

Werner Holly (Chemnitz)

„Beschwerte“ Unterhaltsamkeit in einem politischen Aufklärungsversuch. Schwarzeneggers audiovisueller Social-Media-Auftritt gegen Putin

Abstract: In my analysis of a piece of counter-propaganda, namely the social media video of Arnold Schwarzenegger against Putin, I argue that even in a very serious context such as the Ukraine war, we find elements of entertainment. Thus the overall principle of television and more so the emotionalised intimacy of social media called “Halböffentlichkeit” (semi-public) by Habermas allow a Hollywood-style of narrative introduction to convince the target audience of the speaker’s friendship and credibility. The multi-modal “zip”-structure (speech and voice, subtitles in two languages, pictures of speaker, setting, photographs, film scenes) offers a strictly conventional rhetoric procedure with all three elements of persuasion: *logos*, *ethos*, *pathos*. The hybrid text not only abandons the traditional separation of information and entertainment. It is also entertainment of a special, a dramatic kind. In his general model of entertainment Klein (1997) proposes “Unbeschwertheit” (light-heartedness) as an appropriate principle, analogous to truth in information-oriented communication (Grice). However, instead of “happy-go-lucky” feelings, in Schwarzenegger’s video, heavy emotions such as suffering and despair play a central role and lead the speaker to pathetic appeals for resistance.

Keywords: multimodality, social media, counter-propaganda, entertainment, light-heartedness vs. dramatic emotions, rhetoric

1 Unterhaltsamkeit in öffentlicher Kommunikation: Tendenz steigend

Als im 2. Weltkrieg der große Thomas Mann aus seinem amerikanischen Exil über die BBC Radioansprachen an die „Deutschen Hörer“ richtete, um als bekanntester und vielgerühmter deutscher Schriftsteller der Nazi-Propagandamaschine entgegenzuwirken, tat er dies im Stil der Zeit mit beben-der Stimme und mit allem Pathos, das der Ernst der Lage nach den Traditionen der klassischen Rhetorik nahelegte.¹ Ganz anders Arnold Schwarzenegger im

1 Die 55 Reden aus den Jahren 1941 bis 1945 sind sowohl verschriftet als auch als Tondokumente zugänglich (Mann 1975, Mann 1988/2003).

Jahre 2022.² Ein Hollywood-Star, ein ehemaliger Bodybuilder, der mit Action-Filmen ein weltweites Fan-Publikum erreicht hat, dann sogar als Gouverneur von Kalifornien eine politische Karriere machte, spricht mit Ton, Schrift und Bildern zu seinen „russischen Freunden“ in der lockeren Manier seiner Filmrollen, seinen österreichischen Akzent im Amerikanischen nicht verleugnend. Die auffällige Stildifferenz ist nicht nur unterschiedlichen Persönlichkeiten und Images zuzuschreiben, sie zeugt auch von einem neuerlichen „Strukturwandel der Öffentlichkeit“,³ den eine rasante Medienentwicklung in der Zwischenzeit hervorgebracht hat.

Diese Entwicklung ist seit den 1980er Jahren unübersehbar, als zunehmende Hybridbildung von Information und Unterhaltung – letztere als Superideologie des damals noch als Leitmedium geltenden Fernsehens (vgl. Postman 1985) – zu „Infotainment“ und „Politainment“ (Dörner 2001). Das Fernsehen als Zentrum einer audiovisuellen „Mediokratie“ (Meyer 2001) wird auch bei uns anstelle des Parlaments zur Arena der öffentlichen Meinungsbildung, gestaltet diese nun aber um, und zwar nach seinen inhärenten Prinzipien Theatralisierung, Inszenierung, Ästhetisierung (vgl. Meyer/Ontrup/Schicha 2000), Visualisierung, Boulevardisierung (vgl. Holly 2005; 2008), und wird so zu einem Phänomen der Populärkultur, getrieben von den Mechanismen der Werbeindustrie (vgl. Schwarzenberg 1980; Holly 1991). Was als „Inszenierung von Propaganda als Diskussion“ (Holly/Kühn/Püschel 1986; 1989) oder „Redeshows“ (ebd.) begann, mündete in Auftritte von Politikern in Unterhaltungsshow (vgl. Holly 1990), ritualisierte „TV-Duelle“ (Klein 1990) und „Confrontainment“ (Holly 1994a; Holly/Schwitalla 1995), bis hin zu den audiovisuell perfekt durchinszenierten Talkshows, die den Akteuren immer größere Teile ihrer Aktionsautonomie und -kontrolle entziehen (vgl. Holly 2012; 2018).

Das Aufkommen der Sozialen Medien hat dann noch deutlicