

## MULTIMODALES ERKLÄREN IN UNTERRICHT & VIDEO EIN MODELL ZUR ANALYSE UND PRODUKTION VON ERKLÄRVIDEOS

---

Kirsten Schindler  
Bergische Universität Wuppertal | kschindler@uni-wuppertal.de

Matthias Knopp  
Universität zu Köln | matthias.knopp@uni-koeln.de

### ABSTRACT

Das Erklären stellt eine alltägliche Praxis in der Schule dar; es ist konstitutiv für die Bildungsinstitution. Erklärt wird in Schulbüchern und im Unterricht, von Lehrkräften und von Schüler\*innen, gezielt, vorbereitet oder auch nebenbei und en passant, im Deutschunterricht und in allen anderen Schulfächern. In den letzten Jahren ist diese Erklärpraxis durch Erklärvideos ergänzt worden, die inner- und außerhalb von Unterricht, zur Nach- und Vorbereitung und teils auch als Ersatz von Unterricht genutzt werden. Wie aber unterscheidet sich das Erklären im Video vom Erklären im Unterricht? In unserem Beitrag modellieren wir das Erklärvideo aus medienlinguistischer und mediendidaktischer Perspektive und bestimmen das multimodale Erklären auch im Verhältnis zum Erklären im Unterricht.

### SCHLAGWÖRTER

— ERKLÄRVIDEO — ERKLÄREN IN DER SCHULE — MULTIMODALITÄT

*Gefördert durch die Deutsche Forschungsgemeinschaft (DFG) – Projektnummer 425885011*

**Copyright** Dieser Artikel wird unter der Creative-Commons-Lizenz CC BY-ND 4.0 veröffentlicht:  
<https://creativecommons.org/licenses/by-nd/4.0/deed.de>

**ABSTRACT (ENGLISH)**

**Multimodal explaining in the Classroom & per video. A model for the analysis and production of explanatory videos.**

Explaining is an everyday practice in schools; it is constitutive for the institution itself. Explanations are given in textbooks and in lessons, by teachers and students, purposefully, prepared or also incidentally and en passant, in German lessons and in all other school subjects. In recent years, this explanatory practice has been supplemented by explanatory videos, which are used in and outside of lessons, for follow-up and preparation and sometimes also as a substitute for lessons. But how does explaining in videos differ from explaining in class? In our article we propose a model for the explanatory video using a media-linguistic and media-didactic perspective and define the video as a specific form of multimodality also in relation to explaining in class.

**KEYWORDS**

— EXPLANATORY VIDEO — EXPLAINING IN SCHOOL — MULTIMODALITY

## 1 — EINLEITUNG

2019 veröffentlicht der Rat für kulturelle Bildung eine Studie zur Nutzung kultureller Bildungsangebote durch 12-19-Jährige; befragt wurden 818 Jugendliche. 86% der Befragten geben an, dass sie YouTube nutzen würden, das Interesse steigt mit zunehmendem Alter, bei den 18-19-Jährigen sind es bereits 93%. Die Hälfte von ihnen schätzt YouTube nicht nur als Freizeitmedium, sondern profitiert von YouTube auch für die Schule. So geben sie an, mit YouTube-Videos könnten Inhalte wiederholt und Hausaufgaben unterstützt werden, Wissen würde durch solche Videos vertieft bzw. gezielt für Prüfungen aufbereitet (vgl. Rat für kulturelle Bildung 2019). In der JIM-Studie von 2021 bestätigt sich dieses Bild. YouTube steht zwar durchaus in Konkurrenz zu vielfältigen anderen Angeboten digitaler Medien (insbesondere TikTok), genutzt wird YouTube neben anderen Funktionen aber auch, um „Wissensformate“ und „Tutorials für Themen aus der Schule“ zu rezipieren (vgl. [Medienpädagogischer Forschungsverbund Südwest 2021, 48](#)).<sup>1</sup> Dabei scheint die frequente Nutzung auch ein Indiz dafür zu sein, dass Schule möglicherweise nicht mehr leisten kann, wozu sie angetreten ist und/oder die multimodalen Möglichkeiten eines videobasierten Erklärens besonders unterstützend wirken sowie attraktiv und niederschwellig sind. In der Studie von Cwielong und Kommer (2020) gibt die Hälfte der 1.000 befragten Schüler\*innen der Klassenstufe 9-11 an, „sie würden die Erklärvideos nutzen, weil ihre Lehrerinnen und Lehrer schlecht erklären“ (42). In der eingangs zitierten Studie wird daraus der Schluss gezogen:

Damit hat die kommerzielle Plattform YouTube eine folgenreiche Änderung der Bildungskonstellation bewirkt. Ein weiterer Akteur wirkt als Hilfs- und Nachhilfelehrer intensiv an den schulbezogenen Bildungsprozessen der Jugendlichen mit. (Rat für kulturelle Bildung 2019, 8)

Inwieweit die verbreitete Nutzung tatsächlich ursächlich auf fehlende Kompetenzen der Lehrkräfte zurückgeht, ist an dieser Stelle nicht nachprüfbar. Interessant ist aber, dass Erklärvideos aus Sicht der Anwender\*innen offenbar eine didaktische Lücke schließen. Was aber macht das Erklären in Erklärvideos aus bzw. was macht es so erfolgreich und unterscheidet es vom Erklären in analogen, face-to-face-Kontexten? In unserem Beitrag modellieren wir die multimodale Erklärpraxis (via YouTube) aus mediendidaktischer bzw. medienlinguistischer Perspektive, auch indem wir es mit dem Beispiel des Erklärens im Unterricht kontrastieren, das wir zunächst bestimmen. Für die Analyse und Produktion von Erklärvideos lassen sich – so unsere Annahme – mit dem Modell wichtige Grundlagen schaffen.

## 2 — ERKLÄREN ERKLÄREN

### 2.1 ERKLÄREN IM UNTERRICHT

Beim Erklären werden zwei Konstrukte zueinander in Bezug gesetzt. Das bislang unbekannte Konstrukt, Explanandum, sowie das bekannte Konstrukt, Explanans (vgl. Klein 2009, 2001; Vogt 2009).<sup>2</sup> Eingebettet ist dies in der Regel in eine kommunikative Situation, in der eine lokale Wissensasymmetrie besteht; eine\*r der Aktant\*innen muss also über ein Wissen verfügen, das der\*die andere nicht zur Verfügung hat (vgl. Neumeister / Vogt 2009, 562). Dabei gibt es aber durchaus unterschiedliche Konstellationen für die Initiierung von Erklärungen:

<sup>1</sup> Vgl. für die Relevanz von Erklärvideos für den Deutschunterricht Hackl 2021.

- Eine Person hat einen Erklärbedarf und expliziert diesen initial oder
- jemand erklärt sich selbstständig zur erklärenden Person, ohne dass jemand die Erklärung erbeten hätte; und in
- schulischen Settings wird der Erklärbedarf häufig durch die erklärende Person etabliert (durch die Lehrkraft), das Erklären dient dann eher der Wissensdemonstration (s. zweiten Aufzählungsstrich), der Aktivierung von (bereits vorhandenem) Wissen und dem Schaffen einer gemeinsamen Wissensbasis sowie der Bearbeitung von etwaigen falschen Vorstellungen von Sachverhalten o.Ä.<sup>3</sup>

Mit dem Erklärvorgang wird die lokale Wissensasymmetrie idealerweise behoben.

Wir wollen dies an einem Beispiel des Erklärens im Unterricht illustrieren und zugleich konkretisieren. Die auf der ViLLA-Plattform der Universität zu Köln verfügbare Aufnahme startet mit dem Beginn einer Deutschstunde in der 9. Jahrgangsstufe an einer Gesamtschule. Die Lehrkraft steht den Schüler\*innen zugewandt vor der Tafel platziert, auf der bereits der Stundenverlauf angeschrieben ist (Abb. 1).

Wir wollen dies an einem Beispiel des Erklärens im Unterricht illustrieren und zugleich konkretisieren. Die auf der ViLLA-Plattform der Universität zu Köln verfügbare Aufnahme startet mit dem Beginn einer Deutschstunde in der 9. Jahrgangsstufe an einer Gesamtschule. Die Lehrkraft steht den Schüler\*innen zugewandt vor der Tafel platziert, auf der bereits der Stundenverlauf angeschrieben ist (Abb. 1).

