

**Katarzyna Sierak**

Szkoła Doktorska Uniwersytetu Warmińsko-Mazurskiego w Olsztynie

<https://orcid.org/0000-0002-6559-0692>

[katarzyna.sierak@uwm.edu.pl](mailto:katarzyna.sierak@uwm.edu.pl)

## *Logovisuelle Didaktik im Fremdsprachenunterricht*

### Logovisual didactic in a foreign language lesson

The multimodal texts in the form of text-image combinations have become an integral part of everyday life. We can notice them at every turn, and the best examples of it are the magazines, posters or the internet. Their omnipresence, which is primarily the result of the rapid development and spread of the new media, contributes to the fact that they can no longer be overlooked in the case of foreign language lessons either. Its huge didactic potential, which is the result of the semiotic division of labour of language-image connections, can not only create new opportunities to promote different language-related and non-language-related skills in foreign language learners but also increase the effectiveness of the process of teaching and learning foreign languages. The prerequisite, however, is that the learner can adequately analyze and understand them.

The aim of this article is to make an overview of the forms and functions of different text-image combinations in the foreign language lesson and bring closer the assumptions of logovisual didactic, which assumes the development of special *visual-reading literacy* to confrontation with them.

Keywords: multimodal texts, text-image-combinations, logovisual didactic

**Słowa kluczowe:** teksty multimodalne, kombinacje tekstowo-obrazowe, dydaktyka logowizualna

## 1. Einführung

Bilder begleiten die Menschen seit Anbeginn der Zeit. Als das erste Kommunikationsmittel traten sie viele Jahrtausende vorher auf, bevor sich der Mensch bewusst eines geordneten Schriftsystems, des Alphabets, zu bedienen begann (z. B. Höhlenmalereien, Hieroglyphen, Petroglyphen etc.). Obwohl die Sprache mit der Zeit zum dominantesten Medium wurde<sup>1</sup>, das die Art und Weise revolutionierte<sup>2</sup>, in der wir kommunizieren, in der wir Informationen weitergeben und speichern und in der wir lernen, verliert sie gegenwärtig zugunsten des Visuellen erneut an Bedeutung (vgl. Łosiewicz, 2009: 205). Diese mit der Zeit in der Mediengeschichte vollzogenen Veränderungen erklärt Assmann (2004) folgendermaßen:

Zwei Wenden kennzeichnen die Mediengeschichte, die aber nicht revolutionär, sondern evolutionär verlaufen: Die erste ist die Wende von den Dingen selbst zu den Bildern der Dinge, den Hieroglyphen. Das wäre der erste *iconic turn*. Die zweite ist die Wende von der Bilderschrift der Hieroglyphen zur abstrakten phonetischen Alphabetschrift. Das wäre der *aniconic turn*. Und davon ausgehend würde sich nun die Frage stellen, ob das Pendel jetzt in der Gegenrichtung, also weg von der Alphabetschrift und dem damit verbundenen Buchstabenmenschentum, wieder zurück zu den Bildern schwingt. (Assmann, 2004: 387; Kursivschrift im Original)

Die von Assmann angedeutete Situation widerspiegelt den derzeitigen kommunikativen Wandel allerdings nicht ganz. Das, was das gegenwärtige Interesse an Bildern von der ersten genannten Wende unterscheidet, ist nämlich die Tatsache, dass es sich nicht nur auf Bilder selbst fokussiert und darauf beschränkt. Die gegenwärtige Wende besteht eher in einem wachsenden Bewusstsein für die Rolle von Bildern im Zusammenhang mit Sprache und anderen Elementen im Rahmen eines Kommunikates, in ihrer Simultanität und Omnipräsenz im öffentlichen Diskurs. In der deutschsprachigen Forschungsliteratur wird daher dafür plädiert, in diesem Falle statt von weit verbreiteten Formulierungen, wie *pictorial turn* (Mitchell, 1992) oder *iconic turn* (Boehm, 1994), die eng mit der Zeichenmodalität – dem Bild selbst – verknüpft sind, eher von einem *intermedialen turn* (vgl. Hedling, Lagerroth, 2002: 7; Wolf, 2002: 15) oder einem

---

<sup>1</sup> Jäger (2002) bezeichnet die Zeichenmodalität – Sprache ein „Archimedium“ (vgl. ebd. S. 34).

<sup>2</sup> Moses Mendelssohn drückte sich im Jahre 1783 darüber folgendermaßen aus: „Wir sind Buchstabenmenschen. Vom Buchstaben hängt unser ganzes Wesen ab“ und fügte weiter hinzu: „Wir lehren und unterrichten einander nur in Schriften, lernen die Natur und die Menschen kennen nur in Schriften, arbeiten und erholen, erbauen und ergötzen uns durch Schreibung [...]“ (Mendelssohn, 1989: 421f).

*multimodalen turn* (vgl. Bucher, 2012: 54; Raith, 2014: 25; Wildfeuer u.a., 2020: 17) zu sprechen, da diese Begriffe präziser die sich im Moment abspielenden kommunikativen Veränderungen wiedergeben.

Im vorliegenden Beitrag wird die Rolle der *Multimodalität* im Rahmen einer sich neu entwickelnden logovisuellen Didaktik im Fremdsprachenunterricht reflektiert. Der Fokus liegt dabei auf dem am häufigsten vorkommenden Typ von multimodalen Texten, und zwar auf der Verknüpfung von statischem Bild und Sprache. Es wird versucht, folgenden Fragen nachzugehen:

- Was ist Multimodalität?
- Wie wirken Bild und Text<sup>3</sup> zusammen?
- Sind multimodale Texte immer mühelos zu verstehen?
- Welcher Kompetenzen bedarf es, um Text-Bild-Kombinationen angemessen zu interpretieren? Reicht dafür eine Lesekompetenz aus?
- Welche Formen und Funktionen können Text-Bild-Kombinationen einnehmen und welchen Nutzen kann die DaF-Didaktik daraus ziehen?

Die in Kapitel 2 und 3 angestellten theoretischen Überlegungen münden in einen anwendungsorientierten Teil, in dem veranschaulicht wird, wie mit Hilfe unterschiedlicher Formen von Text-Bild-Kombinationen im Fremdsprachenunterricht einige sprach- und nichtsprachbezogene Kompetenzen gefördert werden können.

