

Synergie von Sprachlichkeit und Bildlichkeit im DaF-Unterricht: Eine qualitativ-deskriptive Fallstudie zum Textverständnis in „Deutsch Perfekt“

Hilal Balcı , Istanbul – Handan Köksal , Edirne

 <https://doi.org/10.37583/diyalog.1828020>

Abstract (Deutsch)

In dieser Arbeit wird untersucht, wie die Kombination von Bild und Text in der Zeitschrift *Deutsch Perfekt* das Textverständnis von DaF-Lernenden beeinflusst. Gestützt auf text- und pragmalinguistische Ansätze (Beaugrande / Dressler 1981; Brinker et al. 2014) sowie Kjørups (1978) Theorie der *pictorial speech acts* wird angenommen, dass kohärente Bild-Text-Cluster kognitive Entlastung bieten. Im Rahmen eines qualitativ-deskriptiven Fallstudiendesigns wurde von 18 DaF-Studierenden der Vorbereitungsklasse der Istanbul Universität-Cerrahpaşa ein theoriegeleiteter Fragebogen zu einem isolierten Bild und dessen Einbettung im Artikel beantwortet. Alle Teilnehmenden identifizierten das Nokia-Tastentelefon korrekt, verknüpften es mit Nostalgie und veralteter Technik und erkannten größtenteils eine illokutionäre Funktion. Mehr als die Hälfte der Teilnehmenden berichtete, dass das Bild zur Schaffung von Kohärenz beigetragen und das Textverständnis erleichtert hat. Die Ergebnisse der Arbeit belegen, dass Bilder in DaF-Materialien als eigenständige kommunikative Akte zu betrachten sind. Zudem tragen sie wesentlich zum Lernprozess bei, indem sie lernrelevante Prädikationen liefern und somit nicht lediglich einer dekorativen Funktion dienen.

Schlüsselwörter: Bild-Text-Synergie, Kohärenz, Textverständnis, Deutsch als Fremdsprache (DaF), multimodale Lehrmaterialien.

Abstract (English)

Synergy of linguistic and visual modes in GFL instruction: A case study on text comprehension in “Deutsch Perfekt”

The present study examines the impact of the combination of image and text in the magazine *Deutsch Perfekt* on the text comprehension of learners of German as a Foreign Language (GFL). Drawing on text- and pragmalinguistic approaches (Beaugrande & Dressler 1981; Brinker et al. 2014) and on Kjørup’s (1978) theory of *pictorial speech acts*, it is hypothesized that coherent image–text clusters reduce cognitive load. In a qualitative case-study design, 18 GFL students in the preparatory class at Istanbul University-Cerrahpaşa completed a theory-driven questionnaire on an isolated picture and its integration into the relevant article. All participants correctly identified the Nokia feature phone, associated it with nostalgia and outdated technology, and largely attributed an illocutionary illustrative function to it. More than half reported that the image fostered textual coherence and facilitated comprehension. The findings demonstrate that images in GFL materials should be regarded as independent communicative acts. They

make a substantial contribution to learning by providing content-relevant predications rather than serving a merely decorative purpose.

Keywords: *image–text synergy, coherence, text comprehension, German as a Foreign Language (GFL), multimodal instructional materials.*

EXTENDED ABSTRACT

The objective of this study is to examine the effect of the combination of image and text created in materials prepared for teaching German as a Foreign Language (GFL) on text perception and text comprehension processes. The present study selected *Deutsch Perfekt* magazine, a publication prepared for use by German learners, as its sample material. The magazine's objective is twofold: first, to facilitate the comprehension of linguistic elements by employing images in an effective manner; and second, to present contemporary, original, and cultural subjects to learners of GFL at various language levels.

In this context, the study examines the coherence between text and image based on text- and pragmalinguistic (action-oriented) approaches. Coherence is defined as the semantic integrity created by establishment of contextual connections in a text and thereby supporting meaning-making. As Beaugrande and Dressler (1981) argue, coherence is not only a characteristic of the text itself, but also a consequence of its interaction with the reader's cognitive processes. Therefore, the combination of images and text can facilitate comprehension by making conceptual relationships visible. According to Kress and van Leeuwen (2001), text and image can generate meaning in various ways through the coexistence of different semiotic modes functioning complementarily and reinforcingly. Similarly, Stöckl (2004) notes that images can concretize content, contribute to the structuring of the text, and support semantic clarity through their composition. Thus, text-image interaction is pivotal for GFL comprehension. According to the action-oriented methodology that forms the basis of this research, a relationship between the use of words and images in their respective settings must be established (Ernst 2002). This is related to Searle's (1969) and Austin's (1962) notion of speech acts. According to this idea, the communication context shapes the performative (illocutionary) functions that are present in every linguistic phrase. By recognizing images as communicative actions (pictorial speech acts) and highlighting the fact that images' meanings can only be understood in context, Kjörup (1978) extended the conventional speech act theory to include images. It is posited that pictures serve as a secondary means of reinforcing meaning in this scenario.

The study was designed as a qualitative descriptive case study. A questionnaire was developed based on Kjörup's theory of pictorial speech acts to serve as a data collection instrument. The questionnaire was subjected to a preliminary trial. The primary research was conducted with a sample of 18 students enrolled in the German preparatory class at Istanbul University-Cerrahpaşa. The responses given by the participants were then subjected to analysis, with the analysis being carried out using a theoretical classification system that had been predetermined, formulated based on previous studies in this field. The analysis revealed that the interaction between images and texts was perceived at various levels.

According to the study's findings, participants emphasized the image's nostalgic, technical, and emotional qualities while evaluating it as generally instructive and explanatory. The majority of respondents viewed image-text coherence favorably, with the visual element helping to improve text comprehension. Some participants, however, claimed that the image's contribution to the text was little and that its existence had no effect on how they understood the text.

In conclusion, the findings of this study demonstrate that images are not merely decorative elements, but rather independent communicative acts that contribute directly to the formation of meaning. In particular, the recognition of the performative (illocutionary) functions of images may assist in enhancing the effectiveness of image-text combinations from a didactic perspective. This finding lends support to Mayer's (2002) assertion that complementary text-image pairs facilitate more effective integration of information within the cognitive system. However, individual differences were observed in the perception of images. While some participants found image aids to be beneficial in facilitating understanding, others considered them to be of little or no significance. This situation demonstrates that learning styles can influence the manner in which images are perceived. In this context, Grein (2019) posits that multimodal teaching facilitates more effective management of individual differences and enables personalization of learning.

The present study is constrained by its exclusive focus on a limited sample of participants, thereby restricting the extent to which its findings can be generalized. It is recommended that future

studies be conducted with larger and more heterogeneous samples. In addition, the integration of performance-based evaluation instruments is recommended to attain a more objective estimation of the impact of images on text comprehension. Furthermore, the flexibility offered by digital learning environments to personalize textual and visual elements according to different learning styles also holds significant potential for further research. The capacity of adaptive learning systems to present images according to individual needs and preferences has the potential to contribute to the development of more efficient applications in language teaching by increasing the effectiveness of image-text compatibility.

Einleitung

Sprachlichkeit und Bildlichkeit spielen in Lehrmaterialien für Deutsch als Fremdsprache eine wesentliche Rolle, unter anderem auch in der Zeitschrift *Deutsch Perfekt*, die speziell für DaF-Lernende entwickelt wurde. Neben der deutschen Sprache erhalten die Lernenden auch Einblicke in den Alltag, die Kultur und die Traditionen des deutschsprachigen Raums. Die visuelle Gestaltung durch Illustrationen, Infografiken und Bilder ergänzt die sprachlichen Elemente gezielt, erleichtert somit das Verständnis komplexer Themen. Die Verknüpfung von Sprachlichkeit und Bildlichkeit trägt zur Förderung von Kohärenz und Verständnis bei, die als wesentliche Faktoren gelten, welche den Lernprozess beeinflussen und das Verstehen komplexer sprachlicher Inhalte erleichtern.