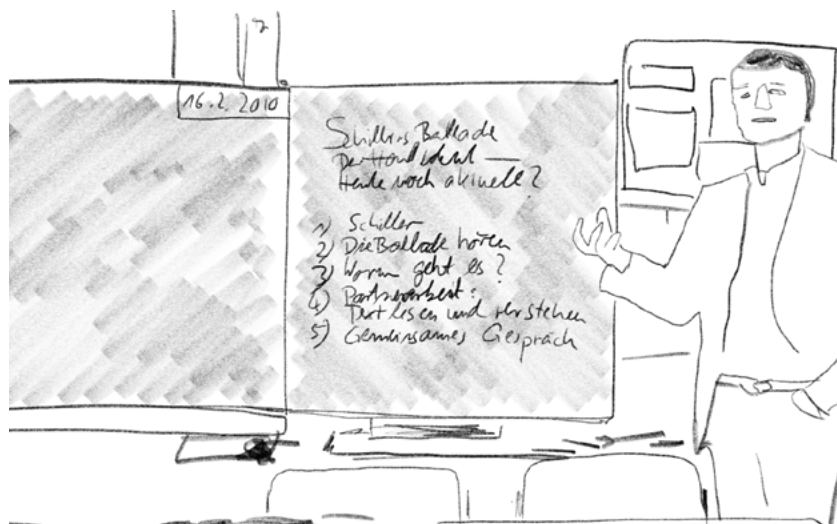


Abb. 1: Adressierung der Schüler\*innen mittels Körperpositur, Blickkontakt (Videoplattform ViLLA an der Universität zu Köln; Video „Schiller – Der Handschuh“, aus: Unruh / Petersen 2021)

Parallel zum begleitenden Sprechen zeigt die Lehrkraft dann auf die einzelnen, an der Tafel nummerierten Punkte (Aufzählung/Liste), die die Unterrichtsphasen kennzeichnen (Abb. 2).

<sup>2</sup> Vgl. zu schriftlichen Formen und sprachlichen Handlungen des Erklärens u.a. Feilke / Rezat 2021.

<sup>3</sup> Wir danken Vivien Heller für wertvolle Hinweise.

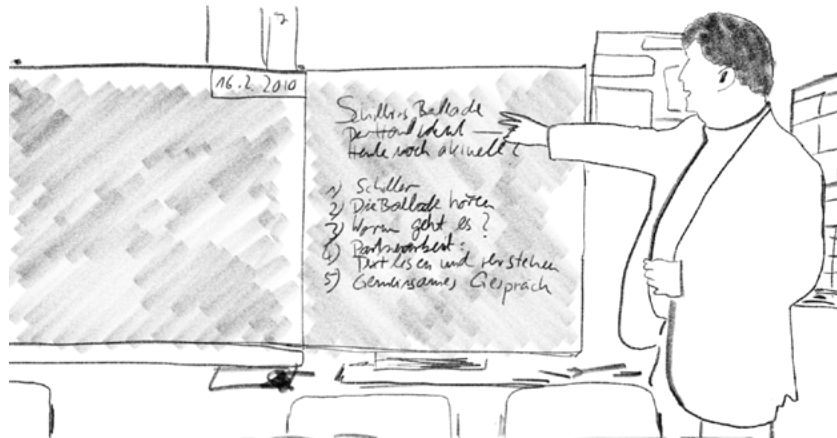


Abb. 2: Lehrerhand – begleitendes Zeigen am Tafelbild (Videoplattform ViLLA an der Universität zu Köln; Video „Schiller – Der Handschuh“, aus: Unruh / Petersen 2021)

Eingebettet ist die gesamte Sequenz gewissermaßen in eine übergeordnete Frage, die auch oben auf der Tafel verkürzt notiert ist: „Ist Schillers Ballade ‚Der Handschuh‘ heute noch aktuell?“ Wenngleich im Anschluss nicht ausschließlich und nicht allein von der Lehrkraft erklärt wird, so sind alle nachfolgenden kommunikativen Handlungen (dazu gehört das Beschreiben, Informieren, Illustrieren etc.) dem Zweck untergeordnet, diese Eingangsfrage zu beantworten.

Mit der Frage „Ist xy noch aktuell“ wird also eine Erklärsequenz initiiert, die sich dem Typ ERKLÄREN-WARUM zuordnen lässt. Beim ERKLÄREN-WARUM werden Phänomene, Vorgänge etc. kausal begründet, es wird also erläutert, was z.B. die Ursache eines Phänomens ist oder worin Gründe für eine bestimmte Einschätzung liegen. Neben dem ERKLÄREN-WARUM unterscheidet Klein (2009) noch zwei weitere Erklärungstypen, das ERKLÄREN-WAS und das ERKLÄREN-WIE. Beim ERKLÄREN-WAS (ergänzt durch ERKLÄREN-WER) werden Begriffe, Phänomene oder Personen definiert. Friedrich von Schiller (Explanandum) beispielsweise, der in der Stunde eine zentrale Rolle spielt, könnte dann als „deutscher Dichter des 18. Jahrhunderts“ (Explanans) eingeführt werden, so wie es die Lehrkraft zu Beginn auch tut. ERKLÄREN-WIE schließlich lässt Vorgänge oder Prozesse nachvollziehbar werden. Dies lässt sich besonders anschaulich an naturwissenschaftlichen Phänomenen zeigen, wie z.B. der Photosynthese, dem Blutkreislauf oder der Funktionsweise eines Siedewasserreaktors. Im obigen Beispiel ließe sich das übertragen auf die Wirkungsweisen, die bestimmte sprachliche Phänomene in der Ballade auf die Leser\*innen haben. Was zeigt der Ausschnitt noch?

Erklärungen im (Deutsch-)Unterricht sind in der Regel multimodal:

Multimodale Texte und kommunikatives Handeln beruhen wesenseigen auf der semantischen und funktionalen Verknüpfung bzw. wechselseitigen Integration verschiedener Zeichenmodalitäten – wie Sprache, Bild, Musik, Geräusch in einem textuellen Rahmen (z. B. in Film, Comic oder illustrierter Zeitschrift). Multimodalität schließt aber auch die Realisierung einer Zeichenmodalität in unterschiedlichen Medien (wie z. B. Sprache in Schrift, Rede und Gesten) sowie auch die Kombination mehrere (sic!) medialer Formate im Kontext der allgemein wachsenden Medienkonvergenz (z. B. Fernsehen, Lesen, Chatten, Twittern, Skypen) ein. (Klug / Stöckl 2016, VII)

Die Lehrkraft spricht (verbaler bzw. auditiver Kanal, auditive Semiosen und akustisches Zeichensystem; vgl. Habscheid 2005), verweist gestisch auf bestimmte Materialien (Tafelanschrieb, im Verlauf mit Magneten befestigte Ausdrucke an der Tafel (z.B. Fotografie von Schiller) oder Personen (um Schüler\*innen den Turn zu geben) (optischer Kanal, visuelle Semiosen; optische Zeichensysteme), positioniert sich im Klassenraum, in dem sie wechselnd stehend und sitzend agiert (Proxemik).

Eine solche Erklärpraxis stellt für die Schüler\*innen zunächst einmal den Regelfall (Prototyp) dar, da sie damit weitgehend sozialisiert wurden. Das Erklären im Erklärvideo ähnelt diesem Erklären zwar, nutzt aber – bedingt durch den digitalen medialen Raum beispielsweise andere Verweismittel und (teils) andere Modalitäten.

Nachdem es in den 2000er-Jahren in der Deutschdidaktik eine breitere Forschung, vor allem zum mündlichen Erklären, gegeben hat (siehe auch Heller 2021), ist in den letzten Jahren mit den Arbeiten von Feilke / Rezat (2021) u.a. der Blick stärker auf schriftliche Formen und sprachliche Handlungen des Erklärens gerückt worden. Mit den neueren Arbeiten zu Erklärvideos (u.a. Anders et al. 2019; Cwielong / Kommer 2020; Dube 2021; Dube / Boerdgen 2022; Ebel 2021; Findeisen / Horn / Seifried 2019; Frei et al. 2020; Knopp / Schindler 2021; Normann 2020; Pertzel 2021; Schindler / Knopp (in Begutachtung); Schöne / Wedler 2021; Torrau 2021; von Brand 2020; Wolk 2018) erweitert sich dieser Blick auf multimodale Aspekte des Erklärens, Sprache (geschrieben und gesprochen) agiert hier im Zusammenspiel mit anderen Modalitäten. Wie dieses Zusammenspiel genau zu beschreiben ist und wie dies mediendidaktisch perspektiviert werden kann, wollen wir in unserem Beitrag herausarbeiten. Zentral erscheinen uns dabei die zeitliche Sukzession und die Koinzidenz des Auftretens unterschiedlicher Elemente.

## 2.2 ERKLÄREN IM ERKLÄRVIDEO

Kontrastieren wir daher zunächst das Erklären im Unterricht mit einem Erklärbeispiel aus dem Internet. Auch hier dient die Beispielsequenz dem Illustrieren und Konkretisieren typischer Praktiken; eine genaue Analyse des multimodalen Zusammenspiels erfolgt im Anschluss. Das Beispiel entstammt dem Kanal von Studyflix. Auf Studyflix werden kostenfreie bzw. durch Werbung finanzierte Lerninhalte angeboten. Das Portal bewirbt sich mit dem Hinweis auf individuelle Lernzuwächse (ähnlich, wie es auch Nachhilfeinstitute tun) (s. Abb. 3).



Abb. 3: Bildschirmfoto der Startseite von <https://studyflix.de>

Das gezeigte Erklärvideo stellt das Thema *Ballade* in eine Reihe weiterer, verwandter Themen zur Lyrik, nummeriert als Video 2/9. Es ist 4:23 Minuten lang (s. Abb. 4).