## 2. Multimodale Prägung der Kommunikation

Ein charakteristisches Merkmal der gegenwärtigen Kommunikation ist ihre Neigung zur semiotischen Differenzierung. Im öffentlichen Diskurs begegnet man immer seltener reinen Bildern ohne Text oder reinen Texten ohne Bild. Monomodale, schriftliche Texte verlangen von ihren Rezipient\*innen eine zeitintensivere Rezeption, was sich bei der herrschenden Informationsüberlastung sowie einer steigenden Anzahl von Personen mit Konzentrationsschwäche nicht besonders gut bewährt. Viel häufiger handelt man stattdessen mit semiotisch ausdifferenzierten Kommunikaten, die imstande sind, die zu überliefernde Botschaft für die Rezipient\*innen zugänglicher und leichter erfassbar zu machen (vgl. Schmitz, 2011: 25). Gerade diejenigen „Texte und kommunikative Handlungen, die mehrere verschiedene Zeichensysteme (Sprache, Bild, Ton) beinhalten“ (Stöckl, 2011: 45), werden mit dem vagen Begriff *Multimodalität* bezeichnet.

---

<sup>3</sup> Nach Opiłowski (2015: 58) wird im vorliegenden Beitrag angenommen, dass nur sprachliche Ressourcen über einen Textstatus verfügen können. Aus diesem Grunde werden im Weiteren diese beiden Begriffe (Sprache und Text) stets simultan verwendet.

Dieser Begriff gilt als vage, weil er unterschiedliche Bedeutungen umfasst. Welche Assoziationen er überhaupt hervorrufen kann, gibt Stöckl (2011) an, indem er für folgende Definitionsbereiche dieses Begriffs plädiert:

1. Multimodalität ist die Kopräsenz und wechselseitige Verknüpfung mehrerer Zeichenmodalitäten auf verschiedenen Ebenen (z.B. Semantik, Handlungsfunktion etc.) in einem Gesamttext.
2. Multimodalität ist eine gesamtkulturelle Kompetenz und eine individuelle Intelligenz, die das Erschließen von Sinn durch „Transkribieren“, d. h. Kommentieren, Erklären oder Paraphrasieren von Botschaften eines Zeichensystems/Mediums durch ein anderes beinhaltet.
3. Multimodalität ist eine Regeln folgende und Muster bildende semiotisch-kognitive Aktivität in der Produktion und im Verstehen von Texten. (ebd. S. 47)

Der Terminus Multimodalität wird demzufolge nicht nur zur Bezeichnung von Verknüpfungen mehrerer Zeichenmodalitäten in einem Gesamttext gebraucht. Er wird zugleich mit der Fähigkeit assoziiert, den Sinn solcher Handlungen korrekt zu verstehen bzw. zu erklären und mit ihnen kooperieren zu können.

Der Kern der Multimodalität ist, dass alle im Gesamttext vorkommenden, semiotischen Elemente (Zeichenmodalitäten) sehr eng zusammenarbeiten, um eine neue Botschaft entstehen zu lassen, die mehr umfasst als die Summe ihrer Einzelteile. Diese in den Begriffen *Bedeutungsmultiplikation*<sup>4</sup> oder *semiotische Synergie*<sup>5</sup> zum Ausdruck kommende Eigenschaft stellt klar, warum es so wichtig ist, die multimodalen Texte nicht nur durch das Prisma ihrer separaten Einzelteile zu betrachten: Erst durch das holistische Wahrnehmen aller ihrer Elemente lässt sich der richtige Sinn solcher Botschaften erzeugen. Wildfeuer u.a. (2020) verdeutlichen:

Diese Abhängigkeiten sowie Wechselwirkungen zwischen den Elementen ist ein wichtiges und zugleich schwieriges Thema im Kontext von Multimodalität, weil der Gebrauch unterschiedlicher Modalitäten jeweils Einfluss darauf haben kann, *was jede einzelne Modalität zum großen Ganzen beiträgt*. Das bedeutet, dass wir in unserer Analyse sehr viel Wert darauf legen sollten, die Zeichenmodalitäten sowohl in ihrer Individualität als auch in ihrem Zusammenspiel zu betrachten. Wir müssen also zum einen herausfinden, was jede Modalität für sich allein zum komplexen Ganzen tatsächlich beiträgt, dürfen aber zugleich nie aus

---

<sup>4</sup> Diese Metapher wurde am Beispiel des von Jay Lemke 1998 weit verbreiteten Ausdrucks *multiplying meanings* gebildet (vgl. Wildfeuer u.a., 2020: 18).

<sup>5</sup> Siehe: Schmitz, 2015: 12.

dem Blick verlieren, welche Informationen ‚lediglich‘ aus der Tatsache entstehen, dass mehrere Modalitäten miteinander kombiniert werden. Zeichenmodalitäten, die in Kombination vorkommen, müssen *immer auch* mit Blick auf ihre gegenseitigen Wechselwirkungen interpretiert werden und können nicht allein unabhängig voneinander betrachtet werden. (ebd. S. 20, Kursivschrift im Original)

Dazu ein Beispiel (Siehe Abb. 1 unten):



Abb. 1: Werbeplakat von Nutella (Quelle: <https://www.pinterest.de/pin/537195061782357419/>) [Zugriff: 08.08.2021].

Dieses Werbeplakat von Nutella ist ein einfacher multimodaler Text. Ihn bilden mehrere unterschiedliche Zeichenmodalitäten, und zwar: Bild, Sprache und Typographie<sup>6</sup>. Wenn man sich jetzt jede Zeichenmodalität getrennt ansieht, so stellt sich heraus, dass z. B. das Bild, das sich zentral auf dem Plakat befindet, ein Glas Nutella mit Haselnüssen (Hauptzutat der Schokoladencreme) und einen Brotkorb im Hintergrund darstellt. Alle diese Produkte, die man gewöhnlich zum Frühstück verzehrt, liegen auf einem Küchentisch. Das nächste – sprachliche – Element des Gesamttextes, das der kurze Satz *Der Morgen macht den Tag* bildet, ist dagegen ein populärer Spruch, der so viel bedeutet wie: „Wie du den Tag am Morgen beginnst, wirkt sich für dich auf den ganzen Tag aus“. Lässt man den Blick etwas länger auf diesem Satz verweilen, so wird klar, dass die Typographie des Satzes – das dritte Element des

---

<sup>6</sup> Stöckl (2004) unterscheidet zwischen zentralen und peripheren Zeichenmodalitäten. Zu den ersten zählen laut ihm u.a.: Sprache, Bild, Geräusche und Musik, zu den anderen z. B. Typographie:

Periphere Modalitäten zeichnen sich dadurch aus, dass sie nur durch die mediale Realisierung einer zentralen Modalität ins Spiel kommen und eng mit dieser Kernmodalität verbunden bleiben. So sind Typographisches, Paraverbales und Nonverbales periphere Zeichenmodalitäten der zentralen Modalität Sprache, die sie materialisieren helfen, begleiten und überformen. (ebd. S. 16)