DaF-Lernende werden beim Verstehen von Texten oft mit Schwierigkeiten konfrontiert, denn die Erschließung der Inhalte wird durch unbekanntes Wortschatz, komplexe Satzstrukturen sowie fehlende kontextuelle und kulturelle Vorkenntnisse erschwert. Visuelle Anknüpfungspunkte wie Bilder, Illustrationen oder Infografiken können die Erschließung von Textinhalten erleichtern und zugleich die Kohärenz und Verständlichkeit fördern. Ein moderner DaF-Unterricht im 21. Jahrhundert verlangt Lehr- und Lernmaterialien, die didaktisch gut durchdacht und multimodal gestaltet sind. Insbesondere im Kontext des autonomen Lernens, das über das im Unterricht übliche Curriculum hinausgeht, sind multimodal gestaltete Materialien unabdingbar. Durch die sinnvolle Verbindung sprachlicher und visueller Elemente werden die Lernenden nicht nur angesprochen, sondern auch das Verstehen und Einprägen neuer Inhalte wird erleichtert.

Ausgehend von dieser Idee ergibt sich die Frage, welche Beispiele es im DaF-Bereich gibt, die sowohl autonomes Lernen fördern als auch durch eine gezielte Verbindung von Sprachlichkeit und Bildlichkeit das Verständnis unterstützen. Die Zeitschrift *Deutsch Perfekt*, die speziell für Selbstlernende konzipiert ist und auch außerhalb des klassischen Curriculums eingesetzt werden kann, gilt als eine besonders gelungene und aktuelle Umsetzung eines solchen Konzeptes. Aufgrund ihrer gezielten Verzahnung von Text und Bild eignet sich *Deutsch Perfekt* hervorragend, um zu analysieren, inwieweit die bewusste Kombination sprachlicher und visueller Informationen in der Praxis zu einem verbesserten Verstehensprozess beiträgt.

Ziel und Relevanz der Arbeit

Das Ziel dieser Arbeit ist es, am Beispiel der Zeitschrift *Deutsch Perfekt* zu untersuchen, wie die Synergie von Sprachlichkeit und Bildlichkeit in ihren spezifischen Kontexten zur Wahrnehmung, Interpretation und zum Verständnis von Texten beiträgt. Vor dem Hintergrund, dass Bilder sowohl als Stimulus dienen als auch bestimmte Sprechakte initiieren können, wird im Folgenden analysiert, inwieweit dieses Potenzial in *Deutsch Perfekt* tatsächlich wirksam wird. Diesbezüglich wurden folgende Forschungsfragen entwickelt:

1. Wie interpretieren DaF-Lernende ein isoliertes Bild aus *Deutsch Perfekt* hinsichtlich
 - 1.1. des Hauptobjekts bzw. der dargestellten Situation
 - 1.2. der vermuteten illokutionären Funktion und
 - 1.3. der wahrgenommenen Merkmale?
2. Wie bewerten DaF-Lernende die gezielte Kombination von Bild und Text in *Deutsch Perfekt* bezüglich ihres Textverständnisses?
3. Welche Empfehlungen lassen sich aus den Erkenntnissen für die Gestaltung und den Einsatz vergleichbarer DaF-Materialien ziehen?

Die vorliegende Arbeit zeigt am Beispiel von *Deutsch Perfekt* auf, wie sprachliche und visuelle Elemente in DaF-Materialien kohärent zusammenwirken können und welche kommunikativen Funktionen die Bilder übernehmen. Gerade für die Entwicklung und Erprobung von DaF-Lehrmaterialien ist die systematische Verzahnung von sprachlichen und visuellen Elementen unerlässlich. Dass eine solche Verzahnung lernförderlich ist, belegt etwa die Studie von Çelikkaya und Karasu (2025: 475) zur Bild-Text-Kohärenz im Lehrwerk *Mit Deutsch A1.1*, in der kongruente, komplementäre und elaborative Relationen den Wissenserwerb signifikant unterstützen. Die Arbeit leistet somit einen Beitrag zur Fremdsprachendidaktik, indem sie einerseits einen theoretischen Rahmen zur Analyse von Text-Bild-Kombinationen bereitstellt und andererseits konkrete Empfehlungen zur Auswahl und Gestaltung von DaF-Materialien aufzeigt. Dies ist insbesondere vor dem Hintergrund aktueller Lehr- und Lernkonzepte von Relevanz, da die effiziente Vermittlung von Fremdsprachen nicht allein auf sprachlichen, sondern zunehmend auch auf multimodalen Materialien fußt.

Text-Bild-Kohärenz in der Textlinguistik

Kohärenz ermöglicht es, sprachliche und inhaltliche Verbindungen innerhalb eines Textes herzustellen (vgl. Brinker et al. 2014: 12), was das Verstehen und Behalten von Informationen unterstützt. Beaugrande und Dressler (1981) definieren Kohärenz als „*configuration of concepts and relations which underlie the surface text*“. Sie betonen, dass Kohärenz nicht allein eine Eigenschaft des Textes ist, sondern durch die Interaktion mit den kognitiven Prozessen der Rezipierenden entsteht. Dies verdeutlicht, dass auch Bilder in Text-Bild-Kombinationen eine entscheidende Rolle spielen, indem sie konzeptuelle Relationen sichtbar machen und das inhaltliche Verständnis fördern. Vor allem ist in der Fremdsprachendidaktik die Schaffung kohärenter und verständlicher Inhalte besonders relevant, da sie den Zugang zu sprachlichen Strukturen und die Entwicklung von Kommunikationsfähigkeiten fördert.

Laut Kress und van Leeuwen (2001) kann die gezielte Kombination verschiedener semiotischer Modi, wie Text und Bild, bei der Gestaltung eines Produkts oder Ereignisses unterschiedliche Wirkungen erzielen. Diese können sich einerseits

gegenseitig verstärken, indem sie denselben Inhalt auf verschiedene Arten vermitteln. Doch andererseits können sie komplementäre Rollen einnehmen, indem Text beispielsweise erklärt und Bild veranschaulicht wird. Darüber hinaus ist es möglich, dass Text und Bild hierarchisch angeordnet sind, etwa wenn der Text die Hauptinformation liefert und das Bild ergänzend oder unterstützend wirkt (vgl. Kress / van Leeuwen 2001: 20). Die Kombination dieser Elemente spielt eine entscheidende Rolle für die Synergie zwischen Sprachlichkeit und Bildlichkeit, da sie die Kohärenz und Verständlichkeit von Text-Bild-Kombinationen fördern.

Auch nach Stöckl (2004) erfüllen Bilder mehrere zentrale Funktionen: Sie veranschaulichen Inhalte, tragen zur Strukturierung eines Textes bei und fördern durch ihre Komposition die semantische Klarheit. Dabei beeinflussen sie nicht nur mögliche Bedeutungen, sondern schaffen auch eine soziale Interaktion zwischen den Bildproduzierenden und -rezipierenden (vgl. Stöckl 2004: 22). In diesem Zusammenhang ist die Interaktion zwischen sprachlichen und bildlichen Elementen in Texten besonders bedeutsam. Beaugrande und Dressler (1981) heben hervor, dass Kohärenz durch „*spreading activation of concepts and relations*“ entsteht, wobei sich die mentale Verarbeitung durch visuelle Unterstützung erleichtern lässt. Dies bedeutet, dass visuelle Elemente nicht isoliert betrachtet werden können, sondern in direkter Verbindung mit den sprachlichen Komponenten stehen und gemeinsam zur Verständlichkeit beitragen. Bilder können somit sprachliche Inhalte ergänzen, indem sie abstrakte Konzepte veranschaulichen und durch ihre kontextuelle Einbettung das Verständnis komplexer sprachlicher Strukturen erleichtern. Durch die Einbindung von Bildern wird das Verständnis gezielt gefördert, da visuelle Elemente eine anschauliche Unterstützung bieten und das Erfassen und Verarbeiten von sprachlich komplexen Inhalten erleichtern. Die kontextuelle Einbettung von Bildern fördert somit nicht nur die Verständlichkeit, sondern schafft auch eine Art Bild-Text-Synergie, bei der sich die semantischen und kognitiven Schwächen des einen Zeichensystems durch die Stärken des anderen ausgleichen (vgl. Stöckl 2004: 249). Die Kombination von Bildern und Texten stellt daher ein didaktisch wertvolles Mittel zur Verständlichkeit dar und unterstützt Lernende dabei, Verknüpfungen herzustellen und Inhalte kohärent zu erfassen.