Abb. 4: Bildschirmfoto bei 0:05 auf <https://studyflix.de/deutsch/ballade-2916>

— Die Rezeption des Videos findet – in der Regel – allein vor dem Endgerät statt. Eine interaktive Situation, ähnlich des Unterrichts, ist weder im engeren Sinne möglich, noch wird sie im konkreten Fall von den Macher\*innen des Videos angeboten – denkbar wären beispielsweise Übungsaufgaben mit Feedback-Funktion gewesen; lediglich die Möglichkeit, im Kommentarbereich Rückmeldung zu geben, oder auf die anderen Videos zu klicken, stellen Handlungsmöglichkeiten der Rezipierenden dar (neben dem grundlegend möglichen Starten/Anhalten des Videos und dem Ansteuern bestimmter Zeitpunkte („Vor-/Zurückspulen“)).

— Die Erklärung ist dekontextualisiert, also aus einer konkreten Situation enthoben. Gerade diese Dekontextualisierung macht das wiederholte (orts- und zeitunabhängige, aber eben nicht Medien-/Geräte unabhängige) Rezipieren möglich.

— Die Erklärsequenz ist verdichtet. Anders als im Unterrichtsbeispiel oben, in dem die Erklärsequenz eine gesamte Schulstunde (45 Minuten) umfasst, sind die Erklärungen hier auf einen kurzen Zeitraum konzentriert (in der Regel 3-5 Minuten).

Uns interessieren – auch in der Folge – vor allem die multimodalen und multimedialen Elemente im Zusammenspiel, mit denen die Erklärung konstituiert wird (s. Abb. 5).

Bei dem Beispiel lässt sich das zunächst grob und an einem Ausschnitt (Still) folgendermaßen beschreiben: Umgesetzt auf einer Farbfläche im Hintergrund (lila) werden Beispielsätze aus Goethes *Erkönig* (als typische Ballade) eingefahren und je nach Fokus eingefärbt (z.B. im Hinblick auf das Metrum und Reimschema). Damit wird die Aufmerksamkeit gezielt gelenkt und fokussiert.

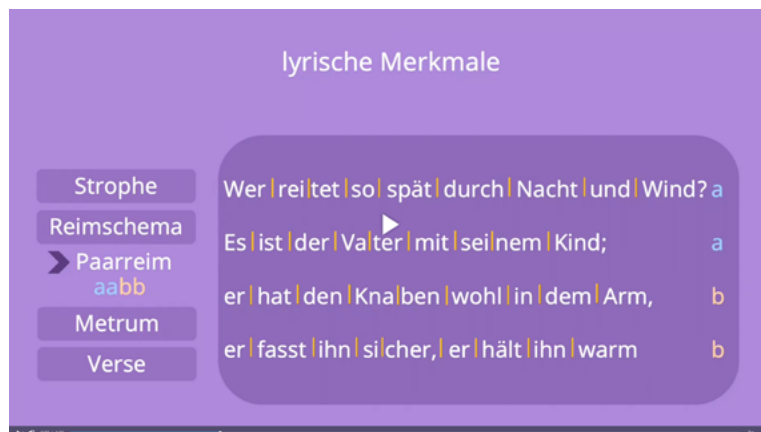


Abb. 5: Bildschirmfoto bei 0:58 auf <https://studyflix.de/deutsch/ballade-2916>

Eine besondere Rolle spielen die Farbflächen. Das Video ist in zwei Bereiche unterteilt, links die Kommentarspalte, mittig die Textspalte; jeweils durch einen dunkleren Lilaton hervorgehoben und überschrieben mit „lyrische Merkmale“. Die Kommentarspalte ist am Text ausgerichtet und wird um einzelne Elemente auf- und zugeklappt (im Still das Beispiel Paarreim für das Reimschema der Ballade).

Zwischen den Bereichen gibt es inhaltliche und formale Bezüge, hier die Erläuterung des Paarreims über die Buchstaben bzw. Buchstabenfolge aabb, die jeweils eingefärbt sind. Anders als multimodale Texte sind Erklärvideos allerdings filmische Sequenzen, die – teils im Sekundentakt – dynamische Bilder kreieren; die Möglichkeit zur Rekonstruktion von Bedeutung wird oftmals über das gleichzeitige (oder zeitlich versetzte) Auftreten von Elementen eröffnet. Für die Analyse sind u.E. diese Sequenzialität und Koinzidenz entscheidend, wir werden dies im Folgenden wieder aufgreifen.



## 2.3 ERKLÄREN UND ERKLÄRMUSTER

Erklärvideos haben nicht nur eine große (und zunehmend größere) Reichweite, die Gattung scheint sich auch sukzessive zu konsolidieren bzw. zu professionalisieren. *Lehrerschmidt* beispielsweise, Schulleiter aus Uelzen, ist inzwischen zu einer wiedererkennbaren Marke geworden (mit einem Logo, dem fast immer getragenen roten Poloshirt, der Selbstreferenz als Lehrerschmidt (in einem Wort), den wiederkehrenden sprachlichen Prozeduren, z.B. bei der Begrüßung etc.). Er hat – Stand Januar 2023 – inzwischen mehr als 1.500 Videos hochgeladen, die über 180 Millionen Aufrufe haben, sein Kanal verzeichnet 1,53 Millionen Abonnent\*innen. All dies verwaltet er sowohl auf seinem eigenen YouTube-Kanal als auch auf einer umfangreichen Webseite. Lehrerschmidt bedient verschiedene Fächer, besonders genutzt werden seine Mathematikvideos. Die Top drei Videos (entsprechend der Abrufzahlen) zeigt Abb. 6.

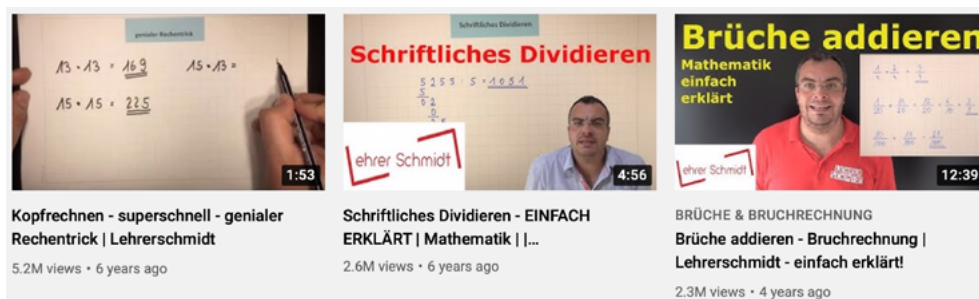


Abb. 6: Bildschirmfotos der am häufigsten aufgerufenen Videos auf <https://www.youtube.com/@lehrerschmidt/videos>

Das sind nicht nur erstaunliche Zahlen für Mathematikvideos, das zeigt auch, dass sich hier ein Format etabliert hat, das einen entsprechend großen Marktanteil hat und mit dem sich Geld verdienen lässt; durchschnittlich werden 1-2 Euro Vergütung für 1.000 Aufrufe gezahlt. Für die insgesamt 209.953.709 Abrufe (Stand 19.01.2023) hätte Lehrerschmidt dann bereits zwischen 209.000 und 420.000 Euro verdient.

Fachliche Spitzenreiter im Erklärvideosegment stellen Mathematik und Naturwissenschaften dar, aber auch das Angebot an Erklärvideos für das Fach Deutsch (auch Deutsch als Fremdsprache) ist breit. Wenngleich die Vielfalt an Gestaltungsmöglichkeiten die Beschreibung als Gattung erschwert, so zeigen sich doch verschiedene Prototypen, zwei möchten wir kurz vorstellen, bevor wir uns genauer den einzelnen Elementen des Erklärvideos widmen.

Den ersten Prototyp nennen wir Personenzentrierung, hier steht die erklärende Person im Fokus, beispielsweise als abgefilmte Personen in natura. Teils fungiert der Hintergrund noch als Tafel mit typographisch gestalteter Schrift (Beispiel in Abb. 7), teils wird darauf ganz verzichtet und die Schrift kommt nur noch am Bildschirmrand vor (Beispiel in Abb. 8).



Abb. 7: Personenzentrierung und gestaltete Schrift auf ‚Tafel‘ in der Erklärung des grammatischen Geschlechts im Deutschen von Lingster Academy (<https://youtu.be/5Vj7giLAiHk?t=46>)



Abb. 8: Personenzentrierung und gestaltete Schrift in Einblendung in der Erklärung der Konjunktion *dass* von Bob Blume ([https://youtu.be/E5RV\\_8pbppw?t=111](https://youtu.be/E5RV_8pbppw?t=111))

Bob Blume, der Produzent des in Abbildung 8 abgebildeten Videos, ist in den Sozialen Medien, zunehmend aber auch in Print- und Hörfunkmedien u.a. unter seinem Twitternamen *Der Netzlehrer* sehr aktiv und bereitet auch Materialien und Statements auf (zuletzt erschienen: *10 Dinge, die ich an der Schule hasse* (Blume 2022)); fast immer taucht er, wiedererkennbar, mit Foto auf seinen Veröffentlichungen auf.