Gesamttextes – dem Namen der Schokoladencreme auf dem Glas sehr ähnelt (dieselbe Art der Schrift, ähnliche Größe der Buchstaben, dieselben Farben – in Schwarz der erste Buchstabe und in Rot die übrigen). Schon von dieser Typographie her lässt sich erkennen, dass diese zwei Elemente, Bild und Sprache, hier im engen Zusammenhang miteinander stehen und sich wechselseitig kontextualisieren werden. Welche Gesamtbotschaft entsteht also aus ihrer Interaktion? Das holistische Wahrnehmen dieser Elemente lässt vermuten, dass dank einem mit Nutella bestrichenen Frühstücksbrot, das immer gut schmeckt und dadurch so viel Freude beim Essen bereitet, die gute Laune tagsüber erhalten wird. Die unterschwellige Botschaft des Gesamttextes (Text und Bild wirken hier persuasiv) ist also: Wer sich einen guten Start in den Tag sichern will, muss unbedingt Nutella kaufen und es am Morgen essen.

Schon aus dieser kurzen Analyse wird ersichtlich, auf welche Art und Weise Text und Bild in multimodalen Texten zusammenwirken. Die sprachlichen Elemente des Gesamttextes beschreiben nicht einfach die bildlichen, die bildlichen Elemente illustrieren nicht einfach die sprachlichen. Sie ergänzen sich gegenseitig und benötigen einander, um gemeinsam den beabsichtigten Sinn herzustellen.

Trotz der Tatsache, dass wir tagtäglich mit ähnlichen multimodalen Texten konfrontiert werden (seien das Internetseiten, Applikationen, Bücher, Zeitschriften, Fernsehprogramme oder andere Werbungen) und diese intuitiv nachzuvollziehen glauben, wissen wir tatsächlich noch relativ wenig darüber, wie genau diese Texte aufgebaut sind und produktiv Sinn ergeben (vgl. Wildfeuer u.a., 2020: 8). Zu einem erfolgreichen Umgang mit ihnen (d.h. zum Verstehen ihrer Bauweise, zur Erschließung ihrer Botschaft, zum Verstehen der Wirkung des Gesamttextes etc.) sind bestimmtes Wissen und eingeübte Fähigkeiten<sup>7</sup> vonnöten. Sowohl als Produzent\*in als auch als Rezipient\*in solcher Texte muss man imstande sein, Verknüpfungen zwischen semiotisch ausdifferenzierten Ressourcen zu verstehen und diese reflektierend zu betrachten.

---

<sup>7</sup> Stöckl (2011), für den Sprache und Bild als Teiltexthe fungieren und im gleichen Maße über den Textstatus verfügen, bezeichnet solche kognitiven und textpraktischen Tätigkeiten, die man zur Lektüre von „Sprache-Bild-Texte“ benötigt, als eine Art *multimodaler Kompetenz*. Ihm zufolge bedeutet multimodale Kompetenz, dass man in der Lage ist:

- Sorten bzw. Typen von Bildern kategorisierend zu erkennen,
- dem Bild eine im Verwendungskontext relevante Bedeutung zuzuweisen,
- den Sprachtext im Abgleich mit der visuellen Botschaft zu verstehen,
- semantisierte Sprache und kontextualisiertes Bild zu integrieren sowie die Bildlichkeit der Sprache und der Textfläche bzw. des Schriftköpers in den Prozess des Gesamtverstehens einzubeziehen. (ebd. S. 45)

### 3. Logovisuelle Didaktik

Der Pionier und Initiator solch einer Art Didaktik ist der deutsche Sprachwissenschaftler Ulrich Schmitz. Schon im Jahre 2004 bemerkte er, dass Text-Bild-Kombinationen anders als übliche Texte interpretiert werden und dass es nötig ist, Lernenden adäquate Verstehensstrategien zum reflektierten Umgang mit ihnen zu vermitteln (vgl. ebd. S. 230). Das Konzept der Didaktik legt der Autor wie folgt dar:

Es geht um konzentrierten Nachvollzug, Verstehen der Bauweise, kritische Reflexion und ggf. Anleitung zu bewusster Eigenproduktion. (Schmitz, 2004: 230)

Die logovisuelle Didaktik verfolgt das Ziel, die Lernenden für die bewusste Aneignung von Text-Bild-Kombinationen sowie eine bewusste Selbstproduktion von Text-Bild-Kombinationen zu sensibilisieren. Sie soll überdies dem Missverständnis entgegenwirken, „dass Text-Bild-Kombinationen leicht zu verstehen, also mühelos für weitgehende Lernprozesse eingesetzt werden können“ (Raith, 2014: 33).

Zurzeit fehlt es an etablierten theoretischen Ansätzen zur logovisuellen Didaktik. Der Grund dafür ist einerseits die Tatsache, dass sie noch ganz am Anfang ihrer Entwicklung steht (es gibt noch zu wenig Forschung auf diesem Gebiet), andererseits wirkt erschwerend, dass im Bereich der Multimodalität (und im Anschluss daran auch der logovisuellen Didaktik) weiterhin keine terminologische Einheitlichkeit herrscht. Die ausgearbeiteten Untersuchungen bedienen sich oftmals unterschiedlicher Begriffe, was den Weg zu einer systematischen, theoretischen Basis dieser Didaktik stark erschwert. Nichtsdestotrotz lässt sich in allen Arbeiten zum Thema eine generelle Tendenz beobachten: Sie gehen von der Annahme aus, dass zur Rezeption von Text-Bild-Kombinationen die Entwicklung und Beherrschung einer erweiterten, durch theoretische Ansätze der Bildwissenschaft unterstützten Lesekompetenz nötig ist, die über die traditionelle Fähigkeit zum Erfassen linearer Texte hinausgeht und die es ermöglicht, Text-Bild-Kombinationen holistisch zu „lesen“ (vgl. z. B. Schmitz, 2004: 230; Raith, 2014: 29; Blell, 2010: 95; Schmitz, 2015: 10; Wittstruck, 2018: 13). Schmitz (2015) verdeutlicht:

Dafür reicht Lesekompetenz, wie sie in Schule und Hochschule vermittelt werden soll, bei Text-Bild-Gefügen nicht aus. Erstens braucht es Bildung für Bilder<sup>8</sup>; darauf können wir an dieser Stelle nicht eingehen. Zweitens, und darum geht es hier, braucht man ein Instrumentarium zum Verständnis der Architektur von Text-Bild-Kompositionen. (ebd. S. 10)

---

<sup>8</sup> Die Rede ist hier von der visuellen Kompetenz (ang. visual literacy). Dazu mehr z. B.: Hallet, 2010; Hansen, 2018.

