

Hinweis der Herausgeber*innen: Dieser Beitrag ist Teil des *Quo-vadis*-Themenhefts und wurde daher keinem Peer-Review-Verfahren unterzogen.

Canvas Reloaded

Zur Analyse multimodaler Strukturen in zeitgemäßen Social-Media-Kurzformaten (Artikel)

Stina Viher & Janina Wildfeuer

Abstract

Die zunehmende Komplexität medialer Umgebungen, etwa durch Technologizität, Programmierbarkeit oder Interface-Design, erfordert eine systematische, begriffliche und methodische Ausdifferenzierung, die in bisherigen Arbeiten oft noch nicht ausreichend granular behandelt wird. Dies hat zur Folge, dass die Vielfalt an Zeichenmodalitäten und plattformspezifischen Kompositionsebenen mit ungleicher Sorgfalt behandelt wird. Der Beitrag schlägt ein materialitätsorientiertes Analysemodell vor: einen erweiterten Canvas-Ansatz (Wildfeuer et al. 2020; Bateman et al. 2026), der die Verschränkung von spezifischen Plattformaaffordanzen und multimodaler Bedeutungskonstitution digitaler Kurzformate (Memes, TikTok-Videos, etc.) hierarchisch operationalisierbar macht. Mithilfe von zwei ausführlichen Analysen zeigen wir, wie sich die Granularität plattformgebundener Kompositionslogiken mithilfe der vorgeschlagenen Herangehensweise rekonstruieren lässt. Diese Vorgehensweise sehen wir als grundlegenden Ausgangspunkt für Analysen multimodaler Kommunikationsformate, auf dem andere, methodisch-spezifische Schritte aufbauen sollten.

Keywords: Multimodalität, Social-Media-Kurzformate, Materialität, Canvas-Ansatz

Kontaktperson:

Stina Viher
Ludwig-Maximilians-Universität München
Schellingstraße 3
80799 München
stina.viher@germanistik.uni-muenchen.de

1 Einleitung: Post-digitale Kommunikationsformate als analytische Herausforderung

Ein Blick auf die gegenwärtige mediale Landschaft zeigt, dass Internetnutzung und soziale Medien längst zu einem selbstverständlichen Bestandteil unseres Alltags geworden sind. Diese Ubiquität digitaler Kommunikationsräume, die im Call zu dieser Sonderausgabe des *Journal für Medienlinguistik* gar als Provokation einer eigenen Epochenbezeichnung, der der Post-Digitalität, verstanden wird, eröffnet spezifische mediale Handlungsmöglichkeiten, die von Nutzer:innen angeeignet und verwendet werden. In diesem Zusammenhang lassen sich neuartige kommunikative Praktiken beobachten, die sich in wiederkehrenden Formaten und verfestigten Mustern manifestieren und als Ausdruck zeitgemäßer Medienkommunikation gelten können.

Insbesondere Kurzformate¹ wie TikToks (Fischer et al. 2025), Reels (Luginbühl 2025), Internet-Memes (Bülow/Merten 2023; Viher 2021), GIFs (Logi/Zappavigna 2024) und Instagram-Postings (Luginbühl 2025; Merten 2024; Viher im Erscheinen) sind vor dem Hintergrund eines Überangebots an Inhalten und verkürzter Aufmerksamkeitsspannen (vgl. Nam/Jung 2021) allgegenwärtig. Charakteristisch für diese Formate sind ihre hohe semiotische Dichte, ihre Prägnanz sowie eine häufig ästhetisch aufbereitete Gestaltung. Aus linguistischer Perspektive zeichnen sie sich insbesondere durch komplexe multimodale Strukturen aus, in denen sprachliche, visuelle und auditive Modalitäten eng miteinander verschränkt sind (z. B. Stöckl 2020; Grzenkowicz/Wildfeuer 2025).

Sprache tritt in diesen Formaten kaum mehr isoliert auf und ist nahezu immer eingebettet in technisch komplexe, multimodale Kompositionen. Insbesondere die spezifischen Plattformaffordanzen sorgen immer öfter dafür, dass Sprache nur eine von vielen Ressourcen ist, während Videotechniken, Musik, GIFs und andere Animationen in den Vordergrund rücken. Der Prozess der Bedeutungskonstitution vollzieht sich folglich immer durch das Zusammenspiel mehrerer unterschiedlicher Zeichenmodalitäten, die jeweils einen eigenständigen Beitrag zur kommunikativen Gesamtbedeutung leisten, aber niemals in Isolation gesehen werden sollten. Kurzformate (vgl. Georgakapoulou 2023, Guerra 2026) machen damit neuartige Prozesse der Bedeutungsentfaltung sichtbar und stellen zugleich

¹ Wir verstehen diese Kurzformate auch als *Snack-Content* (Viher im Erscheinen), also als ästhetisch aufbereitete, prägnante und dennoch informationsreiche Inhalte. Die vermeintliche Einfachheit dieser Formate weist bei näherer Betrachtung gleichermaßen eine hohe semiotische Komplexität und semantische Verdichtung auf. Snack-Content ist demnach weniger als „leichte Kost“ denn als dicht strukturiertes multimodales Kommunikationsformat zu begreifen.

klassische textlinguistische Annahmen von Kohärenz, Kohäsion und Linearität infrage.

Für linguistische Untersuchungen sind sie insofern von besonderem Interesse, als sich in ihnen sowohl neue Textformen manifestieren als auch grundlegende Eigenschaften medialer Kommunikation paradigmatisch bündeln. Aus methodischer Sicht bringt diese multimodale Komplexität jedoch große Herausforderungen mit sich. Sie erfordert eine holistische Perspektive, die sprachliche und nicht-sprachliche Ressourcen als koordiniertes multimodales Handeln begreift. Wir beobachten, dass linguistische Zugänge gegenüber dieser Auffassung immer noch eine gewisse Skepsis, in jedem Fall aber Zurückhaltung zeigen, Sprache als eine von vielen Ausdrucksformen zu behandeln und vor allem anderen Zeichenmodalitäten eine gleichwertige Funktion zuzuschreiben. Sprache spielt immer die „zentrale Rolle“ (Schmitz 2022: 25), die „tragende Rolle“ (Merten et al. 2024: 9), Sprachlichkeit wird als „zwingende Voraussetzung für die Existenz eines Textes“ (Fix 2024: 17) gesehen. Und weiter noch: einer ganzheitlichen multimodal-analytischen Herangehensweise, die die Vorrangstellung von Sprache in Frage stellt, werden „methodische[n] Schwierigkeiten“, „Scheuklappen gegenüber Zusammenhängen“ und ein „modische[r] Denkstil [...], der allzusehr auf viele Spielarten von Multimodalität fixiert ist“ (Schmitz 2022: 25), vorgeworfen.

Dieser Skepsis und den Vorwürfen wollen wir in diesem Beitrag begegnen, indem wir uns unterschiedlichen digitalen Kommunikationskontexten nähern und systematisch beleuchten, wie Bedeutungen in diesen Kontexten konstituiert wird und welche Rolle die unterschiedlichen Zeichenmodalitäten darin spielen. Dabei konzentrieren wir uns insbesondere auf die *Materialität* und damit einhergehende Granularität der sozialen Medien Instagram und TikTok. Wir setzen Impulse für eine methodische Herangehensweise, die es erlaubt, multimodale Kommunikation in digitalen Kontexten in Form einer hierarchisch organisierten Bedeutungsarchitektur zu analysieren, die die Frage danach, welche Rolle Sprache genau spielt, zu beantworten vermag. Aufbauend auf der Canvas-Terminologie (Wildfeuer et al. 2020; Bateman et al. 2026) schlagen wir ein analytisches Vorgehen vor, das die Dimensionen medialer Materialität systematisch den Dimensionen der Zeichenmodalität zur Seite stellt. Während wir damit also durchaus die vielen Spielarten von Multimodalität in den Blick nehmen, wollen wir zugleich methodisch präzise argumentieren, wie eine linguistische Analyse den „schier unüberschaubaren Herausforderungen“ (Schmitz 2022:14) sicher gegenübertreten kann.

2 Dimensionen medialer Materialität statt Sprachfokus

Während Klug und Stöckl (2016: VIII) Sprache noch als Archimedium bzw. auch als Archimodalität bezeichnen, thematisierten Stöckl et al. (2020) im Sammelband *Shifts toward Image-Centricity in Contemporary Multimodal Practices* einen klaren Wandel zu einer bildzentrierten Medienkommunikation. Die Sicht auf Sprache als primärer Bedeutungsträger verliert an Plausibilität und nimmt statt einer übergeordneten Rolle eine gleichwertige ein (Wildfeuer/Bateman 2018; Wildfeuer et al. 2020; Bateman/Tseng 2023, vgl. auch Hess-Lüttich et al. 2017). Obwohl sich in den vergangenen Jahren ein deutlich wachsendes Interesse an Diskursen und kommunikativen Handlungsmustern in digitalen Kontexten beobachten lässt, wovon einschlägige Sammelbände und Studien zeugen (Androutsopoulos/Vogel 2024; Gredel et al. 2022; Kabatnik et al. 2024; Marx et al. 2020; Meier-Vieracker et al. 2023), ist in den meisten dieser Arbeiten doch immer auch eine ungleiche Gewichtung der berücksichtigten Zeichenmodalitäten erkennbar. Während sprachliche Elemente in den Mittelpunkt der Analyse rücken oder häufig mit hoher analytischer Präzision untersucht werden, etwa durch die Identifikation von N-Grammen als wiederkehrende Wortkombinationen (Bülow/Merten 2023; Merten 2024), erfolgt die Einbeziehung weiterer Modalitäten oftmals nachgeordnet und lediglich additiv. Primärer Fokus bleibt überwiegend die Sprache. Werden visuelle Elemente als Analysegegenstände ausgewählt, ist vielfach unklar, nach welchen Kriterien dies geschieht. Hinweise auf ein Close-Reading-Verfahren deuten auf eine interpretative Annäherung hin, die zwar als „feingranulare Kategorisierung“ (Bülow/Merten 2023: 141) ausgewiesen wird, deren methodische Operationalisierung jedoch nicht hinreichend transparent gemacht wird.

Auch besteht ein gewisses Risiko für methodische Verzerrungen, wenn z. B. Klassifikationsschemata auf bestimmten plattformspezifischen Formaten basieren und lediglich „loosely based on the classification“ (Luginbühl 2025: 51) auf ein anderes Format adaptiert werden. Auch die Berücksichtigung bzw. die Verschränkung von Zeichenmodalitäten auf der gleichen oder unterschiedlichen Ebenen (z. B. bei sequentiellen Postings) bleibt häufig unterbestimmt: Zwar werden intermodale Verknüpfungen untersucht und zum Beispiel die Sequenzialität von Karusell-Posts thematisiert (Merten 2024), doch bleibt zugleich fraglich, durch welche Mittel und welche Analyseebenen explizit die Bezugnahme und damit die Bedeutungskonstitution zwischen den Slides hergestellt wird.