Weniger präsent ist Julia von Lingster Academy, in den Kommentaren nutzt sie aber ihren Namen und ihre Wiedererkennbarkeit. Das Video (siehe Abb. 7) hat bereits über 1.4 Millionen Aufrufe.

Filmanalytisch würde man diese Einstellungen als *Nahe* bzw. *Nahe Einstellung* (auch *medium close up*) bezeichnen (vgl. Monaco 2009, 176-214). Die spezifische Darstellungsform korrespondiert im obigen Fall auch mit der Produktionspraxis. Bob Blume, der zu den sprechfreudigsten Entwicklern von Erklärvideos gehört, äußert sich auf diversen Kanälen sehr umfangreich zu seinem eigenen Tun. Zusätzlich war er bereit in einem Interview mit uns Auskunft zu geben (ausführlich in Knopp / Schindler i.Vorb.). Er produziert Erklärvideos zunächst immer mit Blick auf seine Schüler\*innen und rekonstruiert mit dem Video eine Erklärsituation im Unterricht (eine erklärende Person im Vordergrund, Mündlichkeit, der Text (Schrift) ist der mündlichen Erklärsequenz unter- bzw. nachgeordnet).

Ähnlich spontan gelingt aus seiner Erfahrung auch das Video, es benötigt wenig Vorbereitung, allerdings entsprechend Nachbereitungszeit. Konkret nach den Produktionsbedingungen befragt: Die Produktion der Erklärvideos erfolgt entsprechend ‚spontan‘ und nächsprachlich; ein Skript verwendet er nicht. Für ein Video von 5-8 Minuten benötigt er 11 Minuten Vorbereitungszeit. Die technische Nachbearbeitung dauert in der Regel 1,5 Stunden (Interview mit Bob Blume, 17.02.2022).

Die beiden Beispiele in Abbildung 9 und 10, die ohne sichtbare (reale) Personen oder (Comic-)Figuren, Avatare o.Ä. operieren, wollen wir *Bildflächenvideos* nennen.

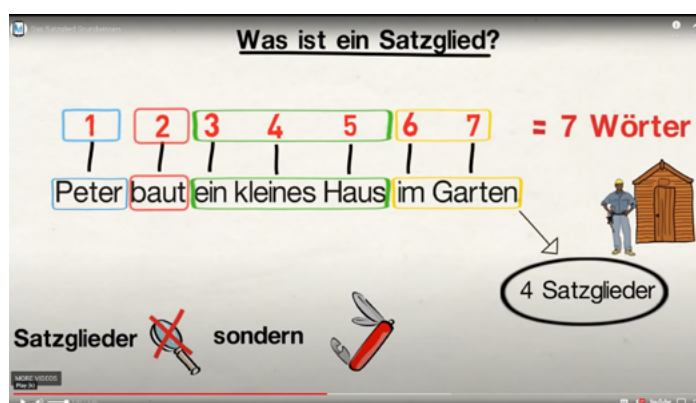


Abb. 9: Beispiel 1 für ein Bildflächenvideo, Bildschirmfotoquelle: Ivi-education ([https://youtu.be/JQXe0\\_Mo9L8?t=102](https://youtu.be/JQXe0_Mo9L8?t=102))

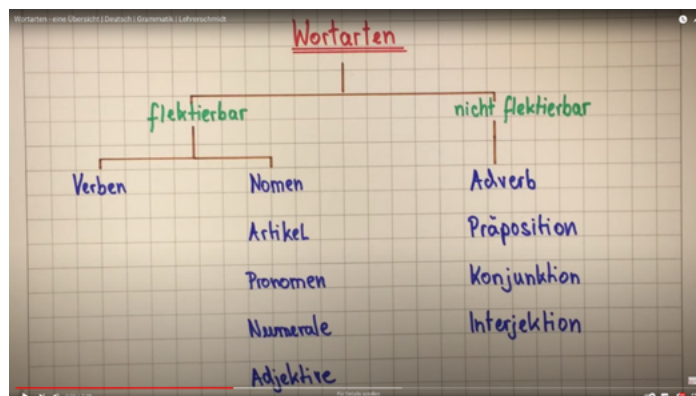


Abb. 10: Beispiel 2 für ein Bildflächenvideo, Bildschirmfotoquelle: Lehrerschmidt (<https://youtu.be/eBEeFYQ6YIE?t=50>)

Sie unterscheiden sich von der Personenzentrierung bereits auf den ersten Blick deutlich. Die Bildfläche ist gänzlich reserviert für die visuelle Darstellung der Erklärsequenz. Die erklärende Person tritt in den (visuellen) Hintergrund und ist nur noch mit der Stimme und teils mit einer schreibenden oder zeigenden Hand präsent.

Produzierende dieser Videos sind beispielsweise Marcus von Amsberg, ein Lehrer aus Hamburg, der das Konzept *Ivi-education* entwickelt hat, sowie Lehrerschmidt aus Uelzen, der bereits eingeführt ist. Lehrerschmidt nutzt auch für seine Grammatikvideos gerne eine karierte Fläche, was vielleicht daran liegt, dass er Mathematiklehrer ist.

Marcus von Amsbergs Erklärvideo repräsentiert nicht nur einen anderen Typ, er geht bei der Erstellung auch anders vor. Wir haben ihn ebenfalls im Februar 2022 zu seiner Arbeit befragt (vgl. Knopp / Schindler i.Vorb.). Ihm ist es wichtig, dass die Erklärung im Fokus steht und nichts von der Erklärung ablenkt, dadurch soll das Video universell einsetzbar sein. Seine Erklärvideos sind umfangreich vorbereitet, die Produktion orientiert sich an einem Skript, das aus seiner Sicht den zeitaufwändigsten und „nervigsten“ Produktionsschritt darstellt. Das Erklärvideo ist dann auch dem mündlichen Erklären überlegen, weil es die Gelegenheit bietet, den Erklärgegenstand umfassend zu durchdringen (Interview mit Marcus von Amsberg, 16.02.2022).

Die bislang genutzte Unterscheidung zwischen der mündlichen Erklärsituation face-to-face und der mündlich/schriftlichen Erklärsituation via Ausgabegerät ist bislang nur skizziert worden und erfasst die Komplexität des Gegenstands noch kaum. Diese zu durchdringen, erscheint uns nicht nur aus fachlicher Perspektive (im Sinne unserer Verortung in der Angewandten Linguistik), sondern auch aus fachdidaktischer Perspektive (hier unserer Verortung in der Deutschdidaktik) notwendig. Dazu benötigen wir aber ein theoretisches Beschreibungsinstrument, das über die Dimension Text hinausgeht und die Multimodalität des Erklärens im virtuellen Raum und in seiner Zeitlichkeit berücksichtigt.

### 3 — MULTIMODALES ERKLÄREN

#### 3.1 KOMMUNIKATIVE PRAKTIKEN IM DIGITALEN RAUM

Erklärvideos sind gegenwärtig i.d.R. mittels digitaler Apparate hergestellte Kommunikate, die eine spezifische Ausformung einer *kommunikativen Praktik* (vgl. Deppermann / Feilke / Linke 2016) darstellen. Der digitale Charakter dieser Kommunikate ist wesentlich, denn Aspekte der Digitalität schreiben sich

— in das Kommunikat (und z.B. entsprechende Anschlusskommunikation) ein und sind

— dieses Kommunikat nachgerade erst ermöglichend (so hat man die Rezeptions- und insbesondere Produktionsmittel für ein Erklärvideo gewissermaßen immer in der Gesäßtasche (vgl. Schmitz 2011, 25) dabei) und bilden

— ein Dispositiv, d.h. die Ermöglichung bestimmter Möglichkeiten, was wiederum Auswirkungen auf den Prozess der Produktion, Rezeption, Distribution und weitere Aspekte hat, die u.E. gut mit den Überlegungen von Stalder zu einer *Kultur der Digitalität* (2021) mit den drei Formen *Referentialität*, *Gemeinschaftlichkeit* und *Algorithmizität* zu fassen sind.

Als zentrale Form der Kultur der Digitalität erachtet Stalder Referentialität (vgl. ebd., 96f.). Unsere kulturellen Praktiken implizieren das Bezugnehmen auf Produkte und Prozesse der Kommunikation anderer (sowie Verfahren der Herstellung kultureller Artefakte). Digitalität – i.S.v. digitalen Medien und durch sie konstituierte Kommunikationsformen und kommunikative Praktiken – ermöglicht das besonders niederschwellige Bezugnehmen, Aneignen, Remixen/Sampeln, Zitieren, Neuarrangieren usw. Diese Praktiken an sich sind nicht neu – neu sind

aber Nutzungsfrequenz, Bandbreite, Einfachheit der Herstellung und Reichweite. Bezogen auf die Praktik *Digitales Erklärvideo* stellen sich hier z.B. Fragen nach der Autorschaft bei kollaborativ produzierten Erklärvideos, dem Verweisen auf Referenzsubjekte/-objekte innerhalb und außerhalb des Videos oder der Teilhabe an kulturellen Prozessen qua publik gemachtem Erklärvideo. Geteilte soziale Bedeutung vollzieht sich hier im ‚größeren Rahmen‘, z.B. von YouTube – auf im Video nicht Erklärtes etwa wird am Ende des Videos z.B. mittels klickbarer URL/Videostill hingewiesen.