Im Folgenden werden einige in diesem Kontext entstandene Überlegungen präsentiert, die Vorgehensweisen für die Analysen von Text-Bild-Kombinationen im Fach Deutsch als Fremdsprache anbieten und die es nach einer weiteren Ausarbeitung und ggf. Modifizierung in Zukunft ermöglichen könnten, eine solide methodische Basis für die logovisuelle Didaktik zu schaffen.

Über die Notwendigkeit der Entwicklung einer spezifischen intermediären Lesekompetenz<sup>9</sup> spricht z. B. Gabrielle Blell (2010). Die Autorin verweist darauf, dass es auch für den Fremdsprachenunterricht vorteilhaft sein kann, die Leseförderung (im Sinne des literarischen Lesens<sup>10</sup>) und die Medienerziehung (im Sinne der Förderung visueller Erziehung) als gemeinsame und einander ergänzende Aufgaben zu betrachten (vgl. Blell, 2004, weiter auch: Blell, 2010: 94). Um dies zu erreichen, ist es nötig, eine Art intermediärer Seh- und Lesekompetenz zu entwickeln, die diese zwei Fähigkeiten verbindet und es ermöglicht, solche Text-Bild-Kombinationen angemessen zu erfassen:

Die Zielstellung der Entwicklung intermediärer und (inter-)kultureller<sup>11</sup> Lese- und Sehkompetenz wird folglich begriffen als die Fähigkeit, (kombinierte) Bild- und Schrifttexte (Werbeplakat, Zeitschriftendoppelseite, *website*, *hyperfiction* etc.) als kulturelle Artefakte intermediär und medial zu lesen und sie im Wechselspiel eigen- und fremdkultureller Bezugskulturen zu interpretieren. (Blell, 2010: 96; Kursivschrift im Original)

Den Ausgangspunkt für Blells Überlegungen bildet die Auffassung, dass die Rezeption von Text-Bild-Kombinationen stets nach einem festen Muster erfolgt. Laut der Autorin liest/schaut der/die Empfänger\*in solche Art der Botschaften

- 1) zuallererst übergreifend schemageleitet, d.h. medienunspezifisch (er/sie nimmt den Gesamttext holistisch wahr);
- 2) erst später spezifisch mediengeleitet, d.h. im Hinblick auf bestimmte Elemente des Gesamttextes: Bild, Sprachtext, Grafik, etc. (vgl. ebd. S. 96).

---

<sup>9</sup> Unter dem Begriff *intermediale Lesekompetenz* versteht die Autorin einerseits: „die Fähigkeit, einzelne Wörter, Sätze und Texte flüssig zu lesen und im Textzusammenhang zu verstehen [...]“, andererseits ist hier: „ein erweitertes Leseverständnis in dem Sinne, ganze Seiten (inklusive Bilder etc.) als „integrated texts“ zu lesen, bzw. zu betrachten“ gemeint (vgl. Blell, 2010: 95).

<sup>10</sup> Dabei geht es um „Lesen als literalisierte schriftsprachliche Praxis“ (vgl. Blell, 2010: 95).

<sup>11</sup> Den Namen dieser Kompetenz ergänzt die Autorin um das Wort *interkulturell*, das darauf hinweisen soll, dass die Rezeption von „Lese- und Bildtext“ immer als subjektiv und kulturell determiniert angesehen wird (vgl. Blell, 2010: 94).



In jeder Phase sind bestimmte Teilkompetenzen zu unterscheiden. Welche genau dazugehören und wie genau der Prozess des „transmedialen Lesens“ abläuft, fasst Blell in Form eines Schemas zusammen (Abb. 2. unten):

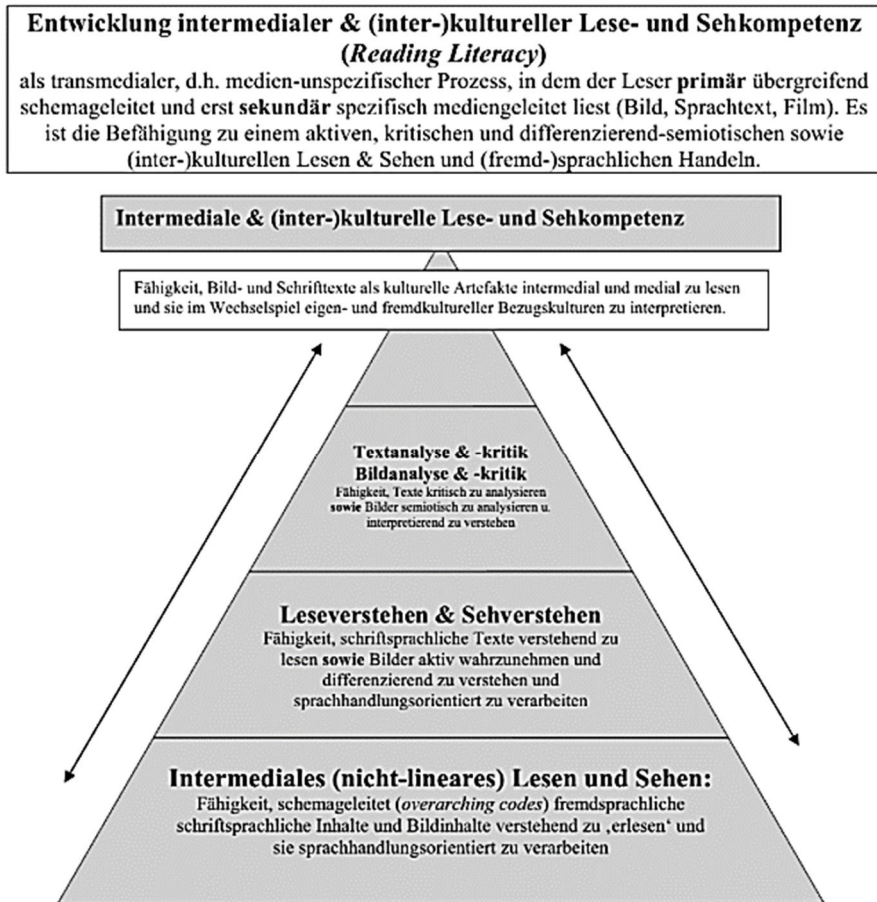


Abb.2: Intermediale und (inter-)kulturelle Lese- und Sehkompetenz (Quelle: Blell, 2010: 97).