Ähnliche Fragen stellen sich bei bestehenden Analysen von komplexeren kommunikativen Situationen oder Medienformaten wie TikTok-Videos (Tschannen/Meier-Vieracker 2024), Tanzgesten

(Kindler-Mathôt 2025) oder Choreographien von Fankommunikation im Stadion (Meier-Vieracker/Hauser 2025). All diese Arbeiten nehmen eine erste Klassifizierung unterschiedlicher Ebenen der „multimodalen Bedeutungsgestalt“ (Kindler-Mathôt 2025: 189) vor und beschreiben „charakteristische Formen der Zeichenprozessierung“ (Meier-Vieracker/Hauser 2025: 2017), das methodische Vorgehen für Transkription und Analyse konzentriert sich dann aber vor allem auf sprachorientierte Ausdrucksformen, bei denen in den meisten Fällen formatspezifische Aspekte (wie z. B. des Videoformats oder der räumlichen Bewegung und Anordnung) außen vor bleiben.

Bei näherer Betrachtung zeigt sich also, dass viele auf Sprache fokussierte Arbeiten – trotz eines häufig proklamierten multimodalen Anspruchs – weiterhin viele nicht-sprachliche Zeichenmodalitäten mit weniger Systematik, Präzision und Transparenz behandeln. Dies stellt insofern ein Problem dar, als kommunikative Handlungen in digitalen Kontexten konstitutiv und inhärent multimodal organisiert sind. Diese grundlegende Eigenschaft der Multimodalität sieht aber zugleich nicht vor, dass spezifische Zeichenmodalitäten dominanter gegenüber anderen Zeichenmodalitäten operieren und deswegen prominenter behandelt werden sollten. Stattdessen hat die Multimodalitätsforschung bewusst den Fokus auf die Gesamtheit und Verbundenheit multimodaler Ensembles gesetzt (Kress 2010), in denen Bedeutungen nicht einzelnen Zeichenmodalitäten zugewiesen werden, sondern sich in ihrem Zusammenspiel multiplizieren (Lemke 1998) und kontextabhängig bestimmt werden müssen. Innerhalb eines multimodalen Ensembles haben Zeichenmodalitäten also einen ‚vorbehaltlichen Status‘ (vgl. Kress 2010: 165), der abhängig ist von den spezifischen Bedingungen der jeweiligen Situation. Wenn Bedeutung also z. B. in sozialen Medien im Zusammenspiel unterschiedlicher Einheiten gesehen wird, muss sich notwendigerweise auch die Sichtweise auf Sprache und sprachliche Muster ändern, ebenso wie sich der Gebrauch dieser in der Praxis ändert. Sprachliche Formen sind dann wie alle anderen Zeichenmodalitäten nicht isoliert zu betrachten, sondern stehen in formalen, semantischen und funktionalen Zusammenhang mit anderen Zeichenmodalitäten (vgl. auch Stöckl 2006: 25).

Im Hinblick darauf kommt eine zusätzliche Problematik multimodaler Analyse zutage, auf die auch Merten et al. (2024) hinweisen und die den weiterhin bestehenden Mangel an quantitativen Studien betrifft. Die Gründe für diese Forschungslücke sind, wie die Autor:innen betonen, offensichtlich:

Sowohl die Zusammenstellung multimodaler und zur Untersuchung aufbereiteter (umfangreicher) Korpora als auch die

nötige Mehrebenen-Annotation sowie Mehrebenen-Analyse sind in zeitlicher wie auch personeller Hinsicht herausfordernde Aufgaben (Merten et al. 2024: 20 f.).

Über den Aspekt des hohen Arbeitsaufwands hinaus verweist diese Problematik auch auf ein methodologisches Defizit: Trotz bereits etablierter Herangehensweisen im angloamerikanischen Raum (vgl. u. a. O'Halloran et al. 2021, Bateman/Hiippala 2021) scheint immer noch die Frage im Raum zu stehen, wie man multimodale Inhalte für die Forschung zugänglich macht. Während die Korpuslinguistik Werkzeuge bereitstellt (z. B. Beißwenger et al. 2025), um sprachliche Daten zu transkribieren, dokumentieren und auszuwerten, scheint der Einbezug anderer Zeichenmodalitäten und vor allem die gleichwertige Behandlung aller Zeichenmodalitäten immer noch eine methodische Herausforderung darzustellen. Mit Blick auf die hier in Rede stehenden Formate bleibt weiterhin die Frage offen, wie die große und teilweise unübersichtliche Menge an Ausdrucksformen in post-digitalen Kommunikationsformaten systematisch beschrieben, annotiert und analysiert werden kann.

Um dieser Komplexitätsproblematik zu begegnen, rückt schon seit einiger Zeit im Kontext interdisziplinärer und damit nicht unbedingt linguistisch-fokussierter Multimodalitätsforschung die Materialität dieser Formate in den Mittelpunkt der Analysen (vgl. Bateman et al. 2017; Meiler 2019; Bateman et al. 2026). Ein ähnliches, wenn auch nicht so starkes Interesse ist in einigen text-, diskurs- und konversationsanalytischen und damit stärker linguistisch geprägten Arbeiten zu erkennen (vgl. z. B. Fix 2008, Spitzmüller 2018, Bender et al. 2022, Mondada 2019). In der multimodalen Semiotik wird Materialität nicht als bloßer technischer Träger verstanden, sondern als konstitutive Eigenschaft der Bedeutungsproduktion: Sie beeinflusst maßgeblich, wie Inhalte wahrgenommen, verarbeitet und interpretiert werden (vgl. Wildfeuer et al. 2020; Wildfeuer 2024; siehe auch Kap. 3.2). Diese Wirkung lässt sich auf materielle Spuren – visuelle, auditive oder greifbare – zurückführen, die in spezifischen medialen Umgebungen realisiert werden. So operiert etwa TikTok primär mit dynamisch-visuellen und auditiven Spuren, wodurch bestimmte kommunikative Praktiken begünstigt und andere strukturell eingeschränkt werden (vgl. Grzenkowicz/Wildfeuer 2026).

Technisch-mediale sowie sozio-kommunikative Merkmale der jeweiligen Plattform rahmen, strukturieren und bringen die kommunikativen Praktiken in spezifischer Weise hervor (vgl. Bender et al. 2022). Hierbei zeigt sich insbesondere, dass Sprache nur eine von vielen Zeichenmodalitäten der kommunikativen Botschaft ist. Dies wird auch mit unseren exemplarischen Analysen im Kapitel 4 sichtbar. Insbesondere aufgrund der technisch-medialen Möglichkeiten

sind in diesen kommunikativen Situationen oft mehrere Materialitäten mit unterschiedlichen Spuren simultan integriert. Es entstehen komplexe, mehrschichtige Konstellationen auf verschiedenen Ebenen: Bei Video-Formaten zeigen sich diese beispielsweise als Textüberlagerungen, Musik oder Split-Screens mit unterschiedlichen Perspektiven. Statische Formate hingegen vereinen in einer Bilderstrecke GIFs, Screenshots oder Layout-Kompositionen etablierter Artefakten (Poster, Cover). Um herauszufinden, wie Zeichenmodalitäten auf unterschiedlichen Oberflächen und im Zusammenspiel mehrerer Oberflächen für kommunikative Zwecke eingesetzt werden, ist es erforderlich, die beteiligten Materialitäten differenziert zu beschreiben und in ihrer Verschränkung zu analysieren. Eine analytische Perspektive, die dieser strukturellen Verschränkung gerecht werden will, muss daher systematisch die Materialitäts- und Kompositionskomplexität von Medienformaten erfassen und sie materialitätsorientiert operationalisieren.

3 Canvas Reloaded: Multimodale Analyse mit Fokus auf Materialität

Eine methodische Herangehensweise, die materialitätsorientiert operationalisiert und dem Anspruch genügt, multimodale Kommunikation in digitalen Kontexten umfassend und detailliert analysieren zu können, beginnt unserer Ansicht nach zunächst mit einer grundsätzlichen semiotischen Fundierung aller Kommunikationsphänomene. Diese erfasst die Dimensionen medialer Materialität, um darauf aufbauend die unterschiedlichen Zeichenmodalitäten zu identifizieren. So wird eine Grundlage für weitere methodisch-spezifische Untersuchungsansätze geschaffen, die auch großangelegte Datenanalysen möglich machen.

3.1 Multimodalität als zentraler Ausgangspunkt kommunikativer Situationen

Unser Ansatz basiert auf einem in früheren Arbeiten vorgestellten theoretischen und methodischen Rahmenwerk zur multimodalen Analyse, der für alle Kommunikationsformate geeignet ist und im Folgenden auf digitale Kontexte zugespitzt werden soll (Wildfeuer et al. 2020; Bateman et al. 2017, 2026). Er versteht multimodale Phänomene grundlegend als (eingebettet in) kommunikative Situationen, die Kombinationen verschiedener Ausdrucksformen wirkungsvoll einsetzen, um Bedeutung zu konstruieren (Wildfeuer et al. 2017: 7). Jede kommunikative Situation wird somit als inhärent multimodal betrachtet, alle in diesen Situationen vorkommenden Ausdrucksformen werden als grundsätzlich relevant angesehen und müssen in der

Analyse berücksichtigt werden. Für Kurzformate wie TikTok-Videos oder Instagram-Postings bedeutet dies, dass nicht nur ihre multimodale Beschaffenheit, sondern auch ihre Einbettung in die Plattform und ihr Gebrauch und ihre Verbreitung durch Zeichennutzer:innen eine Rolle in der Analyse spielen.

Ein multimodal-analytischer Untersuchungsansatz digitaler Kommunikationsformate beginnt also nicht bei oder fokussiert auf einzelne Zeichenmodalitäten innerhalb dieser Formen (wie gesprochene oder geschriebene Sprache), sondern stellt die Tatsache, dass eine kommunikative Situation vorliegt, an den Beginn der Auseinandersetzung. So können nicht nur Medien und Texte als multimodale Artefakte gefasst werden, sondern auch viele andere Aspekte und Herausforderungen in der multimodalen Kommunikation beleuchtet werden, z. B. digitales Lesen als Prozess der Zeichenbildung und -interpretation (vgl. Wildfeuer 2025) oder der Umgang mit KI (Thomson et al. 2025). Ein so zugeschnittener Ansatz versteht sich dann auch nicht als Erweiterung anderer (v. a. linguistischer) methodischer Zugänge, sondern eher als Fundierung oder Unterfütterung solcher Zugänge bzw. als systematische Grundlage, bei der Multimodalität am Beginn und im Zentrum der Analyse steht, während Sprache lediglich ein Bestandteil von vielen in diesen 'multimodalen Gebrauchskontexten' (Schmitz 2022) ist.

Ein Blick auf die kommunikative Situation als Ganzes beginnt dort, wo sie in spezifische soziokulturelle Kontexte eingebettet sind und soziale Praktiken einer Kultur oder Gemeinschaft zeigen. Ein zentraler Schritt in jeder Analyse ist daher die Identifikation dieser sozialen Praktiken, kontextuellen Einflüsse und deren Auswirkungen auf die kommunikative Situation. Dies gelingt am besten durch die Betrachtung der jeweiligen Textgenres und Medien, in denen der Diskurs eingebettet ist. Ein Blick auf das Textgenre bietet Einblicke in sozial konventionalisierte Muster, die in den schematischen Strukturen und der Auswahl von Inhalt und Stil einer bestimmten Klasse kommunikativer Ereignisse sichtbar werden (Swales 1990; Bateman et al. 2026: 76). Ein Medium verstehen wir dagegen als „historisch situierte Praxis“, die eine spezifische Menge an Zeichenmodalitäten bereitstellt und verarbeitet (Wildfeuer et al. 2020: 152). Die Analyse des beteiligten Mediums gibt Aufschluss darüber, wie Kommunikation in einem bestimmten institutionalisierten Kontext tatsächlich stattfindet.