Unter *Gemeinschaftlichkeit* fasst Stalder das Herstellen neuer gemeinschaftlicher Bezugsrahmen („vernetzter Individualismus“, ebd., 143) – am Beispiel Erklärvideo zeigt sich dies gut in den Kommentaren und der Möglichkeit, die Videos mittels der Daumen-hoch/runter-Funktion zu bewerten oder Themenwünsche gegenüber den Videoproduzierenden im Community-Bereich zu äußern. Die Anzahl der positiv/negativ bewertenden Daumen und die Zahl der Aufrufe resp. Nutzer\*innen zeigt die Reichweite – und damit ein Stück weit auch die Qualität – eines Videos an, wir hatten das bei Lehrerschmidt bereits erkennen können. Zugleich werden solche Videos auch wieder schneller von Suchmaschinen gefunden, die hohe Verbreitung schafft also selbst wieder Bedingungen dafür, dass die Sichtbarkeit und damit Verbreitung steigt.

Die durch Digitalität und Digitalisierung anfallenden, erheblich großen Datenmengen lassen sich ohne automatisierte Sortierung nicht bewältigen, *Algorithmizität* – als von Stalder genannte dritte Form – erzeugt dynamische Ordnungen für unterschiedliche dynamische Bereiche (vgl. ebd., 185), z.B. geschieht dies bei Suchanfragen in z.B. Google. „Die Welt wird nicht mehr repräsentiert; sie wird für jeden User eigens generiert und anschließend repräsentiert“ (ebd., 189). Bei Erklärvideos zeigt sich dies beispielsweise, indem Vorschläge für Anschlussvideos, ähnliche Themen bzw. Themen, die zu den bisherigen Suchanfragen passen, gemacht werden.

Kommunikative Praktiken fußen i.d.R. auf der Nutzung von Symbolen/Zeichen, Erklärvideos lassen sich (u.a.) beschreiben als dynamische komplexe (sprachliche) Zeichen, die zwar nicht ausschließlich durch Digitalität determiniert, aber motiviert und restringiert sind (vgl. Androutsopoulos 2006) – und entsprechend unterschiedliche Ausformungen kommunikativer Praktiken ermöglichen. Auf die Spezifik von (sprachlichen) Zeichen im digitalen Raum fokussieren wir im folgenden Kapitel.

### 3.2 ZEICHEN IM DIGITALEN RAUM

Mit Schneider (2017) betrachten wir *Medien* als Verfahren, (sprachliche) Zeichen herzustellen/zu verwenden/zu konstituieren/zu distribuieren/zu rezipieren/zu verändern; sie fußen aber zugleich auch immer auf einer materiellen Grundlage: So sind etwa Zeichen im digitalen Raum ohne ein entsprechendes digitales Ausgabegerät nicht rezipierbar (z.B. dem Computer als Medium). Hierbei werden Zeichen im digitalen Raum ‚fluide‘, die (technische) Niederschwelligkeit ihrer Herstellung, zum Re-

ferenzieren, zur Schaffung von Gemeinschaftlichkeit usw. (s. Kapitel 3.1) führt zu einem hohen Grad an Interaktion und Dynamik. Dabei erfolgt häufig eine Synthese mehrerer Zeichenmodalitäten<sup>4</sup>, z.B. wird in Bildern Text oder Text in Bilder integriert (wie z.B. bei Memes), in Videos Schrift (wie z.B. bei Erklärvideos) oder gesprochene Sprache (z.B. bei TikTok-Videos). Diese Synthese spielt in Erklärvideos u.E. eine substantielle Rolle.

Letzteres zeigt sich aber bereits bei (Alltags-)Texten (vgl. Pappert / Roth 2021), die zwar multimodal sind, die wir aber i.d.R. in ihrer holistischen Gestalt wahrnehmen. Text und Bild beziehen sich aufeinander, beide füllen Flächen und beide verlangen unterschiedlich geartete Abstraktionsleistungen. Schmitz beschreibt dieses Phänomen als Sehfläche (Schmitz 2011). Ihn interessieren dabei vor allem die semiotischen Synergien, die durch Text und Bild erzeugt werden. An einem Beispiel lässt sich Schmitz' Konzept der Sehfläche gut veranschaulichen (s. Abb. 11).



Abb. 11: Vortragsankündigung ZWISCHEN DEN IMPERIEN

Das Plakat (Abb. 11) repräsentiert ein typisches Beispiel einer Vortragsankündigung, wie sie digital auf der Webseite eines wissenschaftlichen Instituts oder ausgedruckt im Hochschulkontext platziert wird.

Der Text beinhaltet für die potenziellen Adressat\*innen, Teilnehmer\*innen des Vortrags, relevante Informationen zur Uhrzeit, zum Veranstalter, zum Titel und Vortragenden, wie zum Veranstaltungsort. Der Text ist zentriert und rechtsbündig ausgerichtet, er ist von oben nach unten zu lesen. Die Schriftgröße markiert die als zentral herausgestellten Informationen (Titel „Zwischen den Imperien“).

<sup>4</sup> Modalität verstehen wir hier mit Stöckl (2016) vergleichsweise umfassend. Neben dem psychologischen Verständnis von Zeichen i.S.v. Zeichentyp/Wahrnehmungskanal (also visuell, auditiv, taktil, olfaktorisch, gustativ) umfasst dieses Verständnis Aspekte von Kodiertheit – was insbesondere für sprachliche Zeichen relevant ist: Welcher Code/welches Zeichensystem wird genutzt? – und Medialität der Zeichen – was insbesondere für Zeichen im digitalen Raum konstitutiv ist: Welche medialen Möglichkeiten der Darstellung werden z.B. eingesetzt? (s. dazu auch hier S. 6)

Der Text ist zudem platziert, denn im Hintergrund ist eine gestaltete und abstrahierte Landkarte zu erkennen, in der die sprachlich ausgewiesenen Länder bzw. die Landesgrenzen abgebildet sind (Ukraine und Baltikum). Interessant ist die sprachliche und bildliche Markierung der Grenze bzw. des Dazwischens. Diese Länder – und das verrät bereits der Titel – befinden sich *zwischen* verschiedenen Imperien. Der Hintergrund korrespondiert also mit dem Haupttitel und verstärkt die sprachliche Aussage.

Auf der Fläche der Länder sind Briefmarken und Geldscheine abgebildet, auf denen Köpfe abgedruckt sind. Diese Köpfe repräsentieren wichtige Nationalautor\*innen der jeweiligen Länder, ihre Texte stellen „Schulklassiker“ dar, wie auch im Untertitel vermerkt wird. Die beiden Zeichensysteme agieren hier komplementär: Sprache als Zeichensystem ist in der Lage, abstrakte Konzepte zu benennen, Bild als Zeichensystem kann diese in holistischer Weise veranschaulichen.

Schmitz ist einer der ersten Linguisten, der die besondere Relevanz von Bildern für sprachliches Verstehen systematisch herausgearbeitet hat (allerdings an statischen, d.h. unbewegten, Sehflächen) und dabei das Zusammenwirken mit Schrift/Text dezidiert berücksichtigt hat. Für eine Analyse von Erklärvideos sind seine Überlegungen von großem Wert, denn

Bild und Text wirken nicht je für sich allein, sondern stecken in ihren Botschaften einander an zu einer simultan und auf einmal wahrgenommenen holistischen Gesamtbotschaft. Niemand (außer uns hier bei der analytischen Detail-Lektüre) zerlegt sie in ihre Bestandteile. (Schmitz 2011, 33)

Dennoch genügt die Idee der statischen Sehfläche nicht, um die Komplexität des ‚bewegten‘ Kommunikats *Erklärvideo* zu erfassen, da neben die visuelle Dimension eine auditive Dimension (Sprache, Musik, Geräusche) tritt und insbesondere das Kommunikat dynamisch ist – das Auftreten von Elementen unterschiedlicher Modalität ist konzertiert auf der Zeitachse (= vierte Dimension, Zeit). Das Auftreten von Schrift (unberücksichtigt in Holly 2004, der das Zusammenwirken von Sprechsprache und bewegten Bildern fokussiert) erscheint uns dabei als konstitutiv für das Gros der Erklärvideos.

#### 4 — EXEMPLARISCHE ANALYSE & DARSTELLUNG IM MODELL

Im Erklärvideo spielen also unterschiedliche Modalitäten zusammen. Wir wollen dies exemplarisch an einem Bildflächenvideo konkretisieren und nutzen dazu zunächst den Still des Videos (s. Abb. 12). Wir unterscheiden bei der Analyse zwischen der visuellen Dimension – dem Gezeigten; und der auditiven Dimension – dem Gehörten.