Storror (2020: 119ff.) hebt hervor, dass die Kohärenz zwischen Text und Bild eine entscheidende Rolle für die Darstellung thematischer Zusammenhänge spielt. Sie zeigt, dass Text-Bild-Cluster dazu beitragen, thematische Zusammenhänge zu verdeutlichen und zusätzliche Informationen bereitzustellen, die weder allein aus dem Text noch aus dem Bild vollständig erschlossen werden können. Die Analyse der Beziehung zwischen Bild, Bildunterschrift und Fließtext ist besonders relevant, da diese entweder ergänzend oder redundant wirken und so die inhaltliche Kohärenz fördern. Nach Opilowski (2013: 220) beeinflussen visuelle Elemente wie Bilder, Grafiken, Schrifttypen oder Farben in unterschiedlichem Maße sowohl die Form als auch die Bedeutung eines Textes. Werden jedoch diese Elemente ignoriert und die Analyse ausschließlich auf den sprachlichen Textteil beschränkt, können die Erkenntnisse unvollständig oder fehlerhaft ausfallen. Auch Dürscheid (2007: 4f.) beschreibt in ihrer

Arbeit, dass die visuelle Gestaltung von Texten, wie Absätze, Einrückungen oder das Layout, eine zentrale Rolle für die Kohärenz und Lesbarkeit spielt. Des Weiteren hebt sie hervor, dass derartige gestalterische Elemente wesentlich dazu beitragen, Informationen klar zu strukturieren und die Aufnahme sowie das Verständnis von Inhalten zu erleichtern. In diesem Zusammenhang verweist sie darauf, dass „der Brückenschlag von der Schrift zum Bild gar nicht so groß [ist]: Der Text vereint nach dieser Auffassung alle Zeichenmodalitäten in sich – und damit auch Schrift und Bild“ (Dürscheid 2007: 5). Hiermit werden die enge Verbindung und das synergetische Zusammenwirken von Schrift und Bild verdeutlicht.

Nachdem die Synergie zwischen Sprachlichkeit und Bildlichkeit aus unterschiedlichen Perspektiven beleuchtet wurde, lassen sich einige zentrale Gemeinsamkeiten ableiten. Gemeinsam ist, dass Sprachlichkeit und Bildlichkeit nicht isoliert wirken und im Kontext des Fremdsprachenunterrichts ihre Synergie eine entscheidende Rolle bei der Verständnisförderung spielt und darüber hinaus ermöglicht, Lerninhalte anschaulicher und zugänglicher zu gestalten. Um die kommunikativen Funktionen dieser Relationen zu erfassen, wird im nächsten Abschnitt deren pragmatische Perspektive thematisiert.

Pragmalinguistische Perspektive und Sprechakttheorie

Ausgehend von den zuvor dargestellten Erkenntnissen stellt sich die Frage, inwiefern diese Synergie tatsächlich wahrgenommen wird und ob sie die beabsichtigte unterstützende Wirkung entfaltet. Denn die Wahrnehmung und Verarbeitung dieser Synergie hängen nicht nur von der sprachlichen und visuellen Gestaltung der Materialien ab, sondern auch von den Lernenden selbst, die ihre kognitiven und kommunikativen Kompetenzen sowie ihre situativen Wahrnehmungen in die Rezeption einbringen. Die Wahrnehmung hängt von einer Kombination struktureller und funktionaler Faktoren ab, die sich anhand theoretischer Ansätze der Textlinguistik und Pragmalinguistik analysieren lassen.

Innerhalb der Textlinguistik unterscheidet man zwischen der sprachsystematisch ausgerichteten Textlinguistik und der kommunikationsorientierten Textlinguistik. Während die sprachsystematisch ausgerichtete Textlinguistik sich auf die syntaktisch-semantische Struktur eines Textes fokussiert, bietet die kommunikationsorientierte Textlinguistik eine erweiterte Perspektive, die sich auf die Funktionalität und situative Einbettung von Texten konzentriert (vgl. Brinker et al. 2014: 13ff.). Aus der Perspektive der sprachsystematisch ausgerichteten Textlinguistik lassen sich die syntaktisch-semantischen Verknüpfungen zwischen Sprachlichkeit und Bildlichkeit analysieren, beispielsweise wie Bilder Textaussagen stützen oder ergänzen. Allerdings reicht die strukturelle Perspektive allein nicht aus, um zu erklären, wie die Synergie von den Lernenden wahrgenommen wird. In diesem Zusammenhang bietet die kommunikationsorientierte Textlinguistik eine funktionale Perspektive, die Texte als sprachliche Handlungen versteht, die in spezifischen Kommunikationssituationen eingebettet sind und bestimmte Intentionen verfolgen, wie das Informieren oder

Appellieren (vgl. Brinker et al. 2014: 16). Somit erfüllen Texte mit eingebetteten Bildern, bestimmte Funktionen wie das Informieren oder Appellieren, die entscheidend beeinflussen, wie Lernende Inhalte wahrnehmen und verarbeiten. Diese Perspektive ermöglicht eine erweiterte Analyse, wie Texte, in Kombination mit Bildern, nicht nur syntaktisch-semantisch kohärent, sondern auch kommunikativ wirksam sein können, um verständnisfördernd zu wirken. Durch die Einbeziehung der sprachsystematischen und der kommunikationsorientierten Textbegriffe wird eine umfassendere Analyse der verständnisfördernden Funktion der Synergie für die Lernenden ermöglicht. Dennoch bleibt die Frage bestehen, wie Lernende diese Synergie wahrnehmen und in ihre individuellen Verstehensprozesse integrieren. Die Pragmalinguistik bietet in diesem Zusammenhang einen erweiterten Rahmen, indem sie untersucht, wie Sprachlichkeit und Bildlichkeit in ihren spezifischen Kontexten wahrgenommen und interpretiert werden.

Im Hinblick der Pragmalinguistik ist Sprache und damit auch ihre visuelle Ergänzung durch Bilder nicht isoliert zu betrachten, sondern immer in Verbindung mit der Umwelt, d. h. der konkreten Situation, in der sie verwendet wird (vgl. Ernst 2002: 12). Diese Situation, der sogenannte Kontext, umfasst sowohl die physischen, sozialen und kulturellen Bedingungen als auch die Intentionen des Senders und das Weltwissen und die Erfahrungen der Rezipierenden, in diesem Fall der Lernenden. In Bezug auf die Synergie von Sprachlichkeit und Bildlichkeit bedeutet dies, dass die verständnisfördernde Funktion stark davon abhängt, wie Lernende den Kontext wahrnehmen und interpretieren. Die kontextabhängigen Prozesse können anhand der Sprechakttheorie von Austin (1962) und Searle (1969) analysiert werden, einem wichtigen Teilbereich der Pragmalinguistik (vgl. Ernst 2002: 98). Nach der Sprechakttheorie besteht jede sprachliche Äußerung aus mehreren Ebenen, die als Sprechakte bezeichnet werden. Austin und Searle haben unterschiedliche, jedoch miteinander verwandte Modelle zur Struktur von Sprechakten entwickelt, die in Tabelle 1 zusammengefasst sind.

| | | | | | |
|------------------|------------------------------------|---|-------------------------------------|---|------------------------------------|
| Austin [1962] | Äußerung von Sprechlauten | Äußerung von Lauten in best. gram. Strukturen | Etwas über etwas aussagen | Angabe der performativen Verwendung der Proposition | intendierte Wirkung des Sprechakts |
| | → Phonetischer Akt | → Phatischer Akt | → Rhetischer Akt | Illokutiver Akt | Perlokutiver Akt |
| Searle [1969] | Äußerungsakt (→ Lokution) | | Propositionaler Akt (→ Proposition) | Illokutiver Akt (→ Illokution) | Perlokutiver Akt (→ Perlokution) |
| | → Referenz / Bezugnahme auf „Welt“ | | → Prädikation / Aussage über „Welt“ | | |

Tab. 1: Struktur von Sprechakten (vgl. Bußmann 2008: 674).

