

Das Bild, das den Satz schon kennt

Visuelle Szenen in der Unterstützten Kommunikation: was sie können, wem sie helfen — und wo das Raster besser bleibt

Lars Tiedemann

Dipl. Heilpädagoge (FH) · AssistUK

4. August 2026

Arbeitsfassung. Dieser Text erklärt, warum ein Foto vom eigenen Frühstückstisch für manche Menschen die bessere Kommunikationsoberfläche ist als jedes Symbolraster — und warum es trotzdem kein Ersatz für ein wachsendes Vokabular ist. Er ist die fachliche Grundlage hinter unserem szenenbasierten Vokabular SceneTalk. Alle Quellen wurden im August 2026 geprüft; graue Literatur ist als solche gekennzeichnet.

Inhalt

1. Zwei Oberflächen, ein Unterschied
2. Was eine visuelle Szene ist
3. Warum Szenen funktionieren: drei Gründe
4. Was die Forschung zeigt
5. Menschen gehören ins Bild
6. Für wen Szenen der richtige Anfang sind
7. Wo das Raster besser bleibt — und wie der Übergang gelingt
8. Wie wir es bei AssistUK halten
9. Grenzen dieses Textes
10. Fazit in fünf Sätzen
11. Literatur (mit Prüfvermerk)

1. Zwei Oberflächen, ein Unterschied

Stellen Sie zwei Bildschirme nebeneinander. Auf dem ersten: zwanzig ordentliche Kästchen, in jedem ein Symbol — Tasse, trinken, mehr, ich. Auf dem zweiten: ein einziges Foto vom heutigen Frühstückstisch, mit der eigenen Tasse, dem eigenen Brettchen, der Mutter am Tisrand.

Für geübte Nutzerinnen und Nutzer der Unterstützten Kommunikation (UK) ist der erste Bildschirm ein Kraftwerk: Aus zwanzig Bausteinen lassen sich hunderte Sätze bauen. Aber für einen Zweijährigen, für eine Frau mit schwerer Aphasie, für einen Mann mit Demenz kann derselbe Bildschirm eine Wand sein — zwanzig kleine Abstraktionen, jede einzeln zu entschlüsseln, keine erzählt von seinem Leben. Das Foto dagegen muss nicht entschlüsselt werden. Es *ist* die Situation. Wer darauf zeigt, ist schon mitten im Gespräch.

Dieser Unterschied hat einen Namen und eine Forschungsgeschichte. Er heißt **Visual Scene Display** — visuelle Szenenanzeige —, und er ist eine der interessantesten Entwicklungen, die unser Feld in den letzten zwanzig Jahren genommen hat.

2. Was eine visuelle Szene ist

Eine visuelle Szene ist ein Foto oder Bild einer **bedeutungsvollen, zusammenhängenden Situation** — der eigene Garten, der Morgenkreis, das Enkelkind auf dem Schoß —, in das **aktive Bereiche** eingebettet sind: Wer die Tasse im Bild berührt oder anblickt, hört „Ich möchte trinken“; wer auf die Mutter zeigt, hört ihren Namen oder startet eine Erzählung. Die Sprache wohnt *im* Kontext, statt in Kästchen daneben zu liegen.

Der Begriff und die erste Systematik stammen aus den frühen 2000er Jahren; Sarah Blackstone hat sie 2004 in einem vielzitierten Übersichtsartikel gebündelt — Praxisliteratur, kein begutachtetes Journal, aber der Text, mit dem das Konzept in die Breite ging. Die begutachtete Forschung dazu kommt vor allem aus zwei Richtungen: aus der Arbeitsgruppe

um Janice Light und Kathryn Drager an der Penn State University, die Szenen für kleine Kinder untersuchte, und aus der Gruppe um David Beukelman, die sie für Erwachsene mit Aphasie entwickelte. Dass dieselbe Idee an beiden Enden des Lebens unabhängig trägt, ist selbst ein Befund.

3. Warum Szenen funktionieren: drei Gründe

Erstens: Sie senken die Abstraktionsschwelle. Ein Symbolraster verlangt eine doppelte Übersetzung — vom Bedürfnis zum Symbol, vom Symbol zur Position im Raster. Eine vertraute Szene verlangt nur Wiedererkennen. Wiedererkennen ist entwicklungsgeschichtlich früher, robuster und bleibt auch dann erhalten, wenn andere Fähigkeiten brüchig werden.

Zweitens: Sie nutzen das episodische Gedächtnis. Das ist der Kern der Aphasie-Linie der Forschung: Menschen mit schwerer Aphasie verlieren den Zugriff auf das Sprachsystem, aber häufig nicht die Erinnerung an ihr eigenes Leben. Ein kontextreiches Foto der eigenen Hochzeit, des eigenen Hundes, der eigenen Werkstatt öffnet diesen erhaltenen Speicher — und trägt damit Gespräche, die über eine Bedürfnistafel nie zustande kämen (Beukelman, Hux, Dietz, McKelvey & Weissling 2015).