Entgegen der viel verbreiteten Annahme, dass ein solcher Ansatz „unüberschaubaren Herausforderungen“ gegenübersteht und viel zeitlichen und personellen Einsatz verlangt (Schmitz 2022, Merten et al. 2024), versteht er sich als richtungsweisender Versuch, Ordnung in die Komplexität und Fülle multimodaler Kommunikation zu bringen. Methodisch geschieht dies mit einem klaren Fokus auf den

materiellen Eigenschaften einer kommunikativen Situation und damit auf ihre interne Organisation.

3.2 Materialität und Canvas

Hierfür setzt unser Ansatz bei der materiellen Grundlage einer kommunikativen Situation an und beleuchtet genauer, wie Bedeutung aus den unterschiedlichen Zeichenmodalitäten konstruiert wird. Diese explizite Auseinandersetzung mit den grundsätzlich verschiedenen Formen von Materialität, die in komplexen multimodalen Situationen zum Tragen kommen, ermöglicht es, sinnvolle Entscheidungen darüber zu treffen, was in eine bestimmte multimodale Analyse einbezogen und was möglicherweise ausgeschlossen werden kann oder soll.

Um diese Formen von Materialität zu fassen, nutzen wir den Terminus Canvas (vgl. Wildfeuer 2020: Kap. 3). Ein Canvas wird verstanden als der Ort, die Umgebung oder auch die Fläche, auf der materielle Regelmäßigkeiten eingeschrieben sind, die dann wahrgenommen und interpretiert werden können (vgl. Wildfeuer et al. 2020: 103-104). Diese Orte und Flächen sind nicht nur real und tatsächlich greifbar, sondern auch virtuell oder digital, also mit technologischer Unterstützung geschaffen, wie zum Beispiel in den sozialen Medien. Der grundlegende Canvas für ein Foto, das zwei Menschen sich auf einem Smartphone-Bildschirm anschauen, ist das Interface auf diesem Bildschirm, das auf dem Mobilgerät zur Verfügung steht. Je nach Beschaffenheit des Artefaktes sind die Eigenschaften dieses Canvases aber unterschiedlich: für ein mit dem Smartphone aufgenommenes Foto ist er statisch, für ein damit ebenso aufgenommenes Video aber dynamisch. So verhält es sich auch auf den unterschiedlichen Plattformen der sozialen Medien; sie können sowohl statische als auch dynamische Postings darstellen und Nutzer:innen als Beobachtende oder als aktiv Teilnehmende der kommunikativen Situation behandeln. Welche materiellen und technologischen Eigenschaften hierfür jeweils zum Tragen kommen, muss Teil einer multimodalen Analyse sein, für die die Bestimmung des Canvas ein erster Ansatzpunkt ist.

Die Beschreibung dieses Canvas bzw. der unterschiedlichen Canvases und Subcavases innerhalb komplexer Situationen ermöglicht es, diese Situationen in verschiedene materielle Ebenen oder Oberflächen zu dekonstruieren und sie entlang mehrerer Kriterien zu klassifizieren, die die beteiligten Zeichenmodalitäten zugänglich machen. Wildfeuer et al. (2020: Kap. 3) schlagen hierfür zunächst Kriterien vor, die sich auf die Repräsentation eines Canvas in Raum und Zeit (dynamisch oder statisch, zwei- oder dreidimensional), seine Transienz bzw. Vergänglichkeit (dauerhaft oder flüchtig bzw.

transient) und die Rolle der Teilnehmenden im und auf dem Canvas (Beobachtende oder Teilnehmende bzw. aktiv Gestaltende) beziehen. Abbildung 1 gibt einen Überblick über die vier grundlegenden Eigenschaften eines Canvases.

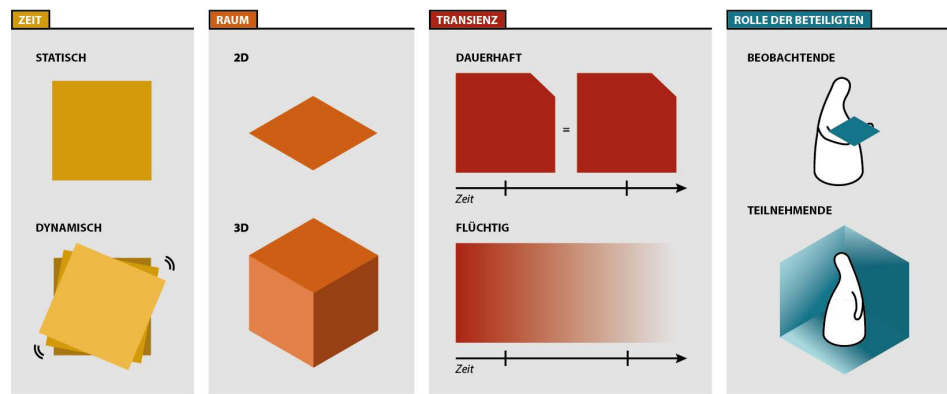


Abbildung 1: Die vier grundlegenden Dimensionen von Materialität, die für Canvases angenommen werden (vgl. Bateman et al. 2017, Wildfeuer et al. 2020, Bateman et al. 2026; Grafik: Jana Pflaeging)

Während also eine klassische Werbeanzeige auf einer Litfaßsäule an einer Bushaltestelle ein statischer, zweidimensionaler und dauerhafter Canvas ist, bei dem Rezipient:innen als Beobachter:innen agieren, ist ein kurzes Video auf TikTok hingegen ein dynamisches, flüchtiges Artefakt, das von Nutzer:innen geteilt oder weiterverarbeitet werden kann. TikTok-Videos kombinieren oft mehrere Canvases, um eine Vielzahl von Zeichenmodalitäten zu integrieren, etwa Screenshots, bewegte Bilder, Texteinblendungen, Animationen, Voice-over und Musik. Ähnlich verhalten sich die Canvases auch bei dem sozialen Medium Instagram, die außerdem durch unterschiedliche, plattforminhärente Formatmöglichkeiten bedingt sind.

3.3 Vom Canvas zur Bedeutungsarchitektur und -hierarchie

Aufbauend auf der initialen Charakterisierung der unterschiedlichen Canvases und mit dem Ziel, auch das komplexe Zusammenspiel dieser Materialitäten zu dekonstruieren, verwenden wir in einem weiteren Schritt hierarchische Darstellungen, wie sie in Abb. 3 und Abb. 5 zu sehen sind, die die Beziehungen der einzelnen Canvases und Subcanvases zueinander organisieren und die komplexe Bedeutungsarchitektur eines multimodalen Artefaktes visualisieren. Auch diese Visualisierungsmethode geht auf jüngste Vorarbeiten zurück (vgl. Bateman et al. 2026: 119, 200), die sie vor allem im Hinblick auf die Vielfalt kommunikativer Situationen testen.

Eine solche Darstellung zeigt die materiellen Dimensionen auf, auf denen die Canvases eines Medienformats beruhen, und macht

zugleich deutlich, wie die unterschiedlichen Canvases und Subcanvases miteinander verzweigt sind. An der Spitze einer visualisierten Hierarchie stellen wir den größten und umfassendsten Canvas der Gesamtsituation dar und fügen dann in den darunter gelegenen Ebenen weitere Subcanvases hinzu. Die materiellen Eigenschaften jedes Canvases oder Subcanvases werden mithilfe von Kurzlabels für jede der in Abschnitt 3.2 beschriebenen Charakteristika aufgelistet: Zeitlichkeit wird als [S]tatisch oder [D]ynamisch beschrieben, Raum wird als Dimensionalität erfasst (0D, 1D, 2D oder 3D), Transienz wird als [D]auerhaft oder [F]lüchtig beschrieben und die Rolle der Beteiligten als [B]eobachtende oder [T]eilnehmende aufgeführt (vgl. Bateman et al. 2017; Wildfeuer et al. 2020; Bateman et al. 2026).

Was den ersten Anwendungen in Bateman et al. (2026) fehlt und wir hier exemplarisch hinzufügen und testen, ist die direkte Verbindung der Canvasanalyse mit der Beschreibung der Zeichenmodalitäten. Jeder einzelne materielle Ausschnitt, der in der Hierarchie identifiziert wird, trägt verschiedene Zeichenmodalitäten, die in unseren Visualisierungen auf einer zusätzlichen Beschreibungsebene, eingebettet in den jeweiligen Canvas, gelistet werden. So können wir einerseits aufzeigen, dass ein einzelner Canvas eine Vielzahl an Zeichenmodalitäten unterstützen kann, und andererseits deutlich machen, dass eine detaillierte Unterscheidung der Modalitäten notwendig ist, um allen Zeichenmodalitäten gerecht zu werden. Denn nicht jeder Canvas unterstützt jede Zeichenmodalität, was wir aufgrund der Charakterisierungskriterien für Canvases erläutern können. Eine visueller statischer Canvas kann bereits eine große Bandbreite an Ausdrucksformen bereitstellen (geschriebene Sprache, Diagramme, Layout, Fotos oder Zeichnungen, etc.), ist aber nicht in der Lage, dynamische Formen wie Bewegtbilder, Animationen oder Musik einzubetten. Das bedeutet, dass z. B. Videos und Musik, die innerhalb sozialer Medien neben statischen Bildern erscheinen, zwangsläufig auf einem anderen Canvas operieren.

Wie zuvor argumentiert und wie wir in unseren Analysen exemplarisch zeigen werden, treten Zeichenmodalitäten in multimodalen Gesamtkontexten grundsätzlich gemeinsam auf einer materiellen Oberfläche auf. Zu betonen ist, dass sie für ihre Identifikation aber nicht in noch weitere, separate Beschreibungsebenen oder materielle Schichten unterteilt werden müssen. Um ihr Zusammenspiel zu untersuchen, zum Beispiel indem Bedeutungsrelationen zwischen einzelnen Modalitäten (zwischen Sprache und Bild oder Bild und Musik) identifiziert werden, ist nicht allein die materielle Ebene ausschlaggebend, sondern müssen auch die formale und kontextuelle bzw. diskurssemantische Ebene von Zeichenmodalitäten hinzugezogen werden (vgl. Wildfeuer/Bateman 2018, Wildfeuer et al. 2020:

Kap. 4), um diese diskursiven Verbindungen untereinander aufzeigen zu können. Unsere eigene Analyse in den folgenden Abschnitten wird nicht so weit gehen und ihren Fokus darauf legen, die unterschiedlichen materiellen Ebenen zweier komplexer Kommunikationsformate ganzheitlich zu betrachten und systematisch zu segmentieren. In einem weiteren Schritt können dann die komplexen Bedeutungsprozesse je nach Untersuchungsinteresse und Forschungsmethode untersucht werden.

4 Rekonstruktion plattformgebundener Kompositionslogiken im Rahmen des Canvas-Ansatzes

Die vorgeschlagene Herangehensweise soll nun an zwei ausgewählten Fallbeispielen aus den sozialen Medien (Instagram-Post und TikTok-Video) exemplarisch veranschaulicht werden. Die Entscheidung für diese Plattformen begründet sich darin, dass sie zum einen zu den beliebtesten Plattformen sozialer Medien weltweit zählen (vgl. Statista 2025), die sowohl im privaten als auch institutionellen Kontext genutzt werden, und darin, dass sie zum anderen sehr komplexe multimodale Strukturen abbilden, die wir zu beleuchten versuchen.