Austin (1962: 94f.) unterscheidet in seinem Ansatz innerhalb des lokutionären Akts drei Ebenen. Der phonetische Akt bezieht sich auf die Artikulation von Lauten, während der phatische Akt die korrekte Bildung grammatischer Strukturen umfasst. Der rhetische Akt hingegen bezieht sich auf die Aussage über ein bestimmtes Thema bzw. auf Sinn und Referenz. Auf dieser Grundlage beschreibt Austin den illokutiven Akt, der die kommunikative Funktion der Äußerung darstellt, wie z.B. eine Behauptung, Frage oder Aufforderung. Schließlich beschreibt der perlokutive Akt die tatsächliche Wirkung der Äußerung (vgl. Bußmann 2008: 674f.). Searle (1969: 22f.) gliedert diese Ebenen anders. Der Äußerungsakt (Lokution), der Austins phonetischem und phatischem Akt entspricht, beschreibt die sprachliche Formulierung einer Äußerung. Der propositionale Akt (Proposition), der Austins rhetischem Akt entspricht, überträgt den Inhalt und die Bedeutung einer Äußerung. Der illokutive Akt (Illokution) bleibt als zentrale Ebene stehen und beschreibt die kommunikative Funktion und Intention der Äußerung. Der perlokutionäre Akt (Perlokution) fokussiert sich auf die tatsächliche Wirkung der Äußerung (vgl. ebd.: 674f.). Nach der Darlegung der Grundlagen der Sprechakttheorie, wie sie von Austin (1962) und Searle (1969) entwickelt wurden, stellt sich allerdings die Frage, inwiefern diese Konzepte auch auf visuelle Kommunikationsformen übertragbar sind.

Bilder als kommunikative Akte: Pictorial Speech Acts

Während die traditionelle Sprechakttheorie sich in erster Linie auf verbale Sprache fokussiert, erweitert Kjörup (1978) dieses Modell in seiner Arbeit *Pictorial Speech Acts* auf die Bildlichkeit und betrachtet Bilder als intentionale kommunikative Akte. Kjörup entwickelt auf der Grundlage von Goodmans Arbeit *Languages of Art: An Approach to a Theory of Symbols*, in der die Theorie der symbolischen Darstellung formuliert wird, die Idee, dass auch Bilder als intentionale Sprechakte zu betrachten sind, die sowohl Referenzen herstellen *representation-of* als auch Bedeutungen ausdrücken können *representation-as* (vgl. Goodman 1969, zit. nach Kjörup 1978: 55f.). Kjörup argumentiert, dass Bilder ähnlich wie verbale Sprache illokutionäre Akte ausführen können, wie etwa das Darstellen, Warnen oder Illustrieren. So sei „der Produzent des Bildes [...] in gewisser Weise in der Lage, durch das Bild eine Handlung vorzunehmen, die als bildlicher Sprechakt betrachtet werden kann, ähnlich dem Sprechakt einer sprachlichen Beschreibung“ (ebd.: 60). Von besonderer Relevanz ist seine Definition von Illustrationen als „sekundäre Nutzung von Bildern“, die in Verbindung mit Texten zur Verdeutlichung zentraler Inhalte und zur Förderung von Verständlichkeit dienen (ebd.: 68). Ein weiterer zentraler Aspekt von Kjörups Ansatz ist die Kontextabhängigkeit von Bildern. Es wird betont, dass die Bedeutung eines Bildes nicht isoliert betrachtet werden kann, denn ihre Referenz und Prädikation werden oft durch den interpretativen Rahmen einer kommunikativen Situation, eines Titels oder eines Textes unterstützt (vgl. ebd.: 63f.).

Auf die Synergie von Sprachlichkeit und Bildlichkeit übertragen, bietet Kjörups Ansatz eine wertvolle Grundlage. Die Betonung der Kontextabhängigkeit von

Bildinterpretation verdeutlicht, dass sowohl sprachliche als auch visuelle Elemente nur im Rahmen ihrer kommunikativen Gesamtsituation vollständig erfasst werden können. Ferner wird hervorgehoben, dass Illustrationen als „sekundäre Nutzung“ von Bildern und in Verbindung mit Texten zur Verständlichkeit beitragen können, indem sie zentrale Inhalte visuell verdeutlichen und kohärente Einheiten schaffen. Dies legt nahe, dass die Effektivität der Synergie von Sprachlichkeit und Bildlichkeit grundlegend davon abhängt, wie gut die einzelnen Elemente aufeinander abgestimmt sind und ob sie in der Lage sind, die kognitiven Anforderungen der DaF-Lernenden in spezifischen Kontexten zu erfüllen. Wie und in welchem Ausmaß diese Synergie tatsächlich zum Verstehensprozess beiträgt, lässt sich durch Kjörups Ansatz untersuchen, der für diese Arbeit einen grundlegenden theoretischen Anknüpfungspunkt bietet. Dieser Ansatz unterstreicht, dass Bilder in einem bestimmten Kontext als illokutionäre Akte wirken können, indem sie referenzielle und prädikative Aussagen über ihren Gegenstand machen und in einer kommunikativen Situation eingesetzt werden (vgl. Kjörup 1978: 63f.).

In *Deutsch Perfekt* ließe sich eine solche Perspektive nutzen, um zu fragen, ob ein Bild allein bereits eine Aufforderung, Warnung oder informierende Funktion erfüllt oder erst im Zusammenspiel mit dem Text seine Bedeutung entfaltet. Die oft beschriebene ‚Text-Bild-Schere zeigt‘, dass Bilder und Texte nicht immer kongruent sind und dadurch neue Sprechakte auslösen können (vgl. Hosch / Macaire 1991: 25). Die Zeitschrift *Deutsch Perfekt* richtet sich an DaF-Lernende verschiedener Niveaustufen und bietet somit ein breites Spektrum an didaktisch aufbereiteten Texten. Neben der sprachlichen Reduktion, wie Vokabelhilfen oder vereinfachte Satzstrukturen, setzt die Zeitschrift konsequent Bilder, Illustrationen und Infografiken ein, die DaF-Lernende beim Erschließen neuer Inhalte unterstützen sollen. Dadurch entsteht eine enge Verknüpfung von Sprachlichkeit und Bildlichkeit, die eine wirksame Bild-Text-Synergie repräsentiert. Darüber hinaus werden in *Deutsch Perfekt* authentische und aktuelle Themen aus dem deutschsprachigen Raum behandelt, welche für das Verstehen kultureller und gesellschaftlicher Bezüge in einer Fremdsprache essenziell sind. Dass Bilder sich als Gesprächsimpuls eignen, verdeutlicht bereits Hosch und Macaire (1991: 25): Lernende können in praktischen Unterrichtssituationen ein Bild auswählen und darüber sprechen. Auf ähnliche Weise ließen sich auch in *Deutsch Perfekt* Bilder als Impuls für Sprechhandlungen einsetzen, etwa um bestimmte lokutionäre Akte oder illokutionäre Funktionen auszulösen.

Forschungsdesign

Die vorliegende Arbeit folgt einem qualitativ-deskriptiven Fallstudiendesign, das darauf abzielt, die Wirkung einer bestimmten Bild-Text-Kombination aus der Zeitschrift *Deutsch Perfekt* im DaF-Unterricht detailliert und theoriegeleitet zu beschreiben, um daraus Rückschlüsse auf die Synergie von Sprachlichkeit und Bildlichkeit im Hinblick auf das Textverständnis in einem DaF-Unterricht zu ziehen.

Nach Creswell und Creswell (2023) ist die Fallstudie eine Form des qualitativen Forschungsdesigns, bei der „eine vertiefte Analyse eines Falles, wie etwa eines Programms, Ereignisses, Prozesses oder einer oder mehrerer Personen, innerhalb eines begrenzten zeitlichen und räumlichen Rahmens“ verfolgt wird (Creswell / Creswell 2023: 38). Die Arbeit ist zugleich deskriptiv ausgerichtet, was bedeutet, dass es nicht um das Erklären oder Verallgemeinern, sondern um die genaue, theoriebasierte Beschreibung eines spezifischen Falles geht. Die deskriptive Fallstudie zielt darauf ab, das untersuchte Phänomen durch eine dichte Beschreibung ‚*thick description*‘, also durch eine detaillierte und vielschichtige Darstellung, erfassbar zu machen. In Anlehnung an Mills und Gay (2019) bedeutet deskriptiv in diesem Zusammenhang: „Wenn man sagt, eine Fallstudie sei deskriptiv, dann bedeutet dies, dass das Endergebnis der Fallstudie, der narrative Bericht, eine ‚dichte Beschreibung‘ des untersuchten Phänomens enthält - das schließt die Einbeziehung vieler Variablen und die Analyse ihrer Wechselwirkungen ein“ (Mills / Gay 2019: 405).