Drittens: Sie verteilen die Last des Gesprächs. Eine gute Szene ist nicht nur Auswahlfläche, sondern gemeinsamer Gesprächsgegenstand: Beide Partner schauen auf dasselbe Bild, beide können zeigen, kommentieren, fragen. Das entlastet die unterstützte kommunizierende Person von der Rolle, jede Äußerung allein aus dem Nichts zu bauen — das Gespräch wird zur Ko-Konstruktion am Bild.

4. Was die Forschung zeigt

Die am häufigsten zitierten Befunde stammen aus zwei Studienreihen mit sehr jungen, sich typisch entwickelnden Kindern — und man muss sie genau lesen, weil sie oft ungenau wiedergegeben werden.

Bei **Zweieinhalbjährigen** (Drager et al. 2003) schnitten Szenen-Layouts im Vergleich relativ am besten ab — aber die ehrliche Hauptbotschaft der Studie ist eine andere: *Alle* dynamischen Anzeigen waren für dieses Alter schwer; die Kinder fanden Vokabular in jeder Bedingung mühsam. Wer aus dieser Studie „Szenen lösen das Problem“ macht, überdehnt sie.

Bei **Dreijährigen** (Drager et al. 2004) wird das Bild klarer: Hier lernten die Kinder das Vokabular im kontextuellen Szenen-Format von der ersten Sitzung an **signifikant besser als im Raster**. Zusammen mit der Überblicksarbeit von Light und Drager (2007) ergibt sich die vorsichtige, belastbare Lesart: **Je jünger das Kind und je früher der Entwicklungsstand, desto stärker spricht die Evidenz für den Einstieg über**

Szenen — nicht, weil Raster schlecht wären, sondern weil sie Voraussetzungen haben, die ein Anfänger noch nicht mitbringt.

Für Erwachsene mit chronischer, schwerer **Aphasie** fasst die Arbeitsgruppe um Beukelman (2015) die Forschung so zusammen: Kontextreiche, persönlich relevante Fotos unterstützen Konversation dort, wo dekontextualisierte Symbole scheitern. Und die aktuelle Standortbestimmung zur Display-Gestaltung insgesamt (Light, Wilkinson, Thiessen, Beukelman & Fager 2019) behandelt Szenen als etablierten Baustein des Werkzeugkastens — mit offenen Forschungsfragen, aber außer Frage stehendem Nutzen für die richtigen Zielgruppen.

5. Menschen gehören ins Bild

Ein Gestaltungsbefund verdient einen eigenen Abschnitt, weil er so schön konkret ist. Krista Wilkinson und Janice Light haben mit Blickmessung untersucht, wohin Menschen schauen, wenn sie Fotos betrachten: **auf die Menschen im Bild** — schneller und länger als auf praktisch alles andere, selbst wenn die Person klein oder am Rand abgebildet ist (2011). Und das gilt nicht nur für typisch Entwickelte: Auch Betrachter mit Autismus, Down-Syndrom oder anderen kognitiven Beeinträchtigungen fixierten die abgebildeten Personen rasch — im Mittel innerhalb von etwa anderthalb Sekunden — und ausdauernd (2014).

Für die Praxis heißt das: **Eine gute Kommunikationsszene ist selten ein Stilleben.** Das Foto vom leeren Frühstückstisch ist schwächer als das Foto, auf dem die Schwester gerade nach der Marmelade greift. Menschen im Bild ziehen die Aufmerksamkeit dorthin, wo Kommunikation entsteht — zu Handlungen, Beziehungen, Gesichtern. Wer Szenen fotografiert, sollte Personen mitfotografieren (und die sozialen Wörter an ihnen verankern), statt sie aus falscher Aufgeräumtheit hinauszuschneiden.

Die zweite Gestaltungsregel folgt aus der Aphasie-Forschung: **persönlich schlägt generisch.** Das eigene Wohnzimmer trägt mehr als das Katalogbild eines Wohnzimmers. Und die dritte kommt aus unserem CVI-Text: Bei zerebraler Sehbeeinträchtigung brauchen Szenen reduzierte Hintergründe und ein einzelnes, deutlich hervorgehobenes Ziel — das Foto des ganzen, vertrauten Gegenstands, nicht ein Suchbild.

6. Für wen Szenen der richtige Anfang sind

Aus Evidenz und Praxis ergeben sich vier Gruppen, bei denen wir Szenen als Einstieg empfehlen:

Sehr junge Kinder und frühe Kommunizierer. Wer gerade erst entdeckt, dass Zeigen etwas bewirkt, braucht keine Grammatikmaschine, sondern einen vertrauten Ort zum Zeigen. Ein Hotspot im Lieblingsfoto — und die erste selbstausgelöste Äußerung ist da.

Menschen mit Aphasie. Das erhaltene episodische Gedächtnis ist ihr Kapital; Szenen verzinsen es. Besonders in der frühen Zeit nach dem Ereignis, wenn das Symbolsystem nicht trägt, halten Szenen das Gespräch offen.

Menschen mit Demenz. Vertraute Orte und Gesichter bleiben lange zugänglich, wenn Neues es nicht mehr ist. Szenen fragen nichts Neues ab.

Menschen mit zerebraler Sehbeeinträchtigung in der mittleren Phase. Fotos realer, vertrauter Objekte sind hier ohnehin das Mittel der Wahl — die Szene liefert sie mit Kontext, sofern der Hintergrund beherrscht wird.

Quer dazu gilt: Szenen sind auch für geübte Kommunizierer ein legitimes **Zusatzwerkzeug** — für das Erzählen („Schau, das war unser Ausflug“), für Übergabesituationen, für alles, wo der Kontext die halbe Botschaft ist.

7. Wo das Raster besser bleibt — und wie der Übergang gelingt

Jetzt die Grenze, und sie ist grundsätzlich: **Eine Szene kann nur sagen, was in ihr vorkommt.** Wer „nochmal“, „nicht das“, „warum“ sagen will — die kleinen Wörter, aus denen Sprache besteht —, findet sie in keinem Foto. Generative Sprache, das Bauen nie gehörter Sätze aus wiederverwendbaren Bausteinen, ist die Stärke des Rasters mit Kernvokabular und die strukturelle Schwäche der Szene. Deshalb ist unsere Formel: **Die Szene kommuniziert, bevor das generative System sitzt — sie ersetzt es nicht.**

Der Übergang ist kein Bruch, sondern ein Anbau. Bewährt hat sich die Hybridform: An der Szene wächst eine kleine, konstante Leiste mit Kernwörtern — „mehr“, „fertig“, „anders“, „Frage“. Das Kind zeigt erst nur ins Bild; dann kombiniert es Bild und Leiste („mehr“ + Schaukel im Foto); dann trägt die Leiste allein immer größere Teile. Die Szene bleibt als Erzähl- und Kontextraum erhalten, während das Raster die Satzarbeit übernimmt. So verliert niemand den vertrauten Boden, und niemand bleibt auf ihm stehen.

8. Wie wir es bei AssistUK halten

SceneTalk, unser szenenbasiertes Vokabular, ist genau nach diesen Linien gebaut: eigene Fotos statt Katalogbilder, frei setzbare aktive Bereiche, verknüpfbare Szenen für zusammenhängende Situationen — und bewusst als Ergänzung im Gefüge positioniert, nicht als Endpunkt. In unserer Empfehlungslogik steht SceneTalk am Anfang (frühe Stufen, ein Hotspot im vertrauten Foto) und behält später die Rolle des Erzählraums neben dem wachsenden Rastervokabular. Die Gestaltungsregeln aus Abschnitt 5 — Menschen ins Bild, persönlich vor generisch, bei Sehbeeinträchtigung reduziert und hervorgehoben — geben wir jeder Beratung mit; ein eigener Modus, der Hintergründe für die CVI-Zielgruppe automatisch beruhigt, steht auf unserer Entwicklungsliste und ist noch nicht im Produkt.

Und weil dieser Text die Grundlage unseres eigenen Produkts ist, sage ich es ausdrücklich: Die Grenze aus Abschnitt 7 gilt auch für uns. Wo eine Beratung bei der Szene stehen bleibt,

obwohl die Person mehr könnte, ist das derselbe Fehler wie das verfrühte Raster — nur in die andere Richtung.

9. Grenzen dieses Textes

Die Kinder-Studien der Penn-State-Gruppe wurden mit sich **typisch entwickelnden** Kindern durchgeführt; die Übertragung auf Kinder mit Beeinträchtigungen ist plausibel und praxisbewährt, aber nicht im selben Maß belegt. Die Aphasie-Evidenz beruht überwiegend auf kleinen Stichproben und Einzelfallserien — die Richtung ist konsistent, große kontrollierte Studien fehlen.

Blackstones Grundlagentext ist Praxisliteratur ohne Peer-Review; ich zitiere ihn als das, was er ist. Und zur Hybridform aus Abschnitt 7 — Szene plus wachsende Kernwortleiste — kenne ich keine Studie, die genau diesen Übergang isoliert prüft; das ist begründete Praxis, keine gemessene Evidenz.