Beide Kommunikationsformate konstituieren eine sequenzielle Bedeutungsentfaltung durch plattformgebundene Logiken. Kohärenz und Kohäsion gehen nicht ausschließlich aus den verwendeten Zeichenmodalitäten hervor, sondern werden durch diese plattformgebundenen Logiken – etwa Swipe-Sequenzen im Karussell-Post oder die zeitliche Organisation audiovisueller Zeichenmodalitäten im TikTok-Video – gestaltet. Den ausgewählten Beiträgen ist gemeinsam, dass sie in einem professionellen Kontext mit sozio-ökonomischer Zielsetzung entstanden sind. Sie werden daher als persuasive Artefakte verstanden, die strategisch auf spezifische kommunikative Intentionen ausgerichtet sind. Ihre Inhalte werden also gezielt produziert und veröffentlicht, um Sichtbarkeit, Positionierung und Reputation zu generieren (vgl. Shepherd 2005). Bei der folgenden Auseinandersetzung mit den Fallbeispielen leitet uns allerdings vor allem das Ziel, die materiellen Unterschiede herauszuarbeiten und den Gebrauch von Zeichenmodalitäten in den Postings aufzuzeigen; eine weitere Analyse persuasiver Strategien oder Positionierungsmuster werden wir nicht vornehmen.

4.1 Instagram

Die Plattform Instagram ist in erster Linie primär als mobile App konzipiert. Deren zentrale Funktionen werden somit v. a. über die

Smartphone-App bereitgestellt und laufend weiterentwickelt (vgl. Instagram 2026). Als „undeniably image-centric“ App (Caple 2020: 153) sind Nutzer:innen dazu aufgefordert, „to share images and videos (using Internet-enabled smart devices) on a particular topic or event“ (Caple 2020: 153). Während lange Zeit das statische Bild im Zentrum stand, lässt sich zunehmend der Einsatz des Bewegtbildes beobachten, indem „reels, stories, and video postings take centre stage“ (Bateman et al. 2026: 116). Für die vorliegende Analyse ist diese grundlegende Charakterisierung von Bedeutung, weil sie auf die materiellen und funktionalen Bedingungen der Produktion und Rezeption von Beiträgen verweist, die die Komposition von Instagram-Beiträgen maßgeblich strukturieren. Im Gegensatz zu sogenannten Single-Posts, die an spezifische materielle Ausprägungen gebunden sind (z. B. statisches Bild oder Bewegtbild), erlauben andere Formate wie IG-Stories, Reels und Karussell-Posts die Integration unterschiedlicher materieller Beschaffenheiten innerhalb eines einzelnen Beitrags.

Diese komplexen, durch die Plattform vorstrukturierten Anordnungen und die gegebenen Möglichkeiten innerhalb der jeweiligen Formate machen eine feingliedrige, mehrschichtige Analyse notwendig, die mit dem von uns vorgestellten Canvas-Ansatz vorgenommen werden kann. Charakteristisch für IG-Posts sind beispielsweise „einfache Bilder“ (Merten 2024: 226) im Sinne von Fotografien, „Collagen, die sich aus mehreren Bildern zusammensetzen“ (Merten 2024: 226) oder stark grafisch aufbereitete Inhalte (Viher im Erscheinen). Im Großen und Ganzen variiert die Aufbereitung von Inhalten je nach Ausrichtung des Accounts und dem Trend der Zeit. Neueste Untersuchungen im Marketing-Bereich zeigen, dass viele Content-Creator:innen auf sprachliche Elemente im Bildbereich oder Reels (Text-Overlays) setzen und damit bei den Nutzer:innen für mehr Interaktion mit dem Instagram-Post sorgen. Neben Text-Overlays stellt die App darüber hinaus eine Vielzahl von Werkzeugen bereit, die bei der Vermittlung von Inhalten zum Einsatz kommen können. Dazu zählen u. a. Filter, Layout-Vorlagen sowie auf der Clip-Ebene zusätzlich Funktionen wie Greenscreen, KI-Bildanimationen und unterschiedliche Arten von Sticker (z. B. Avatare, GIFs, Countdown-Sticker, Sticker für Spendenaktionen usw.). Wenn Nutzer:innen diese Techniken anwenden, entstehen anspruchsvolle Analyseobjekte, die verschiedenste Zeichenmodalitäten auf unterschiedlichen medialen Materialitäten und Ebenen miteinander verschränken. Dies wollen wir mithilfe einer detaillierten Analyse eines Instagram-Posts und insbesondere seines Bildbereichs aufzeigen.

4.1.1 Der Instagram-Post „NFTs auf Instagram“ und seine Canvases

Unser erstes Fallbeispiel (siehe die Screenshots in Abbildung 2) zeigt einen typisch aufbereiteten Inhalt des IG-Accounts *@babygotbusiness*.² Das Social-Media-Profil wird von Ann-Katrin Schmitz gemeinsam mit ihrem Team betrieben. Als Expertin für digitale Markenkommunikation und Content-Marketing richtet sie sich primär an Selbstständige, Content-Creator:innen und Unternehmen aus der Creator Economy. Neben der Vermittlung praxisorientierter Inhalte zur Social-Media-Nutzung bietet der Account zugleich metakommunikative Einblicke in die Mechanismen erfolgreicher Social-Media-Strategien, um u. a. die Leistungen des Unternehmens zu bewerben. Der am 26.05.2022 veröffentlichte Beitrag mit dem Thema NFTs im Kontext von Instagram lässt sich primär als Wissensvermittlung und zugleich als Bewerbung und Vermarktung der eigenen Leistung (auch Markenpositionierung) verstehen.

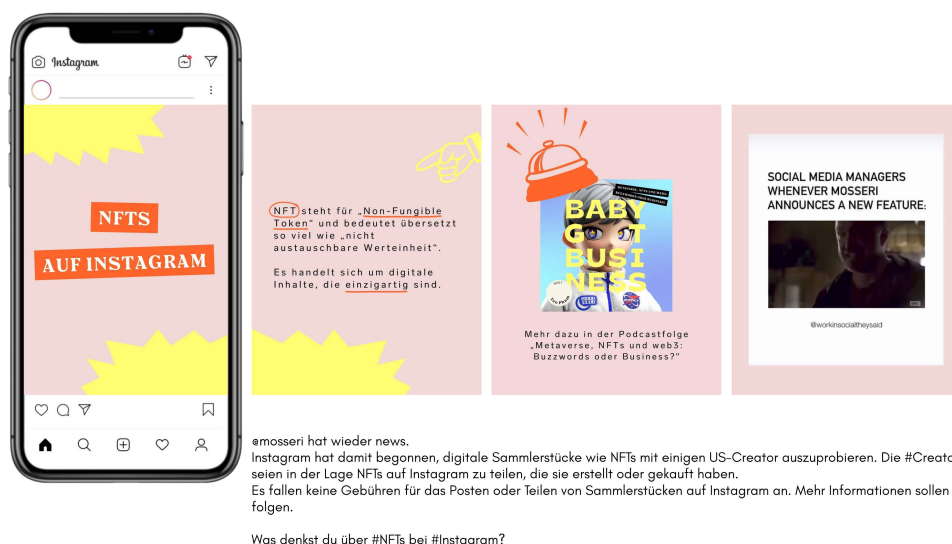


Abbildung 2: Screenshots aus dem Instagram-Post der Unternehmerin und Influencerin Ann-Katrin Schmitz (*@babygotbusiness*, 26.05.2022, https://www.instagram.com/p/CeA-VWYKoCF/?img_index=1), in eigener Darstellung zusammengefügt.

² Das Beispiel stammt aus einem Datensatz mit 657 IG-Postings des Instagram-Accounts *babygotbusiness* (vgl. Viher 2026). Die Beiträge wurden entweder als Einzelpostings oder als Karussell-Postings in Form von statischen Bildern und Bewegtbildern veröffentlicht. Die Datenerhebung fand im Mai 2024 mithilfe eines Python-Scripts statt.

Nehmen wir als Analysierende eine Rezipientenperspektive³ ein und betrachten das App-Interface holistisch als einen kommunikativen Raum mit unterschiedlichen Möglichkeiten, lassen sich grundsätzlich mehrere Canvases identifizieren (vgl. Bateman et al. 2026: Kap. 5). In unserer Analyse konzentrieren wir uns auf das Interface, das uns einen spezifischen Post anzeigt, nachdem wir diesen über den dazugehörigen Account angesteuert haben. Dieser Karussell-Post wird auch beim Öffnen der App und dem anschließenden Durchscrollen der Timeline angezeigt, allerdings sieht der unmittelbare Kontext etwas anders aus. Die komplexe Vielschichtigkeit dieser kommunikativen Situation stellen wir in Abbildung 3 dar, in welcher wir die Gesamthierarchie aller Canvases visualisieren. Diese Hierarchie setzt sich zunächst aus einem Account-Interface und einem Post-Interface zusammen. Die beiden Subcanvases sind zweidimensional, statisch und dauerhaft, weil die Nutzenden sie nicht modifizieren können und somit eine beobachtende Rolle einnehmen. Das im Fokus stehende Post-Interface erweist sich bei der näheren Betrachtung wiederum als mehrschichtig: es umfasst mehrere Handlungsspielräume mit spezifischen technisch-medialen Merkmalen und spezifischen kommunikativen Möglichkeiten. So lassen sich weitere untergeordnete (Sub-)Canvases identifizieren, die je nach Zielsetzung als einzelne Analysegegenstände und/oder in Bezug zueinander betrachtet werden können: Es gibt einen Bildbereich, der entweder statisch oder dynamische Inhalte unterstützt, einen Interaktions-Canvas (mit dem gelikt, kommentiert, gepostet, geteilt oder gespeichert werden kann), eine Caption sowie einen Kommentar-Canvas. Diese Subcanvases sind funktional differenziert und positionieren Akteur:innen mal in die Rolle von Beobachtenden (im Falle des Bildbereichs und der Caption), mal in die Rolle von Teilnehmenden bzw. aktiv Gestaltenden (auf dem Interaktions- und Kommentar-Canvas). Bei einem Karussell-Post sind Sound- bzw. Musik-Interface selten, aber möglich. Die Musikspur läuft dann parallel zum Bildbereich, d.h. über alle Slides hinweg und bestimmt somit die Bedeutungskonstitution mit.

4.1.2 Analysefokus: Bildbereich

Den Canvas des Bildbereichs schauen wir uns im Folgenden genauer an. Die kommunikative Botschaft des Bildbereichs ist im Sinne einer Bilderreihe über mehrere Slides und damit auf unterschiedlichen Canvases verteilt, die von Nutzer:innen durch horizontales Wischen nach links erschlossen werden: Beginnend mit der thematischen Einleitung (Slide 1 und 2) wird auf eine weiterführende Podcastfolge

³ Zur Unterscheidung der Perspektiven siehe Bateman et al. (2017: 362–364).

(Slide 3) des Formats *baby got business* verwiesen, die sich mit NFTs im Kontext von Web 3 und Metaverse auseinandersetzt. Das letzte Bild des Posts (Slide 4) greift das stetige Verkünden von neuen Funktionen der App auf und bewertet diese in Form von Video-Memes. Die einzelnen Slides innerhalb des Bildbereiches übernehmen somit bestimmte kommunikative Funktionen, die wiederum aufgrund von formalen und inhaltlichen Merkmalen der Informationsvermittlung eigenständige Medienartefakte (Podcast-Cover, Internet-Meme) erkennen lassen, die deswegen nach Wildfeuer et al. (2020) als 'dargestellte Medien' bezeichnet werden sollten.