Das ausgewählte Video ist vom Unternehmen *Learning Level Up* produziert, das Erklärvideos zu unterschiedlichen Themen und Fächern anbietet (s. <https://learning-levelup.education/>). Während einzelne der Videos zunächst kostenfrei auf YouTube zur Verfügung stehen, verbleibt das Nutzungsrecht beim Unternehmen. Schulen können Einzelplatz- oder Schullizenzen erwerben, um die Videos (und weitere Lernmaterialien) im Schulkontext zu verwenden. Wir haben dieses Video ausgewählt, da es vergleichsweise schlicht gehalten ist und einen besonders klaren Fall darstellt, was die Nutzung typographischer und gestalterischer Mittel (und der Agilität) angeht.

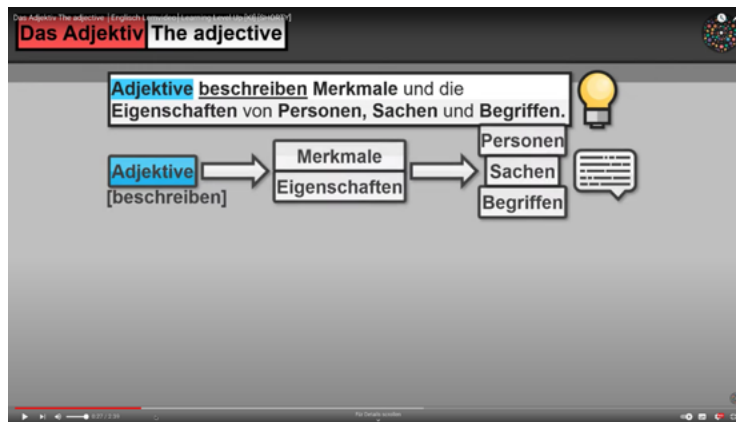


Abb. 12: Still aus dem Lernvideo *Das Adjektiv | The adjective* (<https://youtu.be/WXwDnpJ6nWA?t=26>)

Analog zu Schmitz' Konzept der Sehfläche beschreiben wir zunächst die visuelle Dimension: Auf den ersten Blick lassen sich bereits eine Vielzahl bestehender Elemente unterscheiden, die in spezifischer Weise angeordnet sind, so dass sie sich auf der Sehfläche sinnfällig verteilen. Wir haben es hier mit einem „Operationsraum Schrift“ nach Krämer (2005) zu tun.

Die Fläche als Ganzes ist in drei Bereiche gegliedert: die Überschrift, das Textfeld und die Navigationsleiste (zuweilen automatisch ausgeblendet).

— Innerhalb des Textfeldes zeigt sich die weitere Architektur in einzelnen Zeilenbändern: der Erklärung (Adjektive beschreiben) und der formelhaften Verdichtung (Adjektive ->). Entsprechend handelt es sich hier um eine Verdopplung, die einzig durch die gestalterische Form variiert, inhaltlich aber identisch ist.

— In der Überschrift findet sich der zweisprachige Titel des Videos (*Das Adjektiv | The adjective*) und das Logo des Anbieters (rechts oben farbig).

— Das Logo wird in der Navigationsleiste noch einmal aufgegriffen (links unten s/w).

Im Textfeld selbst finden sich neben Schrift verschiedene Symbole: die Glühbirne (hinter der Erklärung), die Sprechblase und die Pfeile. Alle Symbole sind stark konventionalisiert. Die Glühbirne steht in diesem Fall für eine (neue) Erkenntnis, die Sprechblase im engeren Sinne für eine mündlich getätigte Äußerung, im weiteren Sinne für kommunikative Entitäten insgesamt. Die Pfeile schließlich suggerieren eine Abfolge mit klar markierter Richtung. In der Linguistik, und das Video lässt sich in diesem Zusammenhang verorten, werden mit Pfeilen Referenzen zwischen Form- und Funktionsebene angegeben.

Die Fläche insgesamt ist in Grautönen gehalten. Die Schrift ist in einzelne Textkästen platziert, die schwarz umrandet und weiß gefüllt sind. Die Schrift ist ausschließlich im Fettdruck (in durchgehender Schriftart und -größe), es findet sich zudem das Mittel der Unterstreichung. Der durchgehende Fettdruck erscheint wenig funktional, da die Praxis der Auszeichnung bzw. Hervorhebung damit ad absurdum geführt wird, die Unterstreichung wird daher notwendig, um einzelne Lexeme hervorzuheben.



Deutlich weniger komplex und eher nachgeordnet ist die auditive Dimension gestaltet, die wir als besonders schlicht bzw. konzentriert wahrnehmen. Musik und Geräusche werden nicht verwendet. Es ist lediglich eine computeranimierte, als weiblich wahrzunehmende Sprechstimme zu hören, die dem Text verzögert nachfolgt. Pausen und andere stimmliche Merkmale sind unmarkiert. Betont werden die Inhaltswörter der kurzen Beschreibung (Adjektive, MERKmale, Eigenschaften, PerSONen, SACHen und BeGRIFfen).

Visuelle und auditive Dimension lassen sich zwar getrennt voneinander analysieren, im Erklärvideo agieren sie aber im Zusammenspiel, wir würden hier von Koinzidenz sprechen. Im Beispielvideo ist dies durchgehend in einem bestimmten Muster arrangiert: Die Schrift wird zunächst eingeführt, von links nach rechts einfahrend. Nach kurzer Verzögerung setzt die Sprechstimme ein, die den Text (oben) vorliest. Die Seite wird dann über eine Erweiterung der Definition (unten) aufgefüllt, wobei die Abfolge gleich bleibt. Dieses Muster ist vor allem dadurch geprägt, dass es gleichsam tautologisch agiert. Der Ausgangssatz „Adjektive **beschreiben Merkmale** und die **Eigenschaften** von **Personen, Sachen** und **Begriffen**“ (Hervorhebungen im Original, s. Abb. 13) wird dreifach (und in identischem Wortlaut) präsentiert: Als vollständiger Satz geschrieben, als Satzfragment über Symbole dargestellt (geschrieben) und als Äußerung (gesprochen).

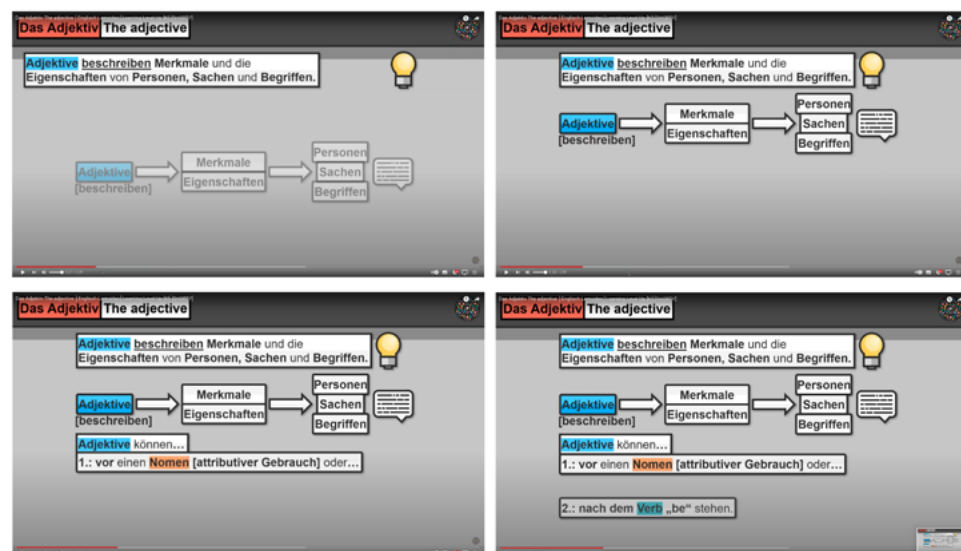


Abb. 13: Zeilenweise von links nach rechts: Einfahren und abnehmende Transparenz der einzelnen Textelemente vom linken Bildrand in die Kadrierung hinein (ohne Ton, Geräusch oder Musik); nach kurzer Verzögerung, fast zeitgleich mit der Stimme

Wir hatten bereits darauf hingewiesen, dass die Dynamik bzw. die Veränderung der bildlichen Ausgangssituation in bisherigen Modellen nicht überzeugend erfasst werden kann. Das Video stellt eben keine statische Darstellung resp. bloße Aneinanderreihung mehrerer statischer Bilder dar (Textbild), sondern ist eine filmisch gestaltete Sequenz, sie ist – filmanalytisch gesprochen – montiert bzw. eine *Montage*.

Die jeweiligen Elemente werden dazu in besonderer Weise, bewusst arrangiert. Neben die Bildgestaltung (visuelle Dimension) und die Tongestaltung (auditive Dimension) tritt die Montage (vgl. Rüssel 2010).