Eine andere Betrachtungsweise bietet Markus Raith (2014) an. Das von ihm vorgeschlagene Prozessmodell des Leseverstehens symbolischer Darstellungen bezieht sich auf das Konzept von Schnotz und Dutke (2004) und erweitert es. Ihm liegt die Auffassung zugrunde, dass Lesen schriftlicher und multimodaler Texte ein interaktiver Konstruktionsprozess ist, in dem „[e]xterne Informationen aus dem Text mit internen Wissensbeständen der Leser zu einem kohärenten mentalen Gebilde verknüpft [werden]“ (Raith, 2014: 33). In diesem Modell sind fünf Hauptphasen zu unterscheiden. Die Rede ist von:

- 1) Allgemeinverständnis;
- 2) Fokussieren und Inferieren jeweils auf Text- und Bildebene;
- 3) Das Heranziehen von Vor- bzw. Kontextwissen zum Fokussieren und Inferieren;
- 4) Analysieren und Interpretieren des Gesamtgefüges;
- 5) Das Reflektieren. (Raith, 2014: 34)

Der Autor betont, dass der Inferenz, die auf der Ebene des Textes, auf der Ebene des Bildes und auf der Gesamtebene der Text-Bild-Kombination erfolgt, die wichtigste Bedeutung zugeschrieben werden muss. „Dabei ist zu beachten, dass beide Symbolsysteme wechselseitig Inferenzprozesse unterstützen können. Was die Rezipienten an Vorwissen aktivieren müssen, um den Text zu verstehen, kann vom Bild unterstützt werden und umgekehrt“ (ebd. S. 34).

Ein drittes Konzept kommt von Schmitz (2015). Treten laut ihm Text-Bild-Kombinationen auf besonderen *Sehflächen*<sup>12</sup> auf, ist die Erzeugung ihres Sinnes nur nach dem Verstehen von deren Funktionsweise möglich. Weil Schauen und Lesen bei solchen Botschaften wechselseitig aufeinander angewiesen sind, bedarf es einer besonderen Kompetenzart, die es ermöglicht, für ihre Analyse und Interpretation diese zwei Aktivitäten zu aktivieren und miteinander zu verknüpfen – die sog. Sehlesekompetenz (Schmitz, 2015: 9). Schmitz hebt vor, dass die Produktion und Rezeption von Text-Bild-Kombinationen anders als bei herkömmlichen, linearen Texten verläuft. Folgende Besonderheiten sollten bei ihrer Lektüre in Anspruch genommen werden:

- Informationen erfolgen nicht linear hintereinander, sondern verteilen sich nach einem kartografischen Plan (Layout und Design) auf der Fläche;
- Wahrnehmung zuallererst als Ganzes;
- Erfassen (top down) wie ein Bild zunächst in der oberflächlichen Gestaltung: Welche Teile sind optisch (zum Beispiel durch Hervorhebungen, Rahmen oder leere Flächen) voneinander unterschieden und in welchem visuellen Verhältnis stehen sie zueinander (zum Beispiel auffällig, unscheinbar, oben, unten, rechts, links, innen, außen, groß, klein);
- Verteilung der einzelnen Gestalten auf der Fläche gibt Indizien für deren Wichtigkeiten und inhaltliche Beziehungen untereinander;
- Text-Bild-Gefüge ermöglichen (im Unterschied zu typischen Bildern) selektive Wahrnehmung. (vgl. Schmitz, 2015: 12f)

Seinem Beitrag fügt der Autor eine Fünfzehn-Punkte-Liste (Hilfsfragen) hinzu, die die Rezipient\*innen zur angemessenen und bewussten Analyse der spezifischen Architektur von Sehflächen hinlenken soll. Sie stellt sich wie folgt dar:

---

<sup>12</sup> „Sehflächen sind Flächen, auf denen Texte und Bilder in geplantem Layout gemeinsame Bedeutungseinheiten bilden.“ (Schmitz, 2011: 25)



1. Welchem kartografischen Plan folgt die Verteilung der Informationen auf der Fläche?
2. Warum steht was an welcher Stelle?
3. Welche grafischen Elemente dienen nur dem Flächendesign, welche tragen auch inhaltliche Bedeutung?
4. Wie groß sind die Anteile von Bild, Grafik und Text und warum?
5. Wie viele Informationseinheiten gibt es?
6. Welche Bilder und Grafiken gehören zu welchen Texten und umgekehrt?
7. Durch welche ästhetischen Mittel kommt die Verbindung der nonverbal-visuellen und der schriftlichen Elemente zustande?
8. Warum sind welche Informationen auf Bilder, Grafiken und Texte verteilt?
9. In welchem formalen und inhaltlichen Verhältnis stehen Bild, Grafik und Text in jeder einzelnen Informationseinheit?
10. Wie müsste der Text – bei etwa gleicher Information – formuliert werden, wenn es keine Bilder oder Grafiken gäbe?
11. Werden durch die Verbindung von Text und Bild unterschwellige Informationen transportiert, die nicht ausdrücklich formuliert werden?
12. Welche der angebotenen Informationseinheiten interessieren mich selbst?
13. Habe ich danach gesucht oder habe ich mich vom Design leiten lassen?
14. Wie intensiv widme ich mich einer Informationseinheit und was veranlasst mich, von hier zu einer (welcher?) anderen zu springen?
15. Was habe ich nach einer halben Stunde noch behalten und warum? (Schmitz, 2015: 13)


Zusammenfassend ist zu sagen, dass bei der Analyse multimodaler Texte (inklusive Text-Bild-Kombinationen) drei wesentliche Schritte zu beachten sind: die ganzheitliche, einführende Erkennung des Gesamttextes (*was für ein multimodaler Text ist das, aus welchen Bestandteilen besteht er, etc.* – Schritt 1), die Betrachtung jeder Zeichenmodalität des Gesamttextes in ihrer Individualität (Schritt 2) und schließlich die Betrachtung aller Zeichenmodalitäten in ihrem multimodalen Zusammenspiel (Schritt 3). Die oben angeführten Theorien legen nahe, dass zur Erfüllung dieser Aufgaben die Entwicklung bestimmter Kompetenzen unerlässlich ist. Die Rede ist von: 1) der in der Schule erworbenen Lesekompetenz (zum Verstehen der sprachlichen Elemente), 2) der von der Bildwissenschaft herzuleitenden visuellen Kompetenz (zum Verstehen der bildlichen Elemente) und 3) der spezifischen, intermedialen Sehlesekompetenz (zum Verstehen des Zusammenspiels beider Zeichenmodalitäten). Nicht zu vernachlässigen ist auch der Erwerb von fundierten theoretischen Kenntnissen bezüglich der allgegenwärtigen Multimodalität, die es ermöglichen, einen Gesamttext in seinen einzelnen Bestandteilen zu erkennen, sowie das Vorhandensein eines (inter)kulturellen Grundwissens, das es erlaubt, einen über die gewohnte soziale Umgebung hinausgreifenden Gesamttext in einen angemessenen Kontext

einzuordnen und kritisch zu interpretieren. Diese Kenntnisse und Fähigkeiten zu vermitteln, sind wesentliche Aufgaben einer logovisuellen Didaktik.