Datenerhebung

Im Rahmen der Arbeit kam der Text ‚Handys – wie früher?‘ aus der Zeitschrift *Deutsch Perfekt*, Ausgabe Nr. 1/ 2025, zum Einsatz, der in Abbildung 1 zu sehen ist. Der Text wurde bewusst ausgewählt, da er aus einer aktuellen Ausgabe stammt und sich durch Alltagsnähe sowie eine ansprechende Darstellung auszeichnet. Die Datenerhebung stützt sich auf einen Fragebogen, der nach Kjørups (1978) Arbeit ‚*Pictorial Speech Acts*‘ entwickelt wurde. Der Fragebogen umfasst sowohl offene als auch geschlossene Fragen zum isolierten Bild und zur Bild-Text-Beziehung, die die Wahrnehmung, Interpretation und funktionale Bewertung von Bild und Bild-Text-Kombinationen erfassen.

The image shows a page from the magazine 'Deutsch Perfekt' with the title 'Handys – wie früher?'. On the left is a photograph of a classic Nokia 3210 mobile phone. The main text discusses the popularity of feature phones and introduces a new version of the Nokia 3210. There are several side columns containing short articles and advertisements, including one for 'Auf Tiktok feiert eine junge Community von Vintage-Fans das Revival der alten Handys.' and another about 'Marktforschungsergebnisse'.

Abb. 1: ‚Handys – wie früher?‘, *Deutsch Perfekt*, Ausgabe Nr. 1/ 2025: 62-64.

Vor der eigentlichen Datenerhebung wurde eine Pilotstudie mit 11 Teilnehmenden durchgeführt. Die Pilotstudie hatte das Ziel, das Erhebungsinstrument hinsichtlich Verständlichkeit und Reliabilität zu erproben, um die inhaltliche Validität der Erhebung sicherzustellen und potenzielle Störfaktoren während der Hauptuntersuchung zu minimieren. Die eigentliche Datenerhebung erfolgte im Rahmen einer Fallstudie mit 18 DaF-Lernenden auf B1-Niveau in der Vorbereitungs-klasse der Istanbul Universität-Cerrahpaşa im akademischen Studienjahr 2024/ 2025. Die mit dem Fragebogen erhobenen Antworten bilden die Grundlage für die anschließende qualitativ-deskriptive Analyse und ermöglichen es, genau nachzuvollziehen, wie die untersuchte Bild-Text-Kombination in diesem Kontext wahrgenommen und bewertet wurde.

Die Auswertung des Fragebogens beruht auf einem theoriebasierten Kategoriensystem, das sich an den zuvor definierten Analysebereichen orientiert, darunter Wahrnehmung der Bildinhalte, illokutionäre Funktionen, visuelle Prädikationen, Synergieeffekte sowie Beitrag zur Verständlichkeit. Ziel war es, funktionale Muster der Bildverwendung zu identifizieren und systematisch herauszuarbeiten, inwiefern Bilder im Zusammenspiel mit Texten zur Förderung von Kohärenz und zur Unterstützung des Textverständnisses beitragen.

Zur besseren Übersicht über die Struktur und theoretische Fundierung des Erhebungsinstruments zeigt Tabelle 2 die einzelnen Fragen des Fragebogens, ihre jeweilige Zielsetzung sowie den zugrunde liegenden Theoriebezug.

| <i>Frage</i> | <i>Ziel der Frage</i> | <i>Theoriebezug</i> |
|---|--|---|
| 1) Worum / um wen geht es in diesem Bild hauptsächlich? | Wahrnehmung der Bildinhalte | Bilder haben keinen ‚automatischen‘ Referenten, sondern verweisen im kommunikativen Kontext auf etwas, das erkannt/entschlüsselt werden muss. |
| 2) Illokution des Bildes: Welchen Zweck könnte das Bild hier erfüllen? | Bildfunktionen und illokutionäre Wirkung | Kjørup spricht von verschiedenen illokutionären Funktionen: ‚depicting‘, ‚warning‘, ‚illustrating‘ etc. |
| 2a) Können Sie bitte erklären, warum Sie die von Ihnen in der vorherigen Frage ausgewählten Optionen gewählt haben? | Detaillierte Erläuterung zu der vorherigen Frage | |
| 3) Welche Eigenschaften oder Merkmale werden in diesem Bild über das dargestellte Objekt/ die Person vermittelt? | Vermittelte Eigenschaften und Eindrücke | Ein Bild ‚prädiert‘ etwas über den Referenten. |
| 4) Wie gut passt das Bild Ihrer Meinung nach zum Inhalt des Textes? | Kohärenz zwischen Bild und Text | Kjørups Unterscheidung zwischen dem ‚Bild als Symbol‘ und der kontextuellen Verwendung. |
| 5) In welchem Ausmaß hat das Bild Ihr Verständnis des Textes beeinflusst? | Einfluss auf das Textverständnis | Kjørup betont, dass Bilder gezielt Informationen hinzufügen oder die Aufmerksamkeit auf bestimmte Aspekte lenken. |

| | | | |
|----|---|---------------------------------|--|
| 6) | Wenn Sie den Text ohne dieses Bild gelesen hätten, wie sehr hätten Ihnen Ihrer Meinung nach Informationen oder Eindrücke gefehlt? | Wahrnehmung der Synergieeffekte | Kjørup betont, dass Bilder zusätzliche oder ergänzende Informationen liefern können. Ist der Text allein verständlich, oder bleibt etwas unklar, wo ein Bild hätte Abhilfe schaffen können? So wird das Fehlen der piktorialen Prädikation thematisiert. |
|----|---|---------------------------------|--|

Tab. 2: Fragebogen nach Kjørup (1978).

Befunde

In den folgenden Tabellen werden die Befunde basierend auf den Antworten der 18 Teilnehmenden dargestellt. Das theoriebasierte Kategoriensystem umfasst fünf Analysebereiche, die jeweils den Zielen der Fragen entsprechen. In der Spalte *n* sind die Teilnehmenden (P¹-P¹⁸) aufgeführt, die sich entsprechend geäußert haben. Die Spalte *N* gibt die absolute Häufigkeit an, mit der ein Code insgesamt vergeben wurde. Eine detaillierte Beschreibung der Anteile erfolgt anschließend in den Befundbeschreibungen.

| | <i>Kategorie</i> | <i>Code</i> | <i>N</i> | <i>N</i> |
|------------------------------------|-----------------------|------------------------|---|----------|
| <i>Wahrnehmung der Bildinhalte</i> | Alter / Zustand | ○ alt, älteres Modell | P ¹ , P ² , P ⁵ | 3 |
| | Bildschirm Elemente | ○ Emoji | P ⁶ , P ¹⁴ , P ¹⁷ | 3 |
| | Emotion / Ausdruck | ○ Angstthema | P ¹⁰ | 1 |
| | | ○ Ausdruck | P ¹⁰ , P ¹³ | 2 |
| | | ○ Gefühlszustand | P ¹³ | 1 |
| | Gerät | ○ Handy (Tastenshandy) | P ¹ -P ¹⁸ (P ⁴ , P ⁵ , P ⁶ , P ⁷ , P ⁸ , P ¹² , P ¹⁶) | 18 (7) |
| | Marke / Modell | ○ Nokia (Nokia 3310) | P ¹ , P ¹¹ , P ¹⁸ (P ¹¹) | 3 (1) |
| | Technologie / Kontext | ○ analog | P ¹⁰ | 1 |
| ○ digital | | P ¹⁰ | 1 | |
| | | Insgesamt | | 41 |

Tab. 3: Wahrnehmung der Bildinhalte.

Im Rahmen des Analysebereichs ‚Wahrnehmung der Bildinhalte‘ wurden sechs Kategorien gebildet und insgesamt 41 Codierungen mit neun unterschiedlichen Codes vorgenommen (Tabelle 3). Das Objekt wurde von allen 18 Teilnehmenden als ‚Handy‘ identifiziert. Sieben Teilnehmende erkannten es als ‚Tastenshandy‘, drei ergänzten die Beschreibung ‚älteres Modell‘ und ebenfalls drei nannten explizit ‚Nokia‘. Drei

Teilnehmende äußerten emotionale Zuschreibungen, drei Teilnehmende verwiesen auf das ‚Emoji‘ auf dem Display und eine/r ordnete es einem technologischen Kontext zu.

| <i>Hauptfunktion</i> | | <i>Begründungs-Code</i> | <i>N</i> | <i>N</i> |
|---|---|--------------------------------|---|----------|
| Darstellung eines Gegenstands / Person (informierend) | ○ | schlichte Beschreibung | P ¹ , P ³ | 2 |
| | ○ | nur ein Objekt | P ⁴ , P ⁶ , P ⁷ | 3 |
| | ○ | Nostalgie / früherer Gebrauch | P ¹ , P ⁵ | 2 |
| | ○ | Textbezug / Illustration | P ⁸ , P ¹³ | 2 |
| | ○ | Überraschung (Emoji) | P ³ , P ¹⁴ | 2 |
| | ○ | Raten / Intuition | P ⁵ , P ⁶ , P ⁸ , P ¹¹ , P ¹⁶ | 5 |
| | ○ | Verständlichkeitsunterstützung | P ¹³ , P ¹⁶ | 2 |
| Illustration / visuelles Textbeispiel | ○ | Textbezug / Illustration | P ² , P ⁸ , P ⁹ , P ¹³ , P ¹⁵ | 5 |
| | ○ | Raten / Intuition | P ⁸ , P ⁹ , P ¹¹ , P ¹⁶ , P ¹⁸ | 5 |
| | ○ | Verständlichkeitsunterstützung | P ¹³ , P ¹⁶ | 2 |
| | ○ | Interesse wecken | P ¹⁵ | 1 |
| Warnung / Hinweis auf ein Problem | ○ | Problem / Warnung | P ¹⁰ , P ¹² , P ¹⁷ | 3 |
| | ○ | Textbezug / Illustration | P ¹⁵ | 1 |
| | ○ | Überraschung (Emoji) | P ¹⁷ | 1 |
| | ○ | Interesse wecken | P ¹⁵ | 1 |
| Appell / Aufforderung | ○ | Appell | P ⁵ , P ¹¹ | 2 |
| | ○ | Nostalgie | P ⁵ , P ¹¹ | 2 |
| | ○ | Raten / Intuition | P ⁵ | 1 |
| Insgesamt | | | | 42 |

Tab. 4: Bildfunktionen und illokutionäre Wirkung.

Im Rahmen des Analysebereichs ‚Bildfunktionen und illokutionäre Wirkung‘ standen vier unterschiedliche Antwortoptionen zur Verfügung (Tabelle 4). Mehrfachantworten waren möglich. Die Teilnehmenden sollten anschließend ihre Auswahl begründen. Aus diesen Antworten wurden insgesamt 25 Codierungen mit 18 unterschiedlichen Codes vorgenommen. Elf Teilnehmende ordneten das Bild der Funktion ‚Darstellung eines Gegenstands bzw. einer Person (informierend)‘ zu und bildeten damit die Mehrheit. Acht Teilnehmende interpretierten das Bild als ‚Illustration bzw. visuelles Textbeispiel‘, vier wählten ‚Warnung bzw. Hinweis auf ein Problem‘ und zwei ‚Appell bzw. Aufforderung‘. Für ‚Darstellung eines Gegenstands bzw. einer Person (informierend)‘ wurden 18 Codierungen vergeben, verteilt auf sieben Begründungscodes. Für ‚Illustration / visuelles Textbeispiel‘ entfallen 13 Codierungen zu vier Begründungscodes. Für ‚Warnung / Hinweis auf ein Problem‘ ergeben sich sechs

Codierungen zu vier Begründungscodes. Für ‚Appell / Aufforderung‘ wurden fünf Codierungen vergeben, verteilt auf drei Codes.

| <i>Kategorie</i> | <i>Code</i> | <i>N</i> | <i>N</i> |
|----------------------|---|---|----------|
| Funktion / Bewertung | ○ aktualisiert | P ¹⁰ | 1 |
| | ○ große Erfindung | P ⁵ | 1 |
| | ○ modern | P ¹⁰ | 1 |
| | ○ lebendig | P ¹⁰ | 1 |
| | ○ unpraktisch | P ⁵ | 1 |
| Marke | ○ Nokia | P ¹⁸ | 1 |
| Farbgestaltung | ○ gelb | P ¹⁸ | 1 |
| | ○ weiß-blau | P ¹⁸ | 1 |
| | ○ rosa | P ¹⁸ | 1 |
| | ○ blau | P ¹⁸ | 1 |
| | ○ rot | P ¹⁵ , P ¹⁸ | 2 |
| Umgebung / Kontext | ○ Fäden (Leinen) | P ¹⁸ | 1 |
| | ○ Teppich | P ⁹ , P ¹⁴ | 2 |
| | ○ Strickmuster | P ⁶ , P ⁹ , P ¹⁴ | 3 |
| | ○ traditionell engerichtetes türkisches Haus | P ¹⁴ | 1 |
| Emotion / Ausdruck | ○ Gefahr | P ¹⁵ | 1 |
| | ○ Wut | P ¹⁵ | 1 |
| | ○ negative Situation | P ¹⁵ | 1 |
| | ○ Entschlossenheit | P ¹⁵ | 1 |
| Physische Merkmale | ○ Emoji | P ⁶ , P ⁹ , P ¹⁰ , P ¹¹ , P ¹² , P ¹⁵ , P ¹⁶ | 7 |
| | ○ Tastenhandy | P ⁴ , P ⁵ , P ⁷ , P ¹¹ , P ¹⁶ | 5 |
| Retro / Nostalgie | ○ Nostalgie | P ⁵ , P ¹⁰ , P ¹⁷ | 3 |
| | ○ alt | P ³ , P ¹⁰ , P ¹³ , P ¹⁷ | 4 |
| | ○ Entwicklung des Handys | P ⁸ | 1 |
| | | Insgesamt | 43 |

Tab. 5: Vermittelte Eigenschaften und Eindrücke.

Im Rahmen des Analysebereichs ‚Vermittelte Eigenschaften und Eindrücke‘ wurden sieben Kategorien gebildet und insgesamt 43 Codierungen mit 24 unterschiedlichen Codes vorgenommen (Tabelle 5). Sieben Teilnehmende nannten das ‚Emoji‘, fünf

‚Tastendy‘. Vier Teilnehmende gaben ‚alt‘, drei ‚Nostalgie‘ und ebenfalls drei ‚Strickmuster‘ an. Die Begriffe ‚rot‘ und ‚Teppich‘ wurden jeweils von je zwei Teilnehmenden genannt. Die übrigen Codes, darunter ‚aktualisiert‘, ‚modern‘, ‚Nokia‘, diverse Farbangaben sowie einzelne Ausdrücke für Emotionen und Umgebung bzw. Kontext, wurden jeweils einmal genannt.

