Vorbereitung der Feier liefert.

Ebene	Beschreibung
Ersteindruck	Das Bild weckt sofort Assoziationen mit einer Geburtstagsfeier.
Aktivierung/Vorwissen	Das Bild aktiviert das Vorwissen über Geburtstagsrituale.
Text-Bildverbindung	Das Bild ist eng mit dem Inhalt des Textes verbunden.
Kongruent	✓ Ja – Torte aus dem Bild entspricht Textinhalt.
Komplementarität	🔗 Ja- Bild verstärkt die emotionale Wirkung.
Elaboration	✗ Keine neue Interpretationsebene.

Tab. 5: Tabellarische Darstellung der Textanalyse

Das Bild unterstützt den Text auf kongruente und teilweise komplementäre Weise, aber es ist nicht elaborativ. Die Text-Bild-Relation ist insgesamt stimmig und fördert das Textverständnis durch emotionale und visuelle Aktivierung.

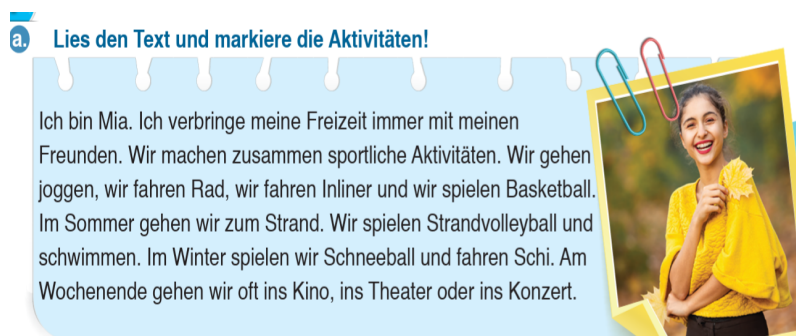


Abb. 3: Beispiel aus dem Schülerbuch “Mit Deutsch” S.81.

In Abbildung 3 erzählt Mia, dass sie ihre Freizeit mit Freunden verbringt und zahlreiche sportliche und kulturelle Aktivitäten unternimmt: joggen, Rad fahren, schwimmen, Volleyball spielen, ins Kino/ Theater/ Konzert gehen.

Das Bild zeigt ein junges Mädchen (vermutlich Mia), das im Freien steht, Herbstblätter hält und ein gelbes Oberteil trägt. Das Bild wirkt positiv, lebendig, naturverbunden.

Ersteindruck: Der Ersteindruck ist freundlich und aktivitätsnah, vermittelt eine aktive, lebensfrohe Persönlichkeit, was mit dem Text übereinstimmt.

Aktivierung des Vorwissens: Der Text aktiviert viele bekannte Begriffe aus dem Freizeit- und Sportbereich, die Schüler aus dem eigenen Alltag kennen:

- Sportarten: Basketball, Radfahren, joggen etc.
- Jahreszeitenbezug: Sommer / Winter
- Kulturelle Aktivitäten: Kino, Theater, Konzert

Auch das Bild aktiviert saisonales Vorwissen (Herbststimmung durch Laub) und emotionale Identifikation mit Mia. Es gibt eine sehr gute Aktivierung sprachlicher und situativer Kontexte.

Kongruente Beziehung: Nur eingeschränkt kongruent: Das Bild zeigt Mia als Person, fröhlich, dynamisch – das passt zur Textaussage. Aber: Keine der genannten Aktivitäten ist im Bild konkret dargestellt (kein Sport, keine Freunde, keine Aktion).

Komplementäre Beziehung:

- Das Bild ergänzt den Text durch visuelle Darstellung von Mias Persönlichkeit: sympathisch, offen, lebensfroh.
- Es unterstützt das, was implizit im Text beschrieben wird (Mia als aktive, gesellige Jugendliche).
- Auch die Herbststimmung (farbige Blätter) passt indirekt zu den im Text erwähnten saisonalen Aktivitäten.