Der Canvas Bildbereich ist grundsätzlich durch eine Layoutstruktur geprägt, die unterschiedliche Darstellungsmuster ermöglicht. In den ersten beiden Slides werden die zentralen Informationen (Titel und begriffliche Erklärung) in sprachlicher Form platziert. Die Fläche wird durch randständig platzierte grafische Elemente (Zickzackwolke) ergänzt, die als saliente Akzente fungieren. Während beim ersten Bild die schriftlichen Elemente „NFTS AUF INSTAGRAM“ zusätzlich durch eine Farbfläche hervorgehoben werden, erfolgt dies beim zweiten Bild durch grafische Markierungen wie Unterstreichungen („Non-Fungible Token“, „einzigartig“) und Einkreisungen („NFT“). Im Zentrum des dritten Bildes befindet sich das Podcast-Cover einer Folge, die thematisch zum Instagram-Post passt. Dieses eigenständige Medienartefakt enthält wiederum eine spezifische Struktur: Das Bild in der Mitte („3D-Figur“), das einen Podcast-Gast abbildet, wird überlagert von dem gelben Logo „baby got business“. In der linken unteren Ecke wird der Name des Podcast-Gastes mit einem kreisförmigen Hintergrund hervorgehoben. Der untergeordnete Bildbereich wird zugleich durch schriftliche Elemente im Sinne einer impliziten Aufforderung „Mehr dazu in der Podcast-Folge [...]“ kontextualisiert. Auf der gleichen Ebene wird mit dem „Alarm-Symbol“ als grafischem Overlay eine weitere Deutungseinheit geschaffen. Das Internet-Meme im letzten Bild bildet eine klar abgegrenzte Fläche, auf der sich der weiße Rahmen klar von der orangenen Farbe des Bildes abhebt. Während der Meme-Canvas selbst zweidimensional statisch ist, ist der Video-Subcanvas dynamisch. Der semantische Inhalt dieser beliebten Formate entsteht durch die Verzahnung von sprachlichen Elementen, dem Bewegtbild und dem O-Ton. Das Artefakt als solches zu erkennen, erfordert im Zuge des Rezeptionsprozesses auch ein spezifisches Wissen, das im Forschungsdiskurs als *meme literacy* (z. B. Lankshear/Knobel 2019) verstanden wird.

Damit das Analysematerial bestmöglich für eine quantitativ-orientierte Auswertung systematisch aufbereitet wird, lohnt es sich, die einzelnen Slides als Analyseebenen zu bestimmen und dabei die of-

fensichtlichen, räumlich abgegrenzten Kompositionssegmente herauszuarbeiten. Ähnlich wie in Untersuchungen zu Postern (vgl. z. B. Zhang/Bateman 2026: 232–243) kann dieser Canvas als seitenbasiert bzw. „page based“ (vgl. Bateman 2008) betrachtet werden. Die einzelnen Bilder bündeln Regelmäßigkeiten, organisieren sie im Layoutraum und tragen zugleich zu einer übergeordneten Layoutstruktur bei. Das Layout wird innerhalb der Multimodalitätsforschung als eigenständige Zeichenmodalität oft übersehen, weil es die grundsätzliche Zusammenstellung eines visuellen, zweidimensionalen Canvases betrifft und dementsprechend oft auch als Designstrategie bezeichnet wird (vgl. dazu aber Bateman 2008). Als Zeichenmodalität ermöglicht es aber die systematische Anordnung und Organisation weiterer Zeichenmodalitäten wie geschriebene Sprache und grafische Elemente und ist somit eine entscheidende Analyseeinheit. Anders als klassische Poster, die meist auf einer materiellen Grundlage beruhen, erlaubt die Plattform Instagram die Integration unterschiedlicher Materialitäten innerhalb eines kommunikativen Artefakts.

Zur systematischen Rekonstruktion des Bildbereichs ist es zunächst sinnvoll, den Inhalt in drei Kategorien bzw. Segmente zu unterteilen: schriftliche Einheiten (mit und ohne Farbfläche), grafische Einheiten (Linien, Umkreisungen, Symbole) und eigenständige Medienartefakte (hier etwa Podcast-Cover und Internet-Meme). Wie in Abbildung 3 ersichtlich, sind diese Segmente auf der gleichen hierarchischen Ebene, jedoch in unterschiedliche Analyseeinheiten verteilt. Diese Unterteilung bringt eine grundlegende Charakterisierung der Inhalte hervor und ermöglicht das schnelle Abrufen von Informationen darüber, wie diese Segmente auf unterschiedliche Ausdrucksmittel zurückgreifen und sie organisieren (vgl. Bateman 2026: 157). Im Sinne einer quantitativen Analyse ist es jedenfalls lohnend, diese Analyse als Grundlage für detaillierte Annotationen zu nutzen, so dass verfestigte Kombinationen mehrerer Ausdrucksformen über mehrere Instagram-Postings vergleichbar und generalisierbar gemacht werden können.

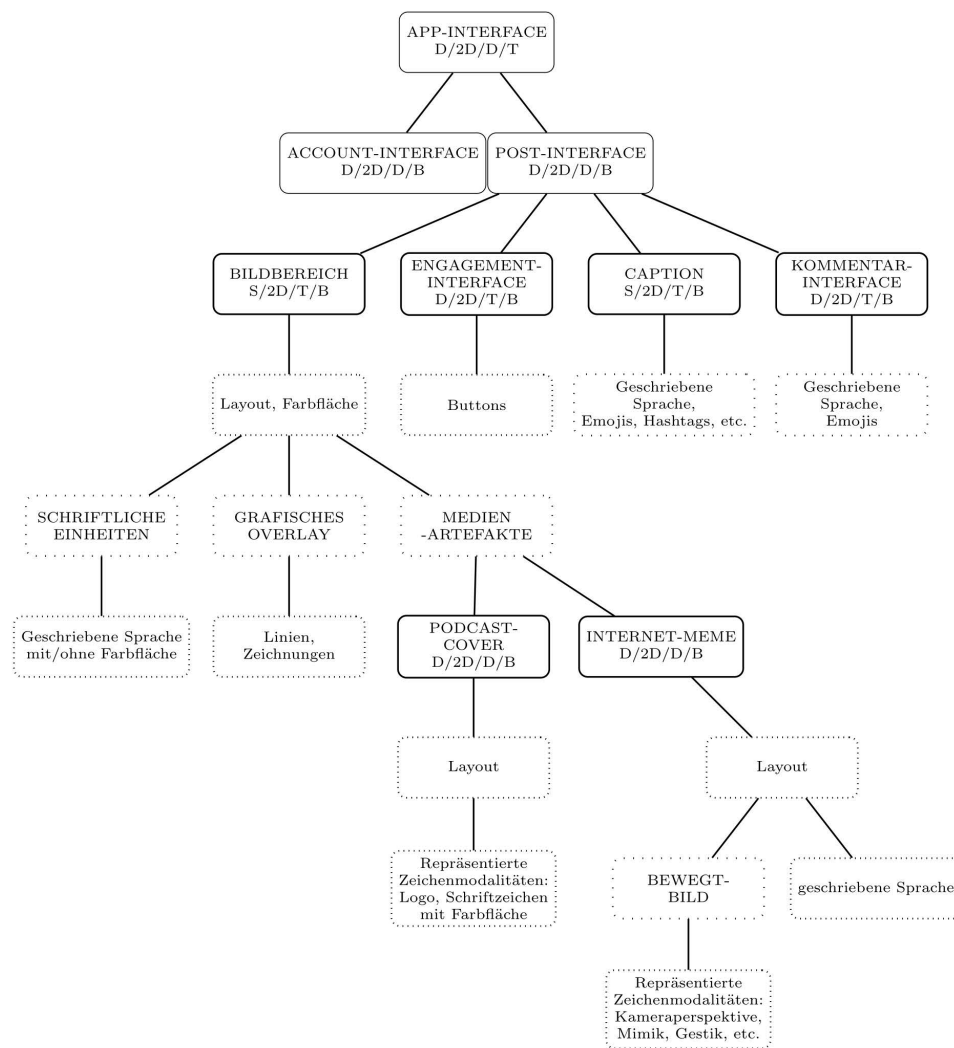


Abbildung 3: Hierarchische Organisation von Canvas und Subcanvases des Karussell-Posts von @babygotbussines

Die Canvas-Analyse schafft für die Untersuchung dieser verfestigten Praktiken einen wichtigen methodischen Schritt, indem der Bildbereich in seiner komplexen Zeichenprozessierung in Segmente gegliedert und differenziert beschrieben wird. Es kann zunächst festgehalten werden, dass Sprache hier eine tragende Rolle für die thematische Rahmung und Informationsvermittlung einnimmt. Zugleich wird jedoch deutlich, dass eine rein sprachzentrierte Betrachtung zentrale Aspekte der kommunikativen Praxis ausblenden würde und die funktionale Logik (die werbende und unterhaltende Intention), die durch die Verschränkung von unterschiedlichen Zeichenmodalitäten realisiert wird, unzureichend erfasst.

4.2 TikTok

Im Gegensatz zu Instagram ist das Interface-Design der Plattform TikTok fast vollständig video- bzw. bewegtbildzentriert. Zwar lassen sich auch hier abweichende Formattypen finden (z. B. auch Posts, die statische Abbildungen im Slide-Format nebeneinander zeigen), allerdings bleiben dynamische, audiovisuelle Artefakte das bevorzugte Format und werden Nutzer:innen vornehmlich dazu angehalten, Videos zu erstellen bzw. hochzuladen. Eine besondere Rolle spielt hier auch die Ton- bzw. Musikebene, für die oftmals externe Quellen verwendet werden, die aber in enger Abstimmung mit der visuellen Ebene eingesetzt werden, zum Beispiel bei Tanzbewegungen oder wenn akustische Signale die Schnittabfolge bestimmen (vgl. auch Tschannen/Meier-Vieracker 2024: 289).

Insgesamt spielt die Plattform eine wichtige Rolle in der heutigen videozentrierten Medienkultur und sticht vor allem durch ihren multimodal-komplexen, partizipativen Charakter in der Masse an sozialen Medien heraus. Mit spezifischen Videoformaten wie ‚Duets‘ oder ‚Stitches‘ sind Nutzer:innen in der Lage, direkt auf die Inhalte anderer Nutzer:innen zu reagieren und mit ihnen zu interagieren, indem sie neue Videos erstellen, z. B. durch Split-Screen-Formate oder mithilfe von Überlagerungen durch Greenscreen-Technik. Audiovisuelle Inhalte werden also grundsätzlich vielfältig gestaltet und durch viele zusätzliche Ebenen wie Voice-Overs, Textüberlagerungen und visuelle Filter weiter ausdifferenziert, so dass „ausgesprochen komplexe, multimodale Gesamttexte“ (Tschannen/Meier-Vieracker 2024: 289, vgl. auch Grzenkiewicz/Wildfeuer 2025) entstehen. Dieser Komplexität zu begegnen und dabei alle Ausdrucksformen in den Blick zu nehmen ist eine Herausforderung, weil sich TikTok-Videos zumeist auch durch ein sehr hohes Tempo und eine hohe Informationsdichte auszeichnen (vgl. Tschannen/Meier-Vieracker 2024: 291).