| | <i>Kategorie</i> | <i>Code</i> | <i>N</i> | <i>N</i> | |
|--|------------------------------|-------------|---|--|---|
| <i>Kohärenz zwischen Bild und Text</i> | Teilweise Passung | ○ | Ausreichend | P ⁴ | 1 |
| | | ○ | sehr relevant | P ⁶ , P ¹⁷ | 2 |
| | | ○ | sehr gut | P ¹² , P ¹⁴ | 2 |
| | Positive Passung | ○ | gut | P ¹ , P ⁵ , P ⁷ , P ⁸ , P ⁹ | 5 |
| | | ○ | widerspiegelt Hauptthema des Textes | P ¹⁰ , P ¹³ , P ¹⁵ , P ¹⁶ | 4 |
| | Unzureichende Passung | ○ | nicht vollständig | P ² , P ³ | 2 |
| | | ○ | fehlt an Inhalt | P ³ | 1 |
| | | ○ | Thema des Textes etwas anders | P ¹³ | 1 |
| | Verständnishilfe / Interesse | ○ | Interesse geweckt | P ¹⁰ , P ¹⁸ | 2 |
| | | ○ | Fokussieren | P ⁸ | 1 |
| | | ○ | Text besser verstanden | P ⁸ , P ¹⁸ | 2 |
| | | ○ | Vorstellung vom Inhalt | P ¹⁰ , P ¹¹ | 2 |
| | | ○ | direkter Zusammenhang | P ¹⁶ | 1 |
| | | ○ | entspricht ehrlich gesagt meinen Gedanken | P ¹⁴ | 1 |
| | Nostalgie / Zeitvergleich | ○ | frühere Zeiten | P ⁴ | 1 |
| | | ○ | alt | P ⁹ | 1 |
| | | ○ | heutige Technologie | P ⁴ | 1 |
| | | ○ | nostalgische Atmosphäre | P ¹⁷ | 1 |
| | | | Insgesamt | 31 | |

Tab. 6: Kohärenz zwischen Bild und Text.

Im Rahmen des Analysebereichs ‚Kohärenz zwischen Bild und Text‘ wurden fünf Kategorien gebildet und insgesamt 31 Codierungen mit 18 unterschiedlichen Codes vorgenommen (Tabelle 6). Fünf Teilnehmende gaben ‚gut‘, vier ‚widerspiegelt Hauptthema‘ an. ‚Sehr gut‘, ‚sehr relevant‘, ‚nicht vollständig‘, ‚Interesse geweckt‘, ‚attraktiver und lesbarer‘ und ‚interessanter gemacht‘ erhielten zwei Nennungen. Alle anderen Codes wurden jeweils einmal vergeben.

| <i>Kategorie</i> | <i>Code</i> | <i>N</i> | <i>N</i> |
|--------------------------------------|---|---|----------|
| begrenzte / keine Hilfe | ○ allgemeine Veranschaulichung des Themas | P ¹⁵ | 1 |
| | ○ Bild allein schwer verständlich | P ¹⁴ | 1 |
| | ○ keinen großen Einfluss | P ³ , P ⁷ , P ⁸ , P ¹¹ , P ¹² | 5 |
| Verständnishilfe | ○ hilfreich | P ¹ , P ⁵ , P ⁹ , P ¹⁷ | 4 |
| | ○ erleichtert | P ⁴ , P ¹³ , P ¹⁴ | 3 |
| | ○ verständlicher | P ⁵ , P ⁹ , P ¹⁰ , P ¹⁴ , P ¹⁷ | 5 |
| | ○ gut | P ³ | 1 |
| Schnelligkeit / Effizienz | ○ leichter konzentrieren | P ⁵ , P ¹⁸ | 2 |
| | ○ Text schneller erfasst | P ¹³ , P ¹⁶ | 2 |
| Visualisierung / mentale Vorstellung | ○ Vorstellung des Themas | P ⁴ , P ¹³ , P ¹⁷ | 3 |
| Vorwissen / Erwartung | ○ Vorwissen zum Thema | P ¹⁰ , P ¹⁸ | 2 |
| | | Insgesamt | 29 |

Tab. 7: Einfluss auf das Textverständnis.

Im Rahmen des Analysebereichs ‚Einfluss auf das Textverständnis‘ wurden fünf Kategorien gebildet und insgesamt 29 Codierungen mit 11 unterschiedlichen Codes vorgenommen (Tabelle 7). Jeweils fünf Teilnehmende nannten ‚keinen großen Einfluss‘ und ‚unverändert‘. Vier gaben ‚hilfreich‘, drei jeweils ‚erleichtert‘ und ‚Vorstellung des Themas‘ an. Alle Codes der Kategorien ‚Schnelligkeit / Effizienz‘ und ‚Vorwissen / Erwartung‘ erschienen zweimal. Alle anderen Codes wurden jeweils einmal vergeben.

| <i>Kategorie</i> | <i>Code</i> | <i>N</i> | <i>N</i> |
|------------------|--|--|----------|
| essenziell | ○ ohne Bild schwieriger | P ¹ , P ⁹ , P ¹⁰ , P ¹³ | 4 |
| | ○ ohne Bild gefehlt | P ⁵ , P ¹³ | 2 |
| | ○ ohne Bild keine Vorstellung zum Thema | P ⁷ , P ¹⁵ | 2 |
| motivierend | ○ interessanter und einprägsamer | P ¹⁸ | 1 |
| | ○ attraktiver und lesbarer | P ¹³ | 1 |
| | ○ ohne Bild langweilig und schwer verständlich | P ¹³ | 1 |
| | ○ Assoziation mit Vergangenheit = Stimmung des Textes besser gefühlt | P ¹⁷ | 1 |
| unklar | ○ nicht sicher | P ¹⁴ | 1 |
| | ○ evtl. andere Gedanken | P ¹² | 1 |
| hilfreich | ○ Verstehen erleichtert | P ³ , P ¹³ , P ¹⁵ , P ¹⁶ | 4 |

| | | | | |
|----------------------------------|---|--|---|----|
| | ○ | Bild unterstützt Text | P ⁵ , P ⁷ | 2 |
| nicht notwendig / verzichtbar | ○ | nicht viel beigetragen | P ⁴ | 1 |
| | ○ | Fehlen des Bildes hätte nichts verändert | P ² , P ⁴ , P ⁶ , P ⁸ , P ¹¹ | 5 |
| | | | Insgesamt | 26 |

Tab. 8: Wahrnehmung der Synergieeffekte.

Im Rahmen des Analysebereichs ‚Wahrnehmung der Synergieeffekte‘ wurden fünf Kategorien gebildet und insgesamt 26 Codierungen mit 13 unterschiedlichen Codes vorgenommen (Tabelle 8). Fünf Teilnehmende nannten ‚Fehlen des Bildes hätte nichts verändert‘. Jeweils vier Teilnehmende gaben die Codes ‚ohne Bild schwieriger‘ und ‚Verstehen erleichtert‘ an. Die Codes ‚ohne Bild gefehlt‘, ‚ohne Bild keine Vorstellung zum Thema‘ und ‚Bild unterstützt Text‘ traten jeweils zweimal auf. Alle anderen Codes wurden jeweils einmal vergeben.

Diskussion, Fazit und Ausblick

Die Befunde bestätigen die zentrale Annahme der Arbeit, dass Bilder in DaF-Materialien nicht lediglich dekorativ wirken, sondern auch eine eigenständige, funktionale Rolle im Bedeutungsaufbau übernehmen. Die zentrale Forschungsfrage dieser Arbeit lautete, wie DaF-Lernende ein isoliertes Bild aus *Deutsch Perfekt* interpretieren. Die Teilnehmenden identifizieren das klassische Nokia-Tastendhandy als Hauptobjekt und ordnen es in einem retro-technischen Kontext ein. Dem Bild wird überwiegend die illokutionäre Funktion der Illustration zugeschrieben. Es werden zudem vielfältige prädikative Merkmale wie Nostalgie, veraltete Technik und Emotion benannt, wodurch dem Bild zeitliche, funktionale und emotionale Bedeutungen zugeschrieben werden. Dieses Ergebnis bestätigt die Annahme einer komplementären Synergie, da sowohl das Bild als auch der Text kohärent auf denselben Referenten verweisen. Das spezifische Bild verleiht dem Text prädikative Mehrwerte, die durch den Text allein möglicherweise nicht in dieser Intensität erzielt werden können.

Auch hinsichtlich der Wirkung der Bild-Text-Kombination zeigen sich signifikante Ergebnisse. Die Hälfte der Teilnehmenden stuft das Bild als Verständnishilfe ein, was zeigt, dass das Bild einen positiven Einfluss auf ihr Textverständnis hatte. Die Antworten der Teilnehmenden belegen, dass Kohärenz auch den kognitiven Verarbeitungsaufwand senkt. Dieser Befund deckt sich mit den Erkenntnissen von Mayer, der zeigt, dass Lernende durch komplementäre Bild-Text-Paare, also durch die gleichzeitige Präsentation von bildlichen und schriftlichen Informationen ein tieferes Verständnis entwickeln, da Informationen über getrennte Kanäle parallel verarbeitet und zu kohärenten mentalen Modellen integriert werden können (Mayer 2002).