Diese Montage lässt sich nun chronologisch als Ablauf darstellen und dann ähnlich eines Skripts (im Film) bzw. Transkripts (mündlicher Interaktionen) lesen und dann für jedes Video konkret bestimmen. Um die dafür notwendigen Elemente aber zuzuordnen, erscheint es uns notwendig, zunächst eine Modellierung (s. Abb. 14) vorzuschlagen, die die drei Dimensionen im abstrakten Zusammenspiel unterscheidet. Das Modell lässt sich als ‚Grammatik der Text-Bild-Verknüpfungen‘ und der Koinzidenz dieser begreifen. Dabei fokussieren wir weniger auf Kognition, denn auf Kommunikation i.S. Hausendorfs et al. (2017, 8) – also auf die sinnlich wahrnehmbaren Erscheinungsformen und nicht die bei der Produktion und Rezeption ablaufenden kognitiven Prozesse. Erklärvideos lassen sich u.E. damit als dynamische, multimodale Texte begreifen, die aus einem Konglomerat von zusammenwirkenden Lesbarkeitshinweisen bestehen. Eng an Hausendorfs et al. (2017) angelehnt wäre demnach ein Erklärvideo nichts anderes als

das Gesamt seiner Lesbarkeitshinweise, die [Rezipierende] in einer konkreten [Rezeptions]situation und [vor] eine[m] konkreten [Rezeptionskontext] wahrnehmungs-, [verbal, ] schriftsprach- und vertrautheitsabhängig zur Geltung bringen und aktualisieren. (ebd., 22)

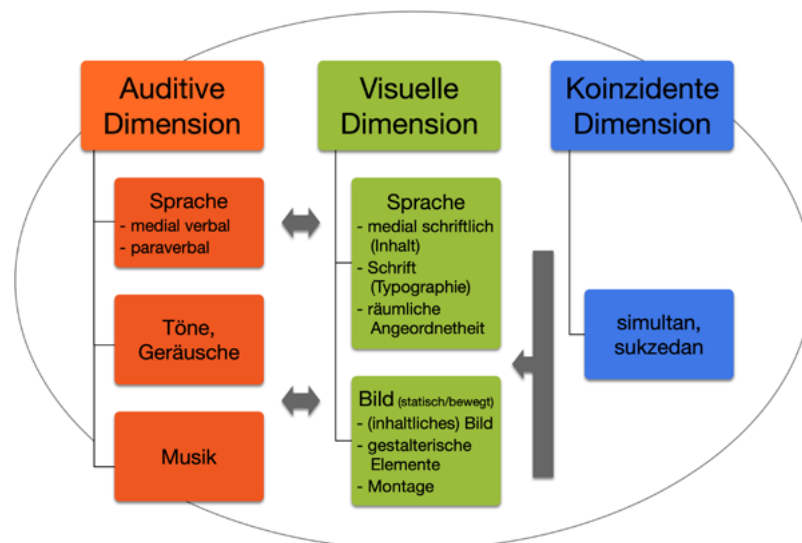


Abb. 14: Modell zur Darstellung des Zusammenwirkens multimodaler Elemente auf drei Dimensionen im Erklärvideo

Wahrgenommen, rezipiert, wird die Erklärung holistisch – so wie es auch in multimodalen Erklärkontexten der schulischen Interaktion der Fall ist, d.h. auditive und visuelle Dimension werden gleichzeitig erfahren – symbolisiert durch den Kreis, der alle Elemente umschließt.

Die bei der Produktion notwendige Anordnung einzelner Elemente ist hingegen zergliedert und muss sinnfällig zusammengestellt werden, wobei enge Bezüge herausgearbeitet werden (z.B. der mündlichen Äußerung, die auf die medial schriftliche Realisierung folgt etc.); symbolisiert durch die Pfeile im Modell. Wie das Zusammenspiel angelegt ist, wird durch die koinzidente Dimension beschrieben, die simultan oder sukzedan erfolgen und sich durchaus auch im Verlauf ändern kann.

## 5 — ZUR PRODUKTION VON ERKLÄRVIDEOS

Lehrkräfte und Schüler\*innen gehen, so haben wir eingangs entfaltet, inzwischen selbstverständlich mit Erklärvideos um. Erklärvideos dienen dazu, Inhalte einzuführen, zu vertiefen, nachzuarbeiten oder einzuüben. Neben diese Rezeptionspraxis tritt zunehmend auch die Forderung, Lehrende wie Lernende zu befähigen, selbst solche Videos zu produzieren. Lehrkräfte können damit gezielt Informationen für ihre Schüler\*innen aufbereiten und differenzieren (vgl. Schindler / von Amsberg 2023). Für die Lernenden lässt sich annehmen: Die Produktion von Erklärvideos setzt nun unterschiedliche Kompetenzen voraus bzw. ermöglicht die Entfaltung bestimmter Kompetenzen. In Bezug auf den Deutschunterricht ließen sich diese als fachliche, fachdidaktische und mediendidaktische Kompetenzen unterscheiden. Lernende müssen den Gegenstand kennen (z.B. ein orthographisches Phänomen) bzw. durchdringen, ihn in seiner Bedeutung und Relevanz einschätzen, und so didaktisch für andere aufbereiten, dass er im Rahmen eines Erklärvideos nachvollziehbar erklärbar wird. Letzteres setzt die Selektion und Montage von Informationen in unterschiedlichen Modalitäten voraus. Dazu ist es notwendig zu verstehen, welche unterschiedlichen Dimensionen wirken und wie sie zusammenspielen können. Unser Modell stellt dafür eine Grundlage dar. Damit dient es zugleich als Matrix, um ein Skript für ein Erklärvideo zu entwerfen und zu schreiben (vgl. Knopp / Schindler i.V.).