#### 4. Text-Bild-Kombinationen im Fremdsprachenunterricht – Funktionen und Einsatzmöglichkeiten

Im Folgenden sind konkrete Ideen für Einsatzmöglichkeiten im DaF-Unterricht aufgelistet, die veranschaulichen, wie im Fremdsprachenunterricht ausgewählte sprach- und nichtsprachbezogene Kompetenzen multimodal gefördert werden können.

Text-Bild-Kombination	Funktion
<p style="text-align: center;"><b>„Die Gabe“</b></p>  <p>Abb. 3: Tafelbild zu ditransitiven Verben (Quelle: Hieronimus, 2014: X; leicht verändert).</p>	<p>Entwicklung der grammatischen Kompetenz</p>
<p><b>Einsatzmöglichkeit und Erläuterung</b></p>	
<p>Grammatikunterricht: Mit Hilfe von Text-Bild-Kombinationen kann man Lernenden den korrekten Gebrauch von unterschiedlichen grammatischen Strukturen erklären. Nach der Analyse des Gesamttextes und dem Entdecken, wie der Satz in solchen Fällen funktioniert und gebildet werden soll, werden die Schüler*innen es in Zukunft leichter haben, diese Szene aus dem Gedächtnis abzurufen, dieses Prinzip auf eine andere Situation zu übertragen und es anzuwenden.</p>	
Text-Bild-Kombination	Funktion
 <p>zu tief ins Glas schauen (lit. to look too deep into the glass)</p>	<p>Entwicklung der phraseologischen Kompetenz</p>
<p>Abb. 4: Deutsche Idiome (Quelle: <a href="https://blog.chatterbug.com/en/20-funny-german-idioms-you-should-know/">https://blog.chatterbug.com/en/20-funny-german-idioms-you-should-know/</a> [Zugriff: 28.06.2021]).</p>	

<p style="text-align: center;"><b>Einsatzmöglichkeit und Erläuterung</b></p> <p>Wortschatzarbeit im Unterricht: Um die phraseologische Kompetenz (sei es Idiome oder Sprüche) bei Lernenden zu fördern, kann man ebenfalls auf Text-Bild-Kombinationen zurückgreifen. Die humoristische Zusammenstellung von Text und Bild kann bei Lernenden längerfristig das Interesse an entdeckendem Lernen wecken (nach der Analyse des Gesamttextes müssen sie die übertragene Bedeutung des Idioms erraten), sie macht Spaß und hat zur Folge, dass die spezifischen Redewendungen für längere Zeit im Gedächtnis gespeichert werden.</p>	
<p style="text-align: center;"><b>Text-Bild-Kombination</b></p>	<p style="text-align: center;"><b>Funktion</b></p>
<div style="text-align: center;">  </div>	<p>Zentrale Funktion: Entwicklung der sprachlichen Fertigkeit (Sprechen); Förderung kulturbezogener und landeskundlicher Kompetenzen; Entwicklung der Medienerziehung</p> <p>Periphere Funktion: Entwicklung der sprachlichen Fertigkeit (Schreiben); Entwicklung der grammatischen Kompetenz</p>
<p>Abb. 5: Werbung eines Lichtspektakels. Potsdam (2019) (Quelle: eigene Aufnahme).</p>	
<p style="text-align: center;"><b>Einsatzmöglichkeit und Erläuterung</b></p> <p>Landeskunde / Film- und Videoarbeit im Unterricht / Medienerziehung (auch peripher: Wortschatzarbeit im Unterricht / Grammatikunterricht): Für die Arbeit mit einem Plakat gibt es mannigfaltige Einsatzmöglichkeiten. Es kann sowohl zum Gegenstand einer ganzen Unterrichtsstunde werden (in diesem Falle wird die Gruppe den Gesamttext detailliert analysieren), als auch als Einführung für die Arbeit mit einem Film dienen. Bei einem Werbeplakat lohnt es sich, dieses auch im Kontext der damit verbundenen <i>Persuasion</i> und ihrer Wirkung auf die Rezipient*innen zu analysieren. Nach dem Besprechen des Gesamttextes, kann weiter mit ihm gearbeitet werden. In Form einer Hausaufgabe können die Lernenden kurze Anschluss-texte mit Bezug auf das präsentierte Plakat verfassen.</p>	