Die gegensätzlichen Meinungen ‚essenziell‘ versus ‚verzichtbar‘ im Analysebereich ‚Wahrnehmung der Synergieeffekte‘ zeigen, dass die Wahrnehmung von Bildern von den individuellen Lesestrategien der Teilnehmenden abhängen kann. Lernende, die sich visuell orientieren, können erheblich von dieser Wirkung profitieren, während textzentrierte Personen den Zusatznutzen geringer einschätzen. Damit rückt die Frage nach unterschiedlichen Lerntypen in den Fokus. Wie Grein (2019: 69) aufzeigt, gibt es „den gehirngerechten Zugang nicht, sondern Menschen lernen sehr unterschiedlich.“ Daher sollten insbesondere verschiedene Wahrnehmungskanäle gezielt angesprochen werden. Ein multimodaler Ansatz, bei dem sprachliche, bildliche und auditive Reize kombiniert werden, kann ein breiteres Repertoire an individuellen Bedürfnissen und Interessen ansprechen und somit die Lernwirksamkeit steigern. Grein (2019: 99) fasst dies wie folgt zusammen: „Gehirngerecht ist also weder nur Grammatikübungen noch dauerhafte Aktivität, sondern Berücksichtigung aller Lernstile im Fremdsprachenunterricht.“ Aus den vorliegenden Ergebnissen kann gesagt werden, dass Bild-Text-Synergien insbesondere für visuelle Lerntypen einen direkten Vorteil bieten, während ein multimodaler Unterrichtsansatz insgesamt dazu beitragen kann, möglichst viele Lernende individuell zu fördern.

Die vorliegende Arbeit basiert auf einer kleinen, homogenen Stichprobe von 18 DaF-Lernenden auf B1-Niveau. Ihre Generalisierbarkeit ist folglich begrenzt. In künftigen Arbeiten sollten größere und heterogenere Gruppen einbezogen werden und neben Selbstauskünften auch Leistungsmessungen erfasst werden, um den Einfluss von Bild-Text-Synergien objektiver zu bestimmen. Aus den gewonnenen Erkenntnissen lässt sich ableiten, dass multimodale Lehrmaterialien ihren größten didaktischen Mehrwert entfalten, wenn bildliche Elemente bewusst auf ihre illokutionäre Funktion hin ausgewählt und inhaltlich eng mit dem Text verzahnt werden. Ein Bild, das prädikative Informationen liefert, erhöht nicht nur die Motivation und die Erinnerung an den Inhalt, sondern schafft auch eine kognitive Brücke, die komplexe Inhalte schneller zugänglich macht. An diesem Punkt setzt das Potenzial digitaler Formate an, denn adaptive E-Books oder Lernplattformen können mithilfe von Lernanalytik situativ entscheiden, ob bild- oder textbetonte Erklärungen angezeigt werden, um visuelle oder verbale Lerntypen gleichermaßen zu unterstützen. Empirische Studien belegen, dass adaptive E-Learning-Umgebungen, die gezielt auf individuelle Lernpräferenzen eingehen, die Lernmotivation und das Engagement der Lernenden signifikant steigern können (El-Sabagh 2021: 4). Aufbauend darauf könnten künftige Arbeiten experimentell prüfen, welche Bildtypen für welche Lernstile den größten Mehrwert bieten, und mithilfe größerer Stichproben die Wirksamkeit evaluieren, mit der Text- und Bildvarianten individuell an Vorlieben angepasst werden. Damit lässt sich feststellen, in welchem Kontext und für welche Zielgruppen Bild-Text-Kombinationen im DaF Kontext besonders effektiv sind und wie sie in zukünftigen digitalen Lernumgebungen optimal eingesetzt werden können. Insgesamt zeigt sich, dass die gezielte Synergie von Bild und Text, unterstützt durch digitale Technologien, einen wichtigen Beitrag zur Effektivitätssteigerung und Individualisierung des Fremdsprachenunterrichts leisten kann.

Literaturverzeichnis

- Austin, John L.** (1962): *How to Do Things with Words*. Harvard University Press.
- Beaugrande, Robert Alain / Dressler, Wolfgang Ulrich** (1981): *Introduction to text linguistics* (Vol. 1). London: Longman.
- Brinker, Klaus / Cölfen, Hermann / Pappert, Steffen** (2014): *Linguistische Textanalyse: Eine Einführung in Grundbegriffe und Methoden - Grundlagen der Germanistik*. (8., durchges. Auflage). Berlin: Erich Schmidt Verlag.
- Bußmann, Hadumod** (2008): *Lexikon der Sprachwissenschaft* (4., überarb. Aufl.). Stuttgart: Alfred Kröner Verlag.
- Çelikkaya, Şerife / Karasu, Gönül** (2025): Wissenserwerb durch Bild-Text-Kohärenz im Fremdsprachenunterricht. In: *Diyalog – Interkulturelle Zeitschrift für Germanistik. Sonderausgabe Band I*, 452–478. <https://doi.org/10.37583/diyalog.1824366>.
- Creswell, John W. / Creswell, J. David** (2023): *Research design: Qualitative, quantitative, and mixed methods approaches* (6th ed.). SAGE Publications.
- Dürscheid, Christa** (2007): Schrift-Text-Bild: Ein Brückenschlag. In: *Zeitschrift für germanistische Linguistik*, 35(3), 269-282.
- El-Sabagh, Hassan A.** (2021): Adaptive e-learning environment based on learning styles and its impact on development students' engagement. In: *International Journal of Educational Technology in Higher Education*, 18(1), 53.
- Ernst, Peter** (2002): *Pragmalinguistik: Grundlagen, Anwendungen, Probleme*. Walter de Gruyter.
- Grein, Marion** (2019): Die Methode des Fremdsprachenlernens? – Berücksichtigung von Lernstilen und Einflussfaktoren als Fachkompetenz. In: Ersch, Christina Maria (Hg.): *Kompetenzen in Deutsch als Fremd- und Zweitsprache*. Frank & Timme, 69-101.
- Hosch, Wolfram / Macaire, Dominique** (1991): Landeskunde mit Bildern: Wahrnehmungspsychologische und methodische Fragen bis zur Entwicklung eines Deutschlandbildes durch Bilder. In: *Fremdsprache Deutsch: Zeitschrift für die Praxis des Deutschunterrichts*, 5, 20-27.
- Kjørup, Søren** (1978): Pictorial Speech Acts. *Erkenntnis*, 55-71.
- Kress Gunther / Van Leeuwen, Theo** (2001): *Multimodal Discourse: The Modes and Media of Contemporary Communication*. Arnold.
- Mayer, Richard E.** (2002): Multimedia learning. In: Ross, Brian H. (Hg.): *Psychology of learning and motivation*, 41. Academic Press, 85-139.
- Mills, Geoffrey E. / Gay, Lorraine R.** (2019): *Educational research: Competencies for analysis and applications* (12. Aufl.). Pearson.
- Opilowski, Roman** (2013): Von der Textlinguistik zur Bildlinguistik. Sprache-Bild-Texte im neuen Forschungsparadigma. In: *Zeitschrift des Verbandes Polnischer Germanisten/Czasopismo Stowarzyszenia Germanistów Polskich (ZVPG)*, 2(3), 217-225.
- Searle, John R.** (1969): *Speech Acts: An Essay in The Philosophy of Language*. Cambridge University Press.
- Stöckl, Hartmut** (2004): *Die Sprache im Bild - Das Bild in der Sprache: Zur Verknüpfung von Sprache und Bild im massenmedialen Text Konzepte. Theorien. Analysemethoden*. Walter de Gruyter.
- Storrer, Angelika** (2020): Textqualität digital: Ein Modell zur Qualitätsbewertung digitaler Texte. In: *Deutsche Sprache: ds*, 48(2), 101-125.