## QUELLENVERZEICHNIS SEKUNDÄRQUELLEN

- **Anders, Petra et al. (2019)**: Erklärvideo. In: Anders, Petra et al. (Hg.): *Einführung in die Filmdidaktik*. Stuttgart: J. B. Metzler, 255-268. — **Androutopoulos, Jannis (2006)**: Introduction: Sociolinguistics and computer-mediated communication. In: *Journal of Sociolinguistics*, Jg. 10, H. 4, 419-438. — **Blume, Bob (2022)**: *10 Dinge, die ich an der Schule hasse*. München: Mosaik. — **Cwielong, Ilona Andrea / Kommer, Sven (2020)**: „Wozu noch Schule, wenn es YouTube gibt?“ Warum eine scheiternde Didaktik neue Formen des selbstorganisierten und selbstbestimmten Lernens fördert. In: Kaspar, Kai et al. (Hg.): *Bildung, Schule, Digitalisierung*. Münster u.a.: Waxmann, 38-44. — **Deppermann, Arnulf / Feilke, Helmuth / Linke, Angelika (2016)**: Sprachliche und kommunikative Praktiken. Eine Annäherung aus linguistischer Sicht. In: Deppermann, Arnulf / Feilke, Helmuth / Linke, Angelika (Hg.): *Sprachliche und kommunikative Praktiken*. Berlin: de Gruyter, 1-24. DOI: <https://doi.org/10.1515/9783110451542-002>. — **Dube, Juliane (2021)**: Literaturunterricht über YouTube. Erklärvideos in heterogenen Lerngruppen. In: *MiDu*, Jg. 3, H. 1, 1-23. DOI: <https://doi.org/10.18716/OJS/MIDU/2021.1.7>. — **Dube, Juliane / Boerdgen, Yolanda (2022)**: Die Spreu vom Weizen trennen. Erklärvideos für den Literaturunterricht mit einem Qualitätsraster auswählen. In: *SLLD-U*. DOI: <https://doi.org/10.46586/SLLD.236>. — **Ebel, Alexandra (2021)**: Einfluss sprachlicher und sprecherischer Merkmale auf die Verständlichkeit von Lernvideos auf YouTube. In: Staubach, Katharina (Hg.): *Multimodale Kommunikation in den Hypermedien und Deutschunterricht. Theoretische, empirische und unterrichtspraktische Zugänge*. Baltmannsweiler: Schneider Hohengehren, 250-275. — **Feilke, Helmuth / Rezat, Sara (2021)**: Erklärttexte lesen und schreiben. In: *Praxis Deutsch*, H. 285, 4-13. — **Findeisen, Stefanie / Horn, Sebastian / Seifried, Jürgen (2019)**: Lernen durch Videos – Empirische Befunde zur Gestaltung von Erklärvideos. In: *Medienpädagogik. Zeitschrift für Theorie und Praxis der Medienbildung*, 16-36. DOI: <https://doi.org/10.21240/mpaed/00/2019.10.01.X>. — **Frei, Mario et al. (2020)**: Die Wirksamkeit von Erklärvideos im Rahmen der Methode Flipped Classroom. In: Kaspar, Kai et al. (Hg.): *Bildung, Schule, Digitalisierung*. Münster u.a.: Waxmann, 284-290. — **Habscheid, Stephan (2005)**: Das Internet – ein Massenmedium? In: Siever, Torsten / Peter Schlobinski, Peter / Runkehl, Jens (Hg.): *Websprache.net. Sprache und Kommunikation im Internet*. Berlin u.a.: de Gruyter, 46-66. — **Hackl, Stefan (2021)**: „How to Deutsch“. Zur Problematik von YouTube-Erklärvideos für das Unterrichtsfach Deutsch. In: Krammer, Stefan / Leichtfried, Matthias / Pissarek, Markus (Hg.): *Deutschunterricht im Zeichen der Digitalisierung*. Innsbruck: StudienVerlag, 116-129. — **Hausendorf, Heiko et al. (2017)**: *Textkommunikation. Ein textlinguistischer Neuanatz zur Theorie und Empirie der Kommunikation mit und durch Text*. Berlin u.a.: de Gruyter. — **Heller, Vivien (2021)**: Die sprachlich-diskursive Darstellung komplexer Zusammenhänge im Fach: Vertextungsverfahren des Erklärens und Argumentierens. In: Quasthoff, Ute / Heller, Vivien / Morek, Miriam (Hg.): *Diskurserwerb in Familie, Peergroup und Unterricht. Passungen und Teilhabechancen*. Berlin: de Gruyter, 303-346. — **Holly, Werner (2004)**: Sprechsprache und bewegte Bilder: Audiovisualität. In: *Mitteilungen des Deutschen Germanistenverbandes*, Jg. 51, H. 2, 122-135. — **Klein, Josef (2001)**: Erklären und Argumentieren als interaktive Gesprächsstrukturen. In: Brinker, Klaus et al. (Hg.): *Text- und Gesprächslinguistik. Ein internationales Handbuch zeitgenössischer Forschung*, 2. Halbband. Berlin u.a.: de Gruyter, 1309-1329. — **Klein, Josef (2009)**: ERKLÄREN-WAS, ERKLÄREN-WIE, ERKLÄREN-WARUM. Typologie und Komplexität zentraler Aspekte der Welterschließung. In: Vogt, Rüdiger (Hg.): *Erklären. Gesprächsanalytische und fachdidaktische Perspektiven*. Tübingen: Stauffenburg, 25-36. — **Klug, Nina-Maria / Stöckl, Hartmut (2016)**: Einleitung. In: Klug, Nina-Maria / Stöckl, Hartmut (Hg.): *Handbuch Sprache im multimodalen Kontext*. Berlin u.a.: de Gruyter, VII-XIII. — **Knopp, Matthias / Schindler, Kirsten (2021)**: Multimodales Erklären im Deutschunterricht. In: *k:ON - Kölner Online Journal für Lehrer\*innenbildung, Sonderausgabe 1 (2021)*, 51-79. DOI: <https://doi.org/10.18716/ojs/kON/2021.s.4>. — **Krämer, Sybille (2005)**: ‚Operationsraum Schrift‘. Über einen Perspektivwechsel in der Betrachtung der Schrift. In: Grube, Gernot / Kogge, Werner / Krämer, Sybille (Hg.): *Schrift. Kulturtechnik zwischen Auge, Hand und Maschine*. München: Wilhelm Fink, 23-61. — **Medienpädagogischer Forschungsverbund Südwest (Hg.) (2021)**: *JIM-Studie 2021. Jugend, Information, Medien. Basisuntersuchung zum Medienumgang 12- bis 19-Jähriger*. Stuttgart. [https://www.mpfs.de/fileadmin/files/Studien/JIM/2021/JIM-Studie\\_2021\\_barrierefrei.pdf](https://www.mpfs.de/fileadmin/files/Studien/JIM/2021/JIM-Studie_2021_barrierefrei.pdf) — **Monaco, James (2009)**: *Film verstehen. Kunst, Technik, Sprache, Geschichte und Theorie des Films und der Neuen Medien*. Reinbek bei Hamburg: Rowohlt. — **Neumeister, Nicole / Vogt, Rüdiger (2009)**: Erklären im Unterricht. In: Becker-Mrotzek, Michael (Hg.): *Mündliche Kommunikation und Gesprächsdidaktik*. Baltmannsweiler: Schneider Hohengehren, 562-583. — **Normann, Jana (2020)**: Komma klar mit Youtube! Chancen und Risiken von Lernvideos. In: *Praxis Deutsch*, H. 283, 22-27. — **Pappert, Steffen / Roth, Kersten Sven (Hg.) (2021)**: *Kleine Texte*. Berlin: Peter Lang. — **Pertzel, Eva (2021)**: Zur Produktion von Erklärvideos im Deutschunterricht – Oder: Der lange Weg vom KMK-Beschluss „Bildung in der digitalen Welt“ zur passenden Stunde im Klassenzimmer. In: Staubach, Katharina (Hg.): *Multimodale Kommunikation in den Hypermedien und Deutschunterricht. Theoretische, empirische und unterrichtspraktische Zugänge*. Baltmannsweiler: Schneider Hohengehren, 250-275.

*praktische Zugänge*. Baltmannsweiler: Schneider Hohengehren, 327-349. — **Rat für kulturelle Bildung (Hg.) (2019)**: *Jugend/YouTube/Kulturelle Bildung – Horizont 2019: Studie: eine repräsentative Umfrage unter 12- bis 19-jährigen zur Nutzung kultureller Bildungsangebote an digitalen Kulturorten*. Essen: Rat für Kulturelle Bildung.

— **Rüsel, Manfred (2010)**: Die Grammatik der Bilder. In: Frizen, Werner / Jansen, Peter (Hg.): *Abi-Box Deutsch – Die Basis-Box für die Einführungsphase der Oberstufe (Schülerarbeitsbuch und Lehrermappe)*. Hannover: Brinkmann Meyhöfer, 68-91. — **Schindler, Kirsten / Knopp, Matthias (in Begutachtung)**: Erklärvideos planen und gestalten – Wissen und Überzeugungen aus Perspektive der Erklärenden. In: Arndt, Lara et al. (Hg.): *Digitale Clips und Lernvideos im Deutschunterricht analysieren, produzieren, nutzen und bewerten. Untersuchungen zu YouTube, VIMEO & Co. aus sprach-, literatur- und mediendidaktischer Perspektive*. SLLD-B. — **Schindler, Kirsten / von Amsberg, Marcus (2023)**: Wir müssen um die Schüler\*innen kämpfen. Das heißt, hier ist ein Innovationsdruck“ – Marcus von Amsberg im Gespräch mit Kirsten Schindler. In: *Der Deutschunterricht*, H. 4 (2023).

— **Schmitz, Ulrich (2011)**: Sehflächenforschung. Eine Einführung. In: Hajo Diekmannshenke, Hajo / Klemm, Michael (Hg.): *Bildlinguistik. Theorien – Methoden – Fallbeispiele*. Berlin: Erich Schmidt, 23-42. — **Schöne, Justine / Wedler, Katharina (2021)**: Erklärvideos in der LehrerInnenausbildung: Textanalytische und produktionsorientierte Zugänge. In: Staubach, Katharina (Hg.): *Multimodale Kommunikation in den Hypermedien und Deutschunterricht. Theoretische, empirische und unterrichtspraktische Zugänge*. Baltmannsweiler: Schneider Hohengehren, 276-302. — **Schneider, Jan Georg (2017)**: Medien als Verfahren der Zeichenprozessierung. Grundsätzliche Überlegungen zum Medienbegriff und ihre Relevanz für die Gesprächsforschung. In: *Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion*, Jg. 18, 34-55. — **Stalder, Felix (2021)**: *Kultur der Digitalität*. Berlin: Suhrkamp.

— **Stöckl, Hartmut (2016)**: Multimodalität – Semiotische und textlinguistische Grundlagen. In: Klug, Nina-Maria / Stöckl, Hartmut (Hg.): *Handbuch Sprache im multimodalen Kontext*. Berlin u.a.: de Gruyter, 3-35. — **Torrau, Sören (2021)**: Anderen und sich selbst die Welt erklären. Schülerinnen und Schüler analysieren Erklärvideos. In: *Praxis Deutsch*, H. 285, 45-53. — **Vogt, Rüdiger (Hg.) (2009)**: *Erklären. Gesprächsanalytische und fachdidaktische Perspektiven*. Tübingen: Stauffenburg.

— **von Brand, Tilman (2020)**: Youtube. In: *Praxis Deutsch*, H. 283, 4-13. — **Wolk, Thorsten (2018)**: Vom passiven Zuschauer zum aktiven Produzenten – Erklärvideos zu grammatikalischen Phänomenen mit der Zeige-Lege-Technik. In: *Deutschunterricht*, Jg. 71, H. 4, 47-53. — **Unruh, Thomas / Petersen Susanne (2021)**: *Guter Unterricht: Praxishandbuch Handwerkszeug für Unterrichts-Profis (Alle Klassenstufen)*. 17. Aufl. Hamburg: AOL.

## ÜBER DIE AUTOR:INNEN

**Prof.in Dr. phil. Kirsten Schindler**, Professorin für Deutschdidaktik (Schwerpunkt Sprachdidaktik) an der Bergischen Universität Wuppertal. aktuelle Arbeits- und Forschungsschwerpunkte: Geschlechterinklusive Deutschunterricht, Schreiben und künstliche Intelligenz, akademisches und berufsbezogenes Schreiben

**Dr. phil. Matthias Knopp** ist Akademischer Rat am Institut für deutsche Sprache und Literatur II an der Universität zu Köln. Seine Forschungs- und Arbeitsschwerpunkte sind u.a. Sprache und sprachliches Lernen mit und in digitalen Medien