Folglich müssen die in der Analyse zu behandelnden Analyseeinheiten auf einer sehr detaillierten Ebene beschrieben werden, wobei nicht nur der Unterschied zwischen statischen und bewegten Bildern erfasst werden sollte, sondern auch verschiedene Kameraperspektiven, der Einsatz von zusätzlichen Objekten wie GIFs, Animationen, Emojis, Bewegungen und Übergänge, die Wahl der Musik und Töne sowie der Gebrauch von gesprochener und geschriebener Sprache – wie es typischerweise in multimodalen Filmanalyseansätzen und audiovisuellen Transkriptionen geschieht (siehe z. B. Baldry und Thibault 2006; Wildfeuer 2014, Bateman et al. 2017: Kap. 13).

Auch für qualitative Analysen einzelner Videos ist es erforderlich, die formalen Aspekte der unterschiedlichen Zeichenmodalitäten, ihre materielle Ebene, an den Beginn der Analyse zu stellen. Wie

Tschannen/Meier-Vieracker (2024: 295) bereits initial demonstrieren, ist hierfür die Canvas-Methode als Beschreibung der hierarchischen Bedeutungsarchitektur hilfreich. Im Folgenden wollen wir anhand eines weiteren Beispiels die Feinheiten dieser Methode weiter beleuchten.

4.2.1 Die Grünen-Video: Make it make sense

Bei unserem Beispiel handelt es sich um ein TikTok-Video der Partei Die Grünen (<https://www.tiktok.com/@diegruenen>), das am 19. Januar 2026 auf der Plattform unter dem Titel „Understanding the Billionaire Tax Debate“⁴ gepostet wurde. Das Video ist sieben Sekunden lang und mit einer Caption verbunden, die lediglich sagt: „Make it make sense“. Das Video lässt sich als Kommentar zur Debatte über die Milliardensteuer in Deutschland verstehen, die zeitgleich mit der Frage danach, ob finanzielle Unterstützung für alle Bürger wirtschaftliche Gerechtigkeit stärken würde, geführt wird.



Abbildung 4: Screenshots der beiden Szenen aus dem TikTok-Video der @diegruenen zur Milliardensteuer (<https://www.tiktok.com/@diegruenen/video/7597023443879562518>)

⁴ <https://vm.tiktok.com/ZGduNWC11/>

Das Video selbst zeigt zunächst einen Screenshot eines Instagram-Posts der Tagesschau, der eine Reihe von 100€-Geldscheinen mit der Headline in weißen Buchstaben „Milliardär:innen werden reicher“ zeigt (siehe Abbildung 4, links). Das Hintergrundbild überlagert wird mithilfe der Greenscreen-Technik eine weibliche Person eingeblendet, deren Videoebene sich dann an den unteren rechten Bildrand bewegt. Von dort kommentiert die Person mithilfe einer Zeige- und Abwinkgeste den zweiten im Hintergrund eingeblendeten Screenshot (Abbildung 4, rechts). Dieser zeigt die zweite Slide des Instagram-Posts der Tagesschau, auf dem in einer Auflistung unter der Headline „Superreiche werden reicher“ über das wachsende Vermögen und eine mögliche Vermögenssteuer gesprochen wird. Zur gleichen Zeit wird ein darüber liegendes Textinsert eingeblendet, das den Text „Und die Regierung spricht weiter über Bürgergeld statt Milliardärssteuer“ sowie ein Clown-Emoji zeigt. Das gesamte Video wird begleitet von Ausschnitten aus dem für TikTok-Videos modifizierten Song „Want to Love - Just Raw“ von Aloboi, die vor allem durch ihre Reverb-Effekte und die Verstärkung des Bassbereichs herausstechen und beide Szenen jeweils mit einem dunklen, hallenden Klang unterlegen, der fast wie ein Störsignal wirkt. Der Song wird auf der Plattform meme-artig genutzt, indem er durch Prozesse der Rekontextualisierung und Replikation vor allem zur Untermalung dramatischer und emotionaler Videosequenzen dient, z. B. bei Charakter-Edits aus Serien, Filmen oder Videospielen.

Obwohl lediglich sieben Sekunden lang, vereint das Video also unterschiedlichste Zeichenmodalitäten zu einem komplexen kommunikativen Gesamtkonstrukt. Um alle diese Zeichenmodalitäten semiotisch fassen zu können, hilft es, die verschiedenen Canvases innerhalb des Videos zu identifizieren, was wir in Abbildung 5 darstellen. Wie in Bateman et al. (2026: Kap. 5 & 12) ausgeführt, gilt es zunächst, die unterschiedlichen Canvases, die die Plattform zur Verfügung stellt, voneinander zu unterscheiden. Für unser Beispiel betrachten wir daher die Benutzeroberfläche der TikTok-App, das Interface, wie es auf dem Bildschirm unseres Mobiltelefons angezeigt wird, als Hauptcanvas, der selbst Raum für verschiedene Medien und Artefakte bietet. Dieser Canvas ist zweidimensional und bleibt innerhalb der Plattform meistens unverändert. Nutzer:innen sind Teilnehmende an der gesamten Kommunikationssituation und entscheiden, welche Aspekte der Plattform sie durchstöbern oder welche Inhalte sie selbst teilen möchten. Wenn sie sich entscheiden, einen bestimmten Inhalt in Form eines Videos anzusehen, sind sie in Bezug auf diesen Inhalt natürlich Beobachtende, weil sie das abgespielte Video auf diesem Interface selbst nicht ändern können.

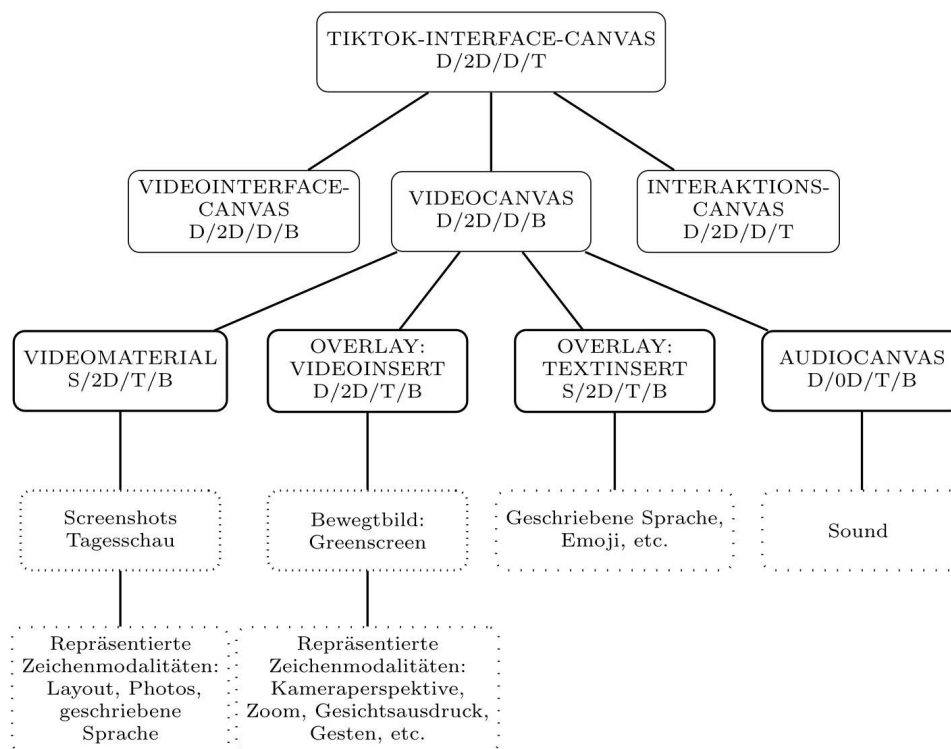


Abbildung 5: Hierarchische Darstellung der Canvases im TikTok-Video der Grünen. Weil der Fokus in der Analyse vor allem auf den Canvases im Video selbst liegt, haben wir der Übersicht halber die Subcanvases im Videointerface und auf dem Interaktionscanvas nicht weiter aufgeschlüsselt.

Was Nutzer:innen von ihrer eigenen Startseite aus sehen können, muss weiter unterteilt werden in den Canvas eines einzelnen Videos, also das audiovisuelle Material an sich, das durch die Plattform wiedergegeben wird, einen Interaktionscanvas, der es Nutzer:innen erlaubt, mit dem Video durch Likes, Teilen oder einen Kommentar zu interagieren, sowie den Videointerface-Canvas, der am unteren Bildrand weitere Informationen zum Video oder der gewählten Musik sowie dem Account bereitstellt. Diese drei Canvases bilden die erste Ebene in der dargestellten Hierarchie in Abbildung 5. Sowohl der Videocanvas als auch der Videointerface-Canvas lassen keine direkte Interaktion zu und machen Nutzer:innen dadurch zu Beobachtenden. Da sie Informationen beinhalten, die sich ändern und sich ggf. auch dynamisch über den Bildschirm bewegen, sind beide Canvases dynamisch und zweidimensional. Auch diese Canvases setzen natürlich unterschiedliche Zeichenmodalitäten ein, die Bestandteil einer Analyse sein können, vor allem wenn es zum Beispiel um Interaktionsmöglichkeiten oder den Einsatz von z. B. geschriebener Sprache und Emojis in einer Caption.

4.2.2 Analysefokus: Das Videomaterial

In unserem Fall konzentrieren wir uns im Folgenden auf das TikTok-Video selbst und damit auf den Canvas, der in unserer Abbildung zentral steht und wiederum in verschiedene Subcanvases aufgeteilt wird. Ähnlich der Analyse von Tschannen und Vieracker (2024) schauen wir uns das Videomaterial und die unterschiedlichen Materialien darin genauer an: Wie bei vielen Kurzvideoinhalten handelt es sich auch hier um ein Artefakt, das wiederum unterschiedliche andere Medienartefakte einbettet. Im Gegensatz zum statischen Post in Abschnitt 4.2 erscheint diese Art von Material aufgrund der Einbettung verschiedener Medienformen noch komplexer.

Unsere Canvasanalyse bestätigt diese Komplexität, indem sie aufzeigt, dass wir es mit vier verschiedenen Subcanvases zu tun haben, auf denen unterschiedliche Zeichenmodalitäten interagieren. Zum einen haben wir das zugrunde liegende Videomaterial, das in diesem Fall ein statischer Canvas ist, der zwei Screenshots darstellt. Bei diesen handelt es sich um Medienformen, die auf einer anderen Plattform (meistens Instagram) generiert und präsentiert wurden und von denen dann eine Abbildung erstellt wurde, die als Bilddatei in das TikTok-Video eingefügt wurde. Diesen Screenshots wurden auf einer weiteren, nun aber plattforminhärenten Ebene sowohl ein Video-Overlay mit Greenscreen-Technik als auch ein Text-Overlay hinzugefügt. Außerdem gibt es einen separaten Audio-Canvas, auf dem ein Musikstück an den Zusammenschnitt der zwei statischen und dynamischen Videosequenzen angepasst wurde. Zwar konstituieren all diese Einzelteile ein kohärentes und audio-visuell abgestimmtes Video, doch wird erst durch die Canvasanalyse deutlich, wie unterschiedlich die Materialitäten tatsächlich sind. Insbesondere die Aufsplittung statischer und dynamischer sowie visueller und auditiver Inhalte ist notwendig, um die unterschiedlichen Zeichenmodalitäten erfassen zu können und in einem weiteren Schritt ihr Zusammenspiel zu verdeutlichen.