Text-Bild-Kombination	Funktion																
<p style="text-align: center;"><b>Wer hat wie viel Freizeit?</b> Pro Werktag verfügen Bundesbürger nach eigenen Angaben über eine durchschnittliche Freizeit von...</p>  <table border="1" style="margin-left: auto; margin-right: auto;"> <tr> <td>Jugendliche (14-17 Jahre)</td> <td>4 Std. 49 Min.</td> </tr> <tr> <td>Ruhestandler (ab 65 Jahre)</td> <td>4:45</td> </tr> <tr> <td>Junge Erwachsene (18-24 Jahre)</td> <td>4:12</td> </tr> <tr> <td>Gesamtdurchschnitt</td> <td>4:03</td> </tr> <tr> <td>Jungsenioren (50-64 Jahre)</td> <td>3:59</td> </tr> <tr> <td>Singles (25-49 Jahre)</td> <td>3:49</td> </tr> <tr> <td>Paare*</td> <td>3:42</td> </tr> <tr> <td>Familien**</td> <td>3:10</td> </tr> </table> <p><small>* Frauen bzw. Männer im Alter von 25-49 Jahren ohne Kinder ** Mütter bzw. Väter im Alter von 25-49 mit Kindern unter 14 Jahre</small></p> <p style="text-align: right;"><small>Stand Februar 2010 © Globus 4499</small></p>	Jugendliche (14-17 Jahre)	4 Std. 49 Min.	Ruhestandler (ab 65 Jahre)	4:45	Junge Erwachsene (18-24 Jahre)	4:12	Gesamtdurchschnitt	4:03	Jungsenioren (50-64 Jahre)	3:59	Singles (25-49 Jahre)	3:49	Paare*	3:42	Familien**	3:10	<p>Entwicklung der sprachlichen Fertigkeit (Sprechen / Schreiben)</p>
Jugendliche (14-17 Jahre)	4 Std. 49 Min.																
Ruhestandler (ab 65 Jahre)	4:45																
Junge Erwachsene (18-24 Jahre)	4:12																
Gesamtdurchschnitt	4:03																
Jungsenioren (50-64 Jahre)	3:59																
Singles (25-49 Jahre)	3:49																
Paare*	3:42																
Familien**	3:10																
<p style="text-align: center;"><b>Einsatzmöglichkeit und Erläuterung</b></p> <p>Wortschatzarbeit im Unterricht / Förderung der Schreibkompetenz: Auch eine Infographik kann die sprachlichen Fertigkeiten der Lernenden entwickeln helfen. Bei der Arbeit mit Infographiken erfolgt das meist in Form einer Interpretation des Gesamttextes oder ihrer schriftlichen Beschreibung. Alle diese Aufgaben setzen sich zum Ziel, auf die Fantasie der Lernenden einzuwirken (die Schüler*innen müssen sich die in der Infografik dargestellte Situation vorstellen), ihre Kreativität zu fördern (sie müssen die zwischen unterschiedlichen Datensets der Infografik auftretenden Korrelationen und Zusammenhänge entdecken und deuten) und sie zum Sprechen / Schreiben anzuregen.</p> <p>* Eine gute Idee bei solchen Übungen (besonders bei Anfängern oder leistungsschwächeren Schüler*innen) ist das Zufügen von Formulierungshilfen, auf die sich die Lernenden beim Sprechen / Schreiben stützen können.</p>																	
<p style="text-align: center;"><b>Text-Bild-Kombination</b></p>	<p style="text-align: center;"><b>Funktion</b></p>																
	<p>Entwicklung der phonetischen und prosodischen Kompetenz</p>																
<p>Abb. 7: Visuelle Phonetik. (Quelle: <a href="https://www.goethe.de/resources/files/jpg576/12---endsilben_vis_phonetik.jpg">https://www.goethe.de/resources/files/jpg576/12---endsilben_vis_phonetik.jpg</a>) [Zugriff: 28.06.2021].</p>																	

Einsatzmöglichkeit und Erläuterung	
<p>Phonetik-Unterricht: Mit Hilfe von Text-Bild-Kombinationen lässt sich auch die phonetische und prosodische Kompetenz bei Lernenden auf originelle Weise fördern. Phonetische Regeln werden besonders häufig in Form linearer Texte angeboten. Lange, oft komplizierte Erklärungen entmutigen die Lernenden und sind dem effektiven Lernen nicht förderlich. Als gute Alternative kann die Posterserie „Visuelle Phonetik“<sup>13</sup> dienen. Die insgesamt 12 Poster stellen auf einfache und zugängliche Art und Weise die wichtigsten deutschen phonetischen Regeln dar, sie können im Unterricht ab A1-Niveau eingesetzt werden. Dank ihrer Hilfe kann es Lernenden leichter fallen, sich die (oft schwierige) Theorie schneller und effektiver zu merken oder auch (wenn die Poster z. B. im Klassenzimmer aufgehängt werden) sich jederzeit daran zu erinnern.</p>	

## 5. Fazit

Weil die Welt derzeit von multimodalen Texten aller Art geprägt ist<sup>14</sup> und die Mehrheit der Menschen diese „blindlings“, d.h. ohne groß darüber nachzudenken, rezipiert, sollte es Anliegen der Schule sein, Lernende für den richtigen Umgang mit ihnen vorzubereiten, sowohl für schulische als auch für außerschulische Zwecke. Bilder treten in der heutigen Welt selten allein auf. Sie werden meistens in Verbindung mit Sprache oder anderen Zeichenmodalitäten angeboten<sup>15</sup>. In der gleichen Form sollten sie ebenfalls im Fremdsprachenunterricht erscheinen. Nicht nur deswegen, weil sie der Alltagskommunikation ähneln, sondern auch deswegen, weil sie über ein enormes didaktisches Potential verfügen. Die in Abschnitt 4 dargestellten Unterrichtsvorschläge haben gezeigt, dass die Text-Bild-Kombinationen auf vielseitige Art und Weise im Fremdsprachenunterricht eingesetzt und gebraucht werden können. Dank ihrer Hilfe lassen sich nicht nur neue Möglichkeiten zur Förderung sprachbezogener und nichtsprachbezogener Kompetenzen schaffen, sondern auch die Effektivität des Fremdsprachenlehr- und Lernprozesses erhöhen.

<sup>13</sup> Mehr dazu online unter: <https://www.goethe.de/ins/jp/de/spr/unt/kum/20912631.html>.

<sup>14</sup> Bucher (2010: 41) hält Multimodalität schon für eine *Universalie des Medienwandels*; Wildfeuer u.a. (2020: 18) sehen sie als *eine schon immer geltende Norm*; Stöckl (2011: 47) dagegen bezeichnet multimodale Texte als der *Normalfall des Kommunizierens*.

<sup>15</sup> Stöckl (2011) verdeutlicht:

„Viele semiotische, philosophische und kommunikationswissenschaftliche Arbeiten vernachlässigen die Tatsache, dass Bilder im kommunikativen Haushalt unserer Gesellschaft selten allein stehen, sondern zumeist mit Sprache und anderen Zeichensystemen verknüpft vorkommen. Selbst so scheinbar typisch „reine“ Bildergebrauchsdomänen wie die Gemädegalerie oder Fotoausstellung sind letztlich auf in anderen Zeichensystemen kodierte und über andere Medien „transportierte“ Erklärungen und Hinweise angewiesen, wie z.B. Bildtitel, Bildlegenden, Künstlerbiografien, Ausstellungsrezensionen, Ausstellungskataloge und Vernissagen (d.h. Gespräche über die betreffenden Kunstgegenstände)“ (ebd. S. 46).