Schließlich: Dieser Text behandelt die Oberfläche, nicht die Partnerseite. Ob eine Szene Gespräche trägt, hängt auch hier an den Menschen, die mitschauen, warten und modellieren — dafür steht der Gesprächsverhaltens-Text neben diesem.

10. Fazit in fünf Sätzen

Eine visuelle Szene ist ein vertrautes Foto, in dem die Sprache dort wohnt, wo sie hingehört: im Kontext. Sie funktioniert, weil Wiedererkennen früher und robuster ist als Entschlüsseln, weil sie das episodische Gedächtnis nutzt und weil beide Gesprächspartner auf dasselbe Bild schauen. Die Evidenz trägt den Einstieg über Szenen besonders für sehr junge Kinder und für Menschen mit Aphasie oder Demenz — und die Blickforschung gibt die schönste Gestaltungsregel gratis dazu: Menschen gehören ins Bild. Aber eine Szene kann nur sagen, was in ihr vorkommt; die kleinen Wörter der Sprache stehen in keinem Foto. Deshalb gilt: mit der Szene anfangen, am Raster wachsen — und beides behalten.

11. Literatur (mit Prüfvermerk)

Alle Quellen am 04.08.2026 auf Existenz und Kernaussage geprüft (✓).

1. ✓ **Drager, K., Light, J., Speltz, J., Fallon, K. & Jeffries, L. (2003).** The performance of typically developing 2½-year-olds on dynamic display AAC technologies with different system layouts and language organizations. *Journal of Speech, Language, and Hearing Research*, 46(2), 298–312. — Szenen-Layout relativ am besten, aber alle Bedingungen für dieses Alter schwer.
2. ✓ **Drager, K. et al. (2004).** Learning of dynamic display AAC technologies by typically developing 3-year-olds: Effect of different layouts and menu approaches. *Journal of Speech, Language, and Hearing Research*, 47(5), 1133–1148. — Dreijährige lernen im kontextuellen Szenen-Format signifikant besser als im Raster.

3. ✓ **Light, J. & Drager, K. (2007).** AAC technologies for young children with complex communication needs: State of the science and future research directions. *Augmentative and Alternative Communication*, 23(3), 204–216. — Überblick; Neugestaltung von Oberflächen für junge Kinder.
4. ✓ **Wilkinson, K. & Light, J. (2011).** Preliminary investigation of visual attention to human figures in photographs. *Journal of Speech, Language, and Hearing Research*, 54(6), 1644–1657. — Menschen im Bild ziehen den Blick schneller und länger an als andere Elemente.
5. ✓ **Wilkinson, K. & Light, J. (2014).** Preliminary study of gaze toward humans in photographs by individuals with autism, Down syndrome, or other intellectual disabilities. *Augmentative and Alternative Communication*, 30(2), 130–146. — Replikation bei Menschen mit Beeinträchtigungen; Fixation binnen ~1,5 s.
6. ✓ **Wilkinson, K., Light, J. & Drager, K. (2012).** Considerations for the composition of visual scene displays. *Augmentative and Alternative Communication*, 28(3), 137–147. — Gestaltungsempfehlungen für Szenen.
7. ✓ **Dietz, A., McKelvey, M. & Beukelman, D. (2006).** Visual Scene Displays: New AAC interfaces for persons with aphasia. *Perspectives on Augmentative and Alternative Communication*, 15(1), 13–17. — Ursprungstext der Aphasie-Linie.
8. ✓ **Beukelman, D., Hux, K., Dietz, A., McKelvey, M. & Weissling, K. (2015).** Using visual scene displays as communication support options for people with chronic, severe aphasia. *Augmentative and Alternative Communication*, 31(3), 234–245. — Forschungszusammenfassung; episodisches Gedächtnis als Träger.
9. ✓ **Light, J., Wilkinson, K., Thiessen, A., Beukelman, D. & Fager, S. (2019).** Designing effective AAC displays for individuals with developmental or acquired disabilities: State of the science. *Augmentative and Alternative Communication*, 35(1), 42–55. — Aktuelle Standortbestimmung zur Display-Gestaltung.
10. ✓ **Blackstone, S. (2004).** Visual scene displays. *Augmentative Communication News*, 16(2), 1–8. — Frühe Systematik; Praxisliteratur ohne Peer-Review, als solche zitiert.
11. ✓ **Tiedemann, L. (2026).** Sehen ist nicht gleich Erkennen · Die zweihundert Wörter, die fast alles sagen · Vom Zugriff zum Vokabular. AssistUK-Grundlagentexte. — CVI-Gestaltung, Kernvokabular und Stufenlogik, an die dieser Text anschließt.

Die fachliche Position, Auswahl und Verantwortung liegen bei mir. — L. T.