Wie bereits in der Analyse des Instagram-Posts deutlich geworden ist, ist diese Art der medialen Darstellung, die wir in diesem TikTok-Video sehen, für soziale Medien sehr typisch und erfüllt eine Reihe weiterer wichtiger kommunikativer Funktionen. Der im TikTok-Video sichtbare Tagesschau-Post, der ursprünglich als eigenständiger Instagram-Post veröffentlicht wurde, erscheint hier in Form von Screenshots als rekontextualisiertes multimodales Artefakt. Dessen spezifische Konfiguration – etwa der Inhalt sowie die Plattformästhetik – erlangt durch die Wiederverwendung im TikTok-Video nun die Funktion eines Video-Kommentars. Wie bei sol-

chen Medienverwendungen typisch, werden zu dieser übernommenen spezifischen Konfiguration weitere plattform-inhärente Ausdrucksformen für die Vermittlung der kommunikativen Botschaft genutzt. Die im Video-Overlay sichtbare Person stellt in jedem Fall einen weiteren, durch die Plattform bereitgestellten Canvas mit eigener Materialität dar, der auf jeden Fall eigenständige Repräsentations- und Darstellungsbeziehungen abbildet. Gestik und Mimik der dargestellten Person operieren auf einem dynamischen, dreidimensionalen Canvas, der im Video aufgenommen und innerhalb des komplexen TikTok-Videos eingebettet wurde.

Hier wird deutlich, dass die Materialität des TikTok-Videos als Hauptcanvas unserer Analyse eine ausreichende Grundlage für den Gebrauch unterschiedlichster Zeichenmodalitäten aus anderen Medien bietet. Obwohl es ein zweidimensionaler, dynamischer Canvas ist, ist es möglich, statische Inhalte einzublenden und sogar dreidimensionale Gegebenheiten zu reproduzieren. Um die Komplexität einer solchen kommunikativen Situation zu begreifen, ist es wichtig, diese unterschiedlichen Mechanismen der Zeichenprozessierung zu differenzieren und zu beschreiben, in welcher Darstellungsbeziehung sie zueinander stehen. Wichtig ist hierbei auch die Feststellung, dass bestimmte Medien wie beispielsweise Screenshots oder spezifische Darstellungen von Mimik und Gestik, wie sie im Overlay-Video eingesetzt werden, immer auch eine bestimmte Rezeptionspraxis mit sich bringen und somit spezifisches Wissen und Medienkompetenz voraussetzen.

Für die weitere Analyse unseres Videos ist es deswegen notwendig, eine feingliedrige Differenzierung aller dieser Aspekte vorzunehmen, wofür unsere Canvas-Analyse eine wichtige methodische Grundlage schafft. Wir können dadurch auch feststellen, dass Sprache in diesem Video lediglich in Form von geschriebener Sprache in den rekontextualisierten Screenshots sowie im Textinsert in Kombination mit einem Emoji vorkommt. Die zentrale kommunikative Funktion übernehmen hingegen videospezifische Aspekte sowie die im Video dargestellte Gestik und Mimik, die maßgeblich für Rahmung und Bewertung des Inhalts sind.

5 Zusammenfassung

Die medienlinguistische Analyse postdigitaler Kommunikationsformate erfordert methodologisch eine präzisere Modellierung der in diesen Formen vorliegenden semiotischen Komplexität und damit materiellen Granularität, als sie bislang im deutschsprachigen Raum vorgenommen wird. Zwar ist die Berücksichtigung plattformspezifischer Kompositionsebenen anerkannt, wird jedoch aus multimodal-

analytischer Perspektive bislang nur zögerlich und häufig mit unzureichender Systematik und Präzision umgesetzt. Mit unserem Beitrag haben wir versucht zu zeigen, dass die Fokussierung auf einzelne Zeichenmodalitäten – vor allem Sprache und Bild – zu kurz greift, um die konstitutiven Praktiken postdigitaler Kommunikationsformate angemessen zu erfassen.

Wir plädieren für eine gleichwertige Berücksichtigung sämtlicher bedeutungstragender Einheiten. Gerade in sozialen Medien, die durch Kurzformate geprägt sind, vollzieht sich der Bedeutungsprozess nicht nur in besonders verdichteter Form, sondern zugleich über mehrere Materialitäten hinweg. Im Sinne der Multimodalitätsforschung, die Bedeutungskonstitution in multimodalen Ensembles als multiplikativ (Lemke 1998) und nicht additiv versteht, erweist sich eine Herangehensweise als unerlässlich, die sowohl die verschiedenen Zeichenmodalitäten mit gleicher analytischer Sorgfalt behandelt als auch die Dimensionen medialer Materialität systematisch erfasst.

Die Fallanalysen verdeutlichen, dass der Canvas-Ansatz (vgl. Wildfeuer et al. 2020; Bateman et al. 2017, 2026) besonders dort analytische Stärke aufweist, wo unterschiedliche Formen von medialer Materialität rekursiv miteinander verschränkt werden. Dadurch können komplexe Repräsentationsbeziehungen sichtbar und in Beziehung gebracht werden – die in der Forschung postulierte „Unüberschaubarkeit“ postdigitaler Kommunikationsformate lässt sich somit systematisch strukturieren. Die Rekonstruktion plattformgebundener Kompositionslogiken sowie die Identifikation von Regularitäten bzw. Ausdruckseinheiten auf der materiellen Oberfläche macht deutlich, dass beide Plattformen mehrere Handlungsspielräume eröffnen, die sich aus spezifischen technisch-medialen und sozio-kommunikativen Möglichkeiten ergeben. Erst die Visualisierung der komplexen Bedeutungsarchitektur des Instagram-Karussell-Posts und des TikTok-Videos zeigt das simultane Auftreten von Ausdruckseinheiten sowohl auf einzelnen Ebenen als auch über mehrere, hierarchisch organisierte Ebenen hinweg. Dabei wird sichtbar, dass eine Vielzahl von Materialitäten beteiligt ist, mithilfe derer Zeichenmodalitäten miteinander interagieren.

Die Komplexität dieser Kurzformate resultiert somit nicht allein aus ihrer semiotischen Dichte, sondern auch aus der Verschränkung unterschiedlicher materieller Dimensionen. Während beispielsweise der Inhalt des Instagram-Bildbereichs durch unterschiedliche Formattypen (Video, Bild- oder Karussell-Post) gekennzeichnet ist, ist das TikTok-Video prinzipiell auf das Bewegtbild ausgerichtet, was wiederum eine spezifische Zeichenprozessierung ermöglicht. Eine materialitätsorientierte Operationalisierung erlaubt es, die Verschränkungen analytisch sichtbar zu machen: Sie erfasst sowohl die

von Plattformen bereitgestellten Werkzeuge bzw. technischen Ressourcen (z. B. Layout-Vorlagen, Greenscreen-Funktion, diverse Sticker sowie Audioebenen) als auch die von ihnen realisierten Zeichenmodalitäten. Dadurch werden plattformübergreifende Ähnlichkeiten ebenso sichtbar wie strukturelle Differenzen in den jeweiligen Oberflächen der Zeichenprozessierung. Die Canvas-Analyse bringt zudem zum Vorschein, dass Sprache in post-digitalen Kontexten nicht länger als primärer Bedeutungsträger gesehen werden kann. Vielmehr ist sie als eine von vielen Zeichenmodalitäten zu begreifen, mittels derer kommunikative Botschaften vermittelt werden.

Unser Beitrag versteht sich dabei ausdrücklich als methodologische Formgebung. Statt einer detaillierten Analyse pragma-semantischer Relationen zwischen Zeichenmodalitäten steht im Zentrum unserer Fallanalysen zunächst eine systematische Segmentierung und zugleich holistische Erfassung komplexer Kommunikationsformate. Die Komplexität dieser Formate variiert immens und entwickelt sich zudem immer weiter. Genau für diese Vielfalt stellen wir einen grundsätzlichen Rahmen bereit, auf dem weitere methodisch- und format-spezifische Schritte aufbauen können. Zweifellos ist so eine Herangehensweise mit einem erheblichen Annotationsaufwand verbunden. Um jedoch die Vielschichtigkeit bestehender und potenziell sich neu entwickelnder Kurzformate und der darin integrierten Ausdrucksmöglichkeiten gerecht zu werden, ist eine systematische, materialitätsorientierte Vorgehensweise unerlässlich. Aktuelle Arbeiten folgen dieser Ausrichtung in unterschiedlichen Dimensionen zur Anwendung auf verschiedene neuere und neueste Formate (vgl. u. a. Mattei 2024; Thiele et al. 2024; Tseng et al. 2026). Die feingliedrige Auseinandersetzung bietet somit einen konkreten Lösungsansatz für viele unterschiedliche Analysen medialer Aspekte. Aus dieser Perspektive sehen wir die Zukunft der Medienlinguistik weniger in der Erweiterung bestehender Analyseverfahren als in einer stärkeren Orientierung an der medialen Materialität. Denn nur wenn deren materielle und plattformspezifische Strukturierung zum Ausgangspunkt der Analyse wird, lassen sich die komplexen Beziehungen zwischen Ausdrucksformen und den damit einhergehenden (Bedeutungs-)Prozessen hinreichend präzise erfassen.

Literatur

- Androutsopoulos, Jannis/Vogel, Friedemann (Hg.) (2024): *Handbuch Sprache und digitale Kommunikation*. Berlin, Boston: De Gruyter.
DOI: 10.1515/9783110744163

IG-Post @Baby got business babygotbusiness: *NFTs*.

URL: https://www.instagram.com/p/CeA-VWYKoCF/?img_index=1 (28.03.2026)

Bateman, John A. (2008): *Multimodality and Genre: a foundation for the systematic analysis of multimodal documents*. Basingstoke: Palgrave Macmillan.

Bateman, John A./Hiippala, Tuomo (2021): From Data to Patterns: On the Role of Models in Empirical Multimodality Research. In: Pflaeging, Jana/Wildfeuer, Janina/Bateman, John A. (Hg.): *Empirical Multimodality Research: Methods, Evaluations, Implications*. Boston: De Gruyter, 65–90.

Bateman, John A./Tseng, Chiao-I (2023): Respondenz von John A. Bateman & Chiao-I Tseng. In: Meiler, Matthias/Siefkes, Martin (Hg.): *Linguistische Methodenreflexion im Aufbruch. Beiträge zur aktuellen Diskussion im Schnittpunkt von Ethnographie und Digital Humanities, Multimodalität und Mixed Methods*. Berlin/Boston: De Gruyter, 315–318.

DOI: 10.1515/9783111043616-016

Bateman, John A./Wildfeuer, Janina/Hiippala, Tuomo (2017): *Multimodality: Foundations, Research and Analysis – A Problem-Oriented Introduction*. Berlin: De Gruyter.

Bateman, John A./Wildfeuer, Janina/Hiippala, Tuomo (2026): *Multimodality. A Hands-On Guide*. Berlin: De Gruyter.

Beißwenger, Michael/Lüngen, Harald/Meier-Vieracker, Simon/Steinsiek, Sarah/Storrer, Angelika/Herzberg, Laura (2025): *Korpusgestützte Analyse internetbasierter Kommunikation. Eine Einführung*. Stuttgart: Metzler.

DOI: 10.1007/978-3-476-05977-2

Bender, Michael/Mell, Ruth M./Wildfeuer, Janina (2022): Zur Spezifik digitaler Medien als Diskursraum: Materialität, Daten, Affordanzen. In: Gredel, Eva (Hg.): *Diskurse – digital: Theorien, Methoden, Anwendungen*. Berlin/Boston: De Gruyter, 27–46.