## BIBLIOGRAPHIE

- Assmann J. (2004), *Die Frühzeit des Bildes – der altägyptische iconic turn*, (w:) Maar C., Burda H. (red.), *Iconic turn. Die neue Macht der Bilder*. Köln: DuMont Literatur und Kunst Verlag, s. 304–322.
- Blell G. (2004), *(New)Media Literacy: Gedanken zur Entwicklung von fremdsprachiger Lesekompetenz bei der Arbeit mit Hyperfiction*, (w:) Bosenius P., Donnerstag J. (red.), *Interaktive Medien und Fremdsprachenlernen*. Frankfurt a. M.: Peter Lang Verlag, s. 45–64.
- Blell G. (2010), *Der Leser als ‚Grenzgänger‘: Entwicklung intermedialer Lese- und Sehkompetenzen*, (w:) Hecke C., Surkamp C. (red.), *Bilder im Fremdsprachenunterricht – Neue Ansätze, Kompetenzen und Methoden*. Tübingen: Gunter Narr Verlag, s. 94–109.
- Boehm G. (1994), *Die Wiederkehr der Bilder*, (w:) Boehm G. (red.), *Was ist ein Bild?* München: Wilhelm Fink Verlag, s. 11–38.
- Bucher H.-J. (2010), *Multimodalität – eine Universalie des Medienwandels: Problemstellungen und Theorien der Multimodalitätsforschung*, (w:) Bucher H.-J., Gloning T., Lehnen K. (red.), *Neue Medien – Neue Formate. Ausdifferenzierung und Konvergenz in der Medienkommunikation*. Frankfurt a. M. – New York: Campus Verlag, s. 41–79.
- Bucher H.-J. (2012), *Multimodalität – ein universelles Merkmal der Medienkommunikation. Zum Verhältnis von Medienangebot und Medienrezeption*, (w:) Bucher H.-J., Schumacher P. (red.), *Interaktionale Rezeptionsforschung. Theorie und Methode der Blickaufzeichnung in der Medienforschung*. Wiesbaden: Springer Verlag, s. 51–82.
- Hallet W. (2010), *Viewing Cultures: Kulturelles Sehen und Bildverstehen im Fremdsprachenunterricht*, (w:) Hecke C., Surkamp C. (red.), *Bilder im Fremdsprachenunterricht – Neue Ansätze, Kompetenzen und Methoden*. Tübingen: Gunter Narr Verlag, s. 26–54.
- Hansen M. (2018), *Das Bild im Fremdsprachenunterricht – facettenreich von der Semantisierungshilfe bis zum Sprech Anlass*, (w:) Kaunzner U. A., Reeg U., Ehrhardt C. (red.), *Bild und Sprache. Impulse für den DaF-Unterricht*. Münster – New York: Waxmann Verlag, s. 149–159.
- Hedling E., Lagerroth U.-B. (2002), *Introduction*, (w:) Hedling E., Lagerroth U.-B. (red.), *Cultural Functions of Intermedial Exploration*. Amsterdam – New York: Brill/Rodopi, s. 7–13.
- Hieronimus M. (2014), *Einleitung*, (w:) Hieronimus M. (red.), *Visuelle Medien im DaF-Unterricht*. Göttingen: Universitätsverlag, s. III–XV.



- Jäger L. (2002), *Transkriptivität. Zur medialen Logik der kulturellen Semantik*, (w:) Jäger L., Stanitzek G. (red.), *Transkribieren: Medien/Lektüre*. München: Wilhelm Fink Verlag, s. 19–41.
- Łosiewicz M. (2009), *Rola obrazu w komunikacji społecznej*, (w:) Obrębska A. (red.), *Komunikacja wizualna w przestrzeni społecznej*. Łódź: Primum Verbum, s. 205–212.
- Mendelssohn M. (1989) [1783], *Jerusalem oder über religiöse Macht und Judentum*, (w:) Thom M. (red.), *Schriften über Religion und Aufklärung*. Berlin: Union Verlag, s. 351–458.
- Mitchell W. J. (1992), *The Pictorial Turn*. "Artforum 30", nr 7, s. 89–94.
- Opiłowski R. (2015), *Der multimodale Text aus kontrastiver Sicht. Textdesign und Sprache-Bild-Beziehungen in deutschen und polnischen Presstexten*. Wrocław – Dresden: ATUT – Neisse Verlag.
- Raith M. (2014), *Multimodales Verstehen und kulturelles Lernen. Zu einer Didaktik des Logovisuellen*, (w:) Hieronimus M. (red.), *Visuelle Medien im DaF-Unterricht*. Göttingen: Universitätsverlag, s. 25–52.
- Schmitz U. (2004), *Bildung für Bilder. Text-Bild-Lektüre im Deutschunterricht*, (w:) Jonas H., Josting P. (red.), *Medien – Deutschunterricht – Ästhetik*. Festschrift für Jutta Wermke. München: Kopaed Verlag, s. 219–232.
- Schmitz U. (2011), *Sehflächenforschung. Eine Einführung*, (w:) Diekmannshenke H., Klemm M., Stöckl H. (red.), *Bildlinguistik: Theorien – Methoden – Fallbeispiele*. Berlin: Erich Schmidt Verlag, s. 23–42.
- Schmitz U. (2015), *Sehlesekompetenz. Text-Bild-Gefüge lesen und verstehen*. "Weiterbildung", nr 3, s. 9–13.
- Schnotz W., Dutke S. (2004), *Kognitionspsychologische Grundlagen der Lesekompetenz: Mehrebenenverarbeitung anhand multipler Informationsquellen*, (w:) Artelt C. i in. (red.), *Struktur, Entwicklung und Förderung von Lesekompetenz. Vertiefende Analysen im Rahmen von PISA 2000*. Wiesbaden: VS Verlag für Sozialwissenschaften, s. 61–99.
- Stöckl H. (2004), *Typographie: Gewand und Körper des Textes – Linguistische Überlegungen zu typographischer Gestaltung*. "Zeitschrift für Angewandte Linguistik", nr 41, s. 5–48.
- Stöckl H. (2011), *Sprache-Bild-Texte lesen. Bausteine zur Methodik einer Grundkompetenz*, (w:) Diekmannshenke H., Klemm M., Stöckl H. (red.), *Bildlinguistik: Theorien – Methoden – Fallbeispiele*. Berlin: Erich Schmidt Verlag, s. 45–70.
- Wildfeuer J., Bateman J. A., Hiippala T. (2020), *Multimodalität. Grundlagen, Forschung und Analyse – eine problemorientierte Einführung*. Berlin/Boston: Walter de Gruyter Verlag.

- Wittstruck W. (2018), *Sprache-Bild-Kombinationen – Lehren und Lernen multimodal: ein Überblick mit Beispielen für die Arbeit im Unterricht DaF*, (w:) Kaunzner U. A., Reeg U., Ehrhardt C. (red.), Bild und Sprache. Impulse für die Arbeit im Unterricht DaF. Münster – New York: Waxmann Verlag, s. S. 11–40.
- Wolf W. (2002), *Towards a Functional Analysis of Intermediality: The Case of Twentieth-Century Musicalized Fiction*, (w:) Hedling E., Lagerroth U.-B. (red.), Cultural Functions of Intermedial Exploration. Amsterdam – New York: Brill/Rodopi, s. 15–34.

Received: 08.10.2021

Revised: 13.01.2022