Elaborative Beziehung

- Keine zusätzliche Bedeutungsebene wird eröffnet.
- Das Bild veranschaulicht nicht die konkreten Aktivitäten oder Orte – es bleibt rein symbolisch/ porträtartig.

Ebene	Beschreibung
Ersteindruck	Positiv, sympathisch, zeigt Mia als aktive fröhliche Figur
Aktivierung/Vorwissen	Ja- Sport, Freizeit, Jahreszeiten, Freundschaft, kulturelle Aktivitäten
Text-Bildverbindung	Im Bild sind keine kulturelle Aktivitäten aus dem Text erkennbar.
Kongruent	 Teilweise- Person (Mia) stimmt, Aktivitäten nicht sichtbar
Komplementarität	 Ja- Bild ergänzt Mias Charakter, emotionale Ausstrahlung
Elaboration	 Keine neue Informationen oder interpretative Tiefe.

Tab. 6: Tabellarische Darstellung der Textanalyse

Das Bild dient nicht zur Veranschaulichung konkreter Inhalte, sondern stellt Mia als Sympathiefigur dar, die Lernende emotional ansprechen soll. Es ergänzt den Text sinnvoll im Hinblick auf Stimmung und Identifikation, erfüllt aber keine kongruente oder elaborative Funktion im engeren Sinne.

Schlussfolgerung

Die vorliegende Analyse zeigt anhand verschiedener Ebenen und Beschreibungen, dass eine gezielte und didaktisch fundierte Gestaltung von Bild-Text-Kohärenz ein zentraler Bestandteil des Lehrbuchs ist. Insbesondere im Lehrwerk *Mit Deutsch A1.1* konnte nachgewiesen werden, dass Bild-Text-Kombinationen häufig erfolgreich eingesetzt werden, um den thematischen Einstieg zu erleichtern, Vorwissen zu aktivieren und das Textverständnis nachhaltig zu fördern. Dabei dominieren kongruente Bild-Text-Beziehungen, bei denen Bild und Text dieselben Informationen vermitteln. Diese Form der Kongruenz unterstützt die Lernenden, indem Informationen parallel über unterschiedliche Modalitäten – visuell (Bild) und verbal (Text) – dargeboten werden. Laut Mayer (2009) erleichtert diese multimodale Informationsverarbeitung die kognitive Integration, da Lernende durch das sogenannte „Multimedia-Prinzip“ von der gleichzeitigen Präsentation redundanter Inhalte profitieren.

Der gezielte Einsatz mehrerer Modalitäten trägt somit nicht nur zur Entlastung des Arbeitsgedächtnisses bei, sondern fördert auch eine tiefere Verarbeitung, wodurch Bild und Text im mentalen Modell miteinander verknüpft werden. Dies verdeutlicht, dass das Lehrwerk *Mit Deutsch A1.1* das Prinzip der Modalitäten konsequent nutzt, um Lernprozesse zu optimieren und ein höheres Maß an Verständnissicherung zu erreichen.

Komplementäre Beziehungen – bei denen Bild und Text sich ergänzen – wurden ebenfalls beobachtet und wirken sich nachweislich positiv auf den Wissenserwerb aus, da sie unterschiedliche Aspekte einer Information abdecken (Schnotz / Bannert 2003). Hingegen wurden elaborative Bild-Text-Beziehungen, welche vertiefte kognitive Prozesse auslösen könnten, im untersuchten Material nicht festgestellt. Diese Form der Verknüpfung gilt jedoch als besonders lernwirksam, da sie dazu beiträgt, abstrakte Inhalte in mentalen Modellen zu verankern (Seufert 2003).

Zugleich wurden auch gestalterische Defizite in der funktionalen Verknüpfung von Bild und Text identifiziert. Diese Mängel unterstreichen die Bedeutung einer bewusst geplanten visuellen Gestaltung bei der Entwicklung eines Lehrbuchs. Bilder sollten nicht rein dekorativer Natur sein, sondern in engem inhaltlichem Bezug zum Text stehen, um einen Mehrwert für den Lernprozess zu bieten (Levin 1981) und durch ihre Wechselwirkung bereichernde Kommunikationsformen ermöglichen (Straßner 2002). Wie Schiffler (1975) betont, erhöht sich die kommunikative Kraft eines Bildes proportional zu seinem Detailreichtum – besonders dann, wenn es vielfältige Interpretationsmöglichkeiten bietet und so dialogorientierte, sprachproduktive Prozesse im Unterricht initiiert.

Auch die sprachlichen Voraussetzungen der Lernenden müssen bei der Auswahl der Bilder berücksichtigt werden. Nur wenn die Lernenden über die sprachlichen Mittel verfügen, um sich zu einem Bild äußern zu können, wird der intendierte kommunikative Nutzen erreicht. Offene, mehrdeutige Bilder fördern Sprachproduktion, regen zum Hypothesenbilden an und bieten zahlreiche Ansatzpunkte zur Erweiterung des Wortschatzes sowie zur Anwendung grammatikalischer Strukturen (vgl. Schiffler 1975).

Für zukünftige Lehrbuchkonzeptionen lässt sich daraus ableiten, dass eine stärkere Berücksichtigung der Bild-Text-Kohärenz auf allen drei Ebenen – nämlich der affektiven (Ersteindruck), der kognitiven (Vorwissensaktivierung) und der semantischen (Text-Bild-Verbindung) – wünschenswert ist. Insbesondere sollte die elaborative Funktion von Bildern stärker berücksichtigt werden, da sie das Potenzial hat, tiefere Verständnisprozesse anzustoßen. Eine ausgewogene und strategische Integration visueller Elemente kann nicht nur das Textverständnis verbessern, sondern auch die Motivation und Lernfreude der Lernenden erheblich steigern (Mayer 2009, Schnotz 2005).

Zusammenfassend lässt sich aus dieser begrenzten Analyse ableiten, dass Bild-Text-Kohärenz im Fremdsprachenunterricht nicht nur eine methodische Ergänzung darstellt, sondern eine zentrale Grundlage für den Wissenserwerb und die Förderung der Sprachkompetenz im Fremdsprachenunterricht bildet. Die Analyse zeigt, dass kohärent gestaltete Bild-Text-Kombinationen das Verständnis komplexer sprachlicher Strukturen erleichtern, indem die Verknüpfung von visuellen und sprachlichen Informationen ermöglichen und so die kognitive Verarbeitung optimieren. Kongruente Beziehungen tragen dabei wesentlich zur Wortschatzarbeit und Strukturvermittlung bei, während komplementäre und elaborative Elemente insbesondere die kommunikative und interkulturelle Kompetenz stärken. Damit wird deutlich, dass Bild-Text-Kohärenz eine didaktisch wirksame Strategie ist, um Sprachverwendung und Sprachverständnis im Unterricht zu fördern. Sie leistet somit einen nachhaltigen Beitrag zur Entwicklung der sprachlich-kommunikativen Handlungsfähigkeit der Lernenden und unterstreicht die Relevanz multimedialer Ansätze für einen modernen, kompetenzorientierten Fremdsprachenunterricht.

Literaturverzeichnis

- Ballstaedt, Steffen-Peter** (2005): Text-Bild Kompositionen im Unterrichtsmaterial. In: *Der Deutschunterricht* 57 (4), 61-70.
- Ballstaedt, Steffen-Peter** (2012): Visualisieren: *Bilder in wissenschaftlichen Texten*. UVK Verlagsgesellschaft. Konstanz und München.
- Berardo, Sacha Anthony** (2006): The use of authentic materials in the teaching of reading. In: *The Reading Matrix*, 6(2), 60-69.
- GER** (2020): *Reference for Languages: Learning, teaching, assessment – Companion volume*, Council of Europe Publishing, Strasbourg, available at www.coe.int/lang-cefr.
- Drewniak, Ute** (1992): Lernen mit Bildern in Texten. *Untersuchung zur Optimierung des Lernerfolgs bei Benutzung computerpräsentierter Texte und Bilder*. Münster: Waxmann.
- Dwyer, Francis M.** (1978): *Strategies for improving visual learning*. State College, Pa.: Learning Services.
- Ehlich, Konrad** (2005): Sind Bilder Texte? In: *Der Deutschunterricht. Beiträge zu seiner Praxis und wissenschaftlicher Grundlegung*. Heft 4, 51-60.
- Eitel, Alexander / Scheiter, Katharina** (2015): Picture or text first? Explaining sequence effects when learning with pictures and text. In: *Educational Psychology Review*, 27(1), 153–180.

- Engelkamp, Johannes** (1998): Gedächtnis für Bilder. In: Sachs-Hombach, Klaus / Rehkämpfer, Klaus (Hrsg.): *Bild - Bildwahrnehmung - Bildverarbeitung. Interdisziplinäre Beiträge zur Bildwissenschaft*. Wiesbaden: Deutscher Universitäts-Verlag, 227-242.
- Heyd, Gertraude** (1991): *Deutsch lehren: Grundwissen für den Unterricht in Deutsch als Fremdsprache*. 2. überarbeitete und erweiterte Aufl. Frankfurt am Main: Moritz Diesterweg.
- Holzbrecher, Alfred** (2006): Foto+ Text: Didaktische Perspektiven. In: Holzbrecher, Alfred / Oomen-Welke, Ingelore / Schmolling, Jan (Hrsg.): *Foto+Text Handbuch für die Bildungsarbeit*. Wiesbaden: VS Verlag für Sozialwissenschaften GWV Fachverlage GmbH, 15-23.
- İşeri, Kamil** (2007): Altıncı sınıf Türkçe ders kitabının ilköğretim Türkçe programının amaçlarına uygunluğunun değerlendirilmesi. In: *Dil Dergisi*, 136, 58-74.
- Karatekin, Pınar/ Seymen, Aylin** (2022): Almanca ders kitaplarının görsel tasarım ilkeleri açısından değerlendirilmesi. In: *RumeliDE Dil ve Edebiyat Araştırmaları Dergisi*, (31), 1044-1058. DOI: 10.29000/rumelide.
- Kast, Bernd, / Neuner, Gerd** (1994): *Zur Analyse, Begutachtung und Entwicklung von Lehrwerken für fremdsprachlichen Deutschunterricht*. Langenscheidt. Berlin.
- Kümmerling-Meibauer, Bettina** (2018): *The Routledge Companion to Picturebooks*. Routledge.
- Levie, Howard W. / Levin, Joel R.** (1982): *Effects of text illustrations: A review of research*. In: D. F. Jonassen (Hg.): *Educational communications and technology: A review of the research*. Association for Educational Communications and Technology, 233-288.
- Levin, Joel R.** (1981): On Functions of Pictures in Prose. In: Pirozzola, Francis J./ Wittrock, Merlin C. (Hrsg): *Neuropsychological And Cognitive Processes In Reading*. Academic Press, New York, 203-228.
- Makowska, Magdalena** (2013): Im multimodalen Dialog. Zum Zusammenspiel von Text und Bild auf den Einbänden von Kinderbüchern. In: *Convivium. Germanistisches Jahrbuch Polen*, 115–142. <https://doi.org/10.18778/2196-8403.2013.07>
- Mayer, Richard E. / Moreno, Roxana** (2003): Nine ways to reduce cognitive load in multimedia learning. In: *Educational Psychologist*, 38(1), 43-52.
- Mayer, Richard E.** (2009): *Multimedia Learning* (2nd ed.): Cambridge University Press.
- Miles Matthew B. / Huberman, Michael** (1995): *Qualitative data analysis: An expanded sourcebook*. 2nd Edition. Thousand Oaks. CA: Sage Publications. London.
- Milli Eğitim Bakanlığı Ders Kitapları ve Eğitim Araçları Yönetmeliği** (2024): *Resmi Gazete* (15.10.2024). Sayı: 32693.
- Mentiş Taş, Ayşe** (2007): Yeni sosyal bilgiler ders kitaplarına ilişkin öğretmen görüşlerinin belirlenmesi. In: *Selçuk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 17, 519-532.
- Paivio, Allan** (1986): *Mental Representations: A Dual Coding Approach*. Oxford University Press. New York
- Pekrun, Reinhard** (2006): The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. In: *Educational Psychology Review*, 18(4), 315–341. <https://doi.org/10.1007/s10648-006-9029-9>.
- Rentel, Nadine** (2011): Die Didaktisierung von Phraseologismen im DaF-Unterricht anhand multimodaler Texte. In: *Linguistik online*, 47(3), 54-66.
- Schiffler, Ludger** (1975): Wie sollen Bilder im Fremdsprachen- Anfangsunterricht gestaltet sein? Empirische Überprüfungen der Sematisierungsfähigkeit verschiedener Bildkonzeptionen. In: *Praxis des neusprachlichen Unterrichts* 22, 65-79.

- Schnotz, Wolfgang** (1994): Wissenserwerb mit logischen Bildnern. In: Weidenmann, Bernd (Hrsg.): *Wissenserwerb mit Bildern: Instruktionale Bilder in Printmedien, Film/Video und Computerprogrammen*. Bern: Huber Verlag, 95–147.
- Schnotz, Wolfgang** (2005): Multimedia Learning: Understanding with Text and Pictures. In: Nunes, T. / Bryant, P. (Hg.): *Handbook of Educational Psychology*. Elsevier.
- Schnotz, Wolfgang / Bannert, Maria** (2003): Construction and interference in learning from multiple representation. In: *Learning and Instruction*, 13(2), 141-156.
- Seufert, Tina** (2003): Supporting coherence formation in learning from multiple representations. In: *Learning and Instruction*, 13(2), 227–237.
- Siever, Christina Margrit** (2014): Multimodale Kompetenz und multimodale Kommunikation im DaF-Unterricht. In: Hieronimus, Marc (Hg.): *Visuelle Medien im DaF-Unterricht*. Göttingen, Universitätsverlag, 381-400.
- Straßner, Erich** (2002): *Text-Bild-Kommunikation-Bild-Text-Kommunikation*. Tübingen: Niemeyer. (Grundlagen der Medienkommunikation, Bd. 13).
- Sweller, John** (1994): Cognitive load theory, learning difficulty, and instructional design. In: *Learning and Instruction*, 4(4), 295-312.
- Ünver, Şerife / Genç, Ayten** (2013): Görselliğin Gücü: Almanca Ders Kitaplarındaki Görselliğe İlişkin Öğretmen Görüşleri. In: *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi* 28(1), 393-404.
- Weidenmann, Bernd** (1990): Wissenserwerb in Bildern. *Forschung für eine visuelle Lernkultur. Unterrichtswissenschaft Zeitschrift für Lernforschung* 18/1, 62-66. DOI: 10.25656/01:29666
- Weidenmann, Bernd** (1994a): Informierende Bilder. In: Weidenmann, Bernd. (Hrsg.): *Wissenserwerb mit Bildern. Instruktionale Bilder in Printmedien, Film/Video und Computerprogrammen*. Bern: Huber Verlag, 9-58.
- Weidenmann, Bernd** (1994b): *Lernen mit Bildmedien: psychologische und didaktische Grundlagen*. Weinheim/Basel: Beltz Weiterbildung.
- Wildfeuer, Janina / Bateman, John A. / Hiippala, Tuomo** (2020): *Multimodalität: Grundlagen, Forschung und Analyse–Eine problemorientierte Einführung*. Walter de Gruyter GmbH & Co KG. Berlin.
- Van Gog, Tamara** (2014): The signaling (or cueing) principle in multimedia learning. In: Mayer, Richard (Hrsg.): *The Cambridge Handbook of Multimedia Learning*. Cambridge University Press. England, 263-278.