DOI: 10.1515/9783110721447-003

Bülow, Lars/Merten, Marie-Luis (2023): Multimodale Metaphern im Kontext von Internet-Memes. Korpuspragmatische und kognitionslinguistische Zugänge zu einem soziokognitiven Online-Phänomen. In: Meier-Vieracker, Simon/Bülow, Lars/Marx, Konstanze/Mroczynski, Robert (Hg.): *Digitale Pragmatik*. Stuttgart: Metzler.

Caple, Helen (2020): Image-Centric Practices on Instagram. Subtle shifts in footing. In: Stöckl, Hartmut/Caple, Helen/Pflaeging, Jana (Hg.): *Shifts Toward Image-Centricity in Contemporary Multimodal Practices*. London: Routledge, 153–176.

- Fischer, Friederike/Meier-Vieracker, Simon/Niendorf, Lisa (Hg.) (2025): *TikTok – Memefication und Performance: Interdisziplinäre Zugänge*. (Digitale Linguistik, 2). Berlin/Heidelberg: Metzler.
DOI: 10.1007/978-3-662-70712-8
- Fix, Ulla (2008): Nichtsprachliches als Textfaktor. Medialität, Materialität, Lokalität. In: *Zeitschrift für germanistische Linguistik* 36 (3), 343–354.
- Fix, Ulla (2024): ‘Text’ – ‘Keine Texte’ – ‘Nicht-Texte’ – ‘Dennoch Texte’? In: *Osnabrücker Beiträge zur Sprachtheorie* (103), 11–29.
- Georgakopoulou, Alexandra/Giaxoglou, Korina/Patron, Sophia (Hg.) (2023): *Small Stories Research: Tales, Tellings, and Tellers Across Contexts*. London: Routledge.
DOI: 10.4324/9781003253563
- Gredel, Eva (Hg.) (2022): *Diskurse – digital: Theorien, Methoden, Anwendungen*. Berlin/Boston: De Gruyter.
DOI: 10.1515/9783110721447
- Grzenkiewicz, Maciej/Wildfeuer, Janina (2025): Addressing TikTok’s multimodal complexity. A multi-level annotation scheme for the audio-visual design of short video content. In: *Digital Scholarship in the Humanities* 40 (4), 1143–1166.
DOI: <https://doi.org/10.1093/llc/fqaf047>
- Guerra, Giulia Isabella (2026): The Short Video Format. Conceptualizing a Ubiquitous Artifact of Visual Platforms. In: *International Journal of Communication* (20), 792–816.
DOI: 10.65476/rdgsy150
- Hess-Lüttich, Ernest/Warnke, Ingo H./Reisigl, Martin/Kämper, Heide (Hg.) (2017): *Diskurs – semiotisch: Aspekte multiformaler Diskurskodierung*. Berlin/Boston: De Gruyter.
- Instagram (2026): *Instagram Blog* (Zugriff am 16.03.2026).
URL: <https://about.instagram.com/blog>
- Kabatnik, Susanne/Bülow, Lars/Merten, Marie-Luis/Mroczynski, Robert (Hg.) (2024): *Pragmatik multimodal* (Studien zur Pragmatik, 7). Tübingen: Narr Francke Attempto.
DOI: 10.24053/9783823395829
- Kindler-Mathôt, Claire (2025): eins zwei drei vor wusch. Dynamische Entfaltung multimodaler Bedeutung einer onomatopoetischen Interjektion. In: Kabatnik, Susanne/Bülow, Lars/Merten, Marie-Luis/Mroczynski, Robert (Hg.): *Pragmatik multimodal*. Tübingen: Narr Francke Attempto, 165–194.
- Kress, Gunther (2010): *Multimodality. A Social Semiotic Approach to Contemporary Communication*. London: Routledge.
- Lankshear, Colin/Knobel, Michele (2019): Memes, Macros, Meaning, and Menace: Some Trends in Internet Memes. In: *Journal of Communication and Media Studies*, 4 (4), 43–57.

- Lemke, Jay (1998): Multiplying Meaning: Visual and Verbal Semiotics in Scientific Text. In: Martin, J. R./Veel, Robert (Hg.): *Reading Science: Critical and Functional Perspectives on Discourses of Science*. London: Routledge, 87–113.
- Logi, Lorenzo/Zappavigna, Michele (2024): GIFs as Social Media Paralanguage. In: *Language@Internet* (22), 32–56.
- Luginbühl, Martin (2025): TV News on Instagram. Affordances, genres, journalistic cultures. In: *Discourse & Communication* 19 (1), 46–71.
- Marx, Konstanze/Lobin, Henning/Schmidt, Axel (2020): *Deutsch in sozialen Medien. Interaktiv – multimodal – vielfältig*. Berlin/Boston: De Gruyter.
DOI: 10.1515/9783110679885
- Mattei, Elena (2024): Approaching tourism communication with empirical multimodality: exploratory analysis of Instagram and website photography through data-driven labeling. In: *Frontiers in Communication* (9).
DOI: 10.3389/fcomm.2024.1355406
- Meier-Vieracker, Simon/Bülow, Lars/Marx, Konstanze/Mroczynski, Robert (Hg.) (2023): *Digitale Pragmatik* (Digitale Linguistik, 1). Berlin/Heidelberg: Metzler.
DOI: 10.1007/978-3-662-65373-9
- Meiler, Matthias (2019): Zur praxeologischen Verhältnisbestimmung von Materialität, Medialität und Mentalität *oder*: Medien als Praxis. In: *Zeitschrift für Semiotik* 41 (1–2), 63–88.
- Merten, Marie-Luis (2024): Körpersemantisierungen im #bodylove-Diskurs. Eine Korpusstudie zu Instagram-Posts als multimodales Positionierungsformat. In: Kabatnik, Susanne/Bülow, Lars/Merten, Marie-Luis/Mroczynski, Robert (Hg.): *Pragmatik multimodal* (Studien zur Pragmatik, 7). Tübingen: Narr Francke Attempto, 223–256.
- Merten, Marie-Luis/Bülow, Lars/Kabatnik, Susanne/Mroczynski, Robert (2024): Pragmatik multimodal – Einführung. In: Kabatnik, Susanne/Bülow, Lars/Merten, Marie-Luis/Mroczynski, Robert (Hg.): *Pragmatik multimodal* (Studien zur Pragmatik, 7). Tübingen: Narr Francke Attempto. 7–25.
DOI: 10.24053/9783823395829
- Mondada, Lorenza (2019): Contemporary issues in conversation analysis: Embodiment and materiality, multimodality and multi-sensoriality in social interaction. In: *Journal of Pragmatics* 145, 47–62.
DOI: 10.1016/j.pragma.2019.01.016

- Nam, Jinyoung/Jung, Yoonhyuk (2021): Digital natives' snack content consumption and their goals: A means-end chain approach. In: *Telematics and Informatics* 63.
DOI: 10.1016/j.tele.2021.101664
- O'Halloran, Kay L./Pal, Gautam/Jin, Minhao (2021): Multimodal Approach to Analysing Big Data. Multimodal Approach to Analysing Big Social and News Media Data. In: *Discourse, Context & Media*, 40.
DOI: <https://doi.org/10.1016/j.dcm.2021.100467>
- Schmitz, Ulrich (2022): Kann und soll Linguistik angesichts multimodaler Kommunikation eine eigene Disziplin sein? In: *Osnabrücker Beiträge zur Sprachtheorie* 100, 11–28.
- Shepherd, Ifan D. H. (2005): From Cattle and Coke to Charlie: Meeting the Challenge of Self Marketing and Personal Branding. In: *Journal of Marketing Management* 21 (5–6), 589–606.
- Spitzmüller, Jürgen (2018): Multimodalität und Materialität im Diskurs. In: Warnke, Ingo (Hg.): *Handbuch Diskurs*. Berlin/Boston: De Gruyter, 521–540.
- Statista (2025): *Social Media – Daten & Fakten*.
URL: <https://de.statista.com/themen/1842/soziale-netzwerke/#topicOverview> (28.03.2026).
- Swales, John M. (1990): *Genre analysis. English in academic and research settings*. Cambridge: Cambridge University Press.
- Thiele, Leandra/Schmidt-Borcherding, Florian/Bateman John A. (2024): All eyes on the signal? – Mapping cohesive discourse structures with eye-tracking data of explanation videos. In: *Frontiers in Communication*.
DOI: 10.3389/fcomm.2024.1356495
- Thomson, T. J./Pfurtscheller, Daniel/Lobinger, Katharina/Laba, Nataliia/Christ, Katharina (2025): Visual Literacies in Transition. Multimodal co-production with generative AI. In: *Journal of Visual Literacy* 44 (4), 384–405.
DOI: 10.1080/1051144X.2025.2566592
- Tschannen, Jana/Meier-Vieracker, Simon (2024): Performing Science. In: Jaki, Sylvia/Meiler, Matthias/Pflaeging, Jana/Wildfeuer, Janina (Hg.): *Multimodalität in Wissensformaten*. Berlin: Peter Lang, 285–322.
- Tseng, Chiao-I/Fielder, Emily/Wenk, Luise Charlotte Claire (2026): Multimodal semiotic content analysis: combining quantitative large-data analysis with qualitative semiotic coding. In: *Visual Communication* (ahead of print) 0 (0), 1–27.
DOI: 10.1177/14703572261418120
- TikTok-Video @diegruenen (2026): *Milliardärssteuer*.
URL: <https://www.tiktok.com/@diegruenen/video/7597023443879562518> (28.03.2026)

- Viher, Stina (2021): *Kommunikation in sozialen Netzwerken via Memes. Zur (Re)konstruktion sowie damit verbundener Form und Funktion der Text-Bild-Hybride auf Facebook*. Göttingen: Vienna University Press.
- Viher, Stina (i. E.): Snack-Content auf Instagram– Leicht verdaulich oder doch ein üppiges Mehr-Gänge-Menü? In: Bülow, Lars/Luttermann, Karin (Hg.): *Digitale Werbung* (Special Issue).
- Wildfeuer, Janina (2014): *Film Discourse Interpretation. Towards a New Paradigm for Multimodal Film Analysis*. London, New York: Routledge.
- Wildfeuer, Janina/Bateman, John A. (2018): Theoretische und methodologische Perspektiven des Multimodalitätskonzepts aus linguistischer Sicht. In: *IMAGE* 28, 5–45.
- Wildfeuer, Janina/Bateman, John A./Hiippala, Tuomo (2020): *Multimodalität. Grundlagen, Forschung und Analyse*. Berlin: De Gruyter.
- Wildfeuer, Janina (2024): Mediality vs. Materiality. A multimodal perspective on the notion of media as procedures. In: Luginbühl, Martin/Schneider, Jan Georg (Hg.): *Media as Procedures of Communication*. Amsterdam: John Benjamins, 16–39.
DOI: 10.1075/pbns.348.02wil
- Wildfeuer, Janina (2025). Multimodale Texte lesen: Theoretische und methodische Bausteine aus der Multimodalitätsforschung. In Staiger, Michael (Hg.): *Schrift / Bild – Lesen: Interdisziplinäre Perspektiven für die Leseforschung*, Berlin/Heidelberg: Metzler, 27–46.
DOI: 10.1007/978-3-662-69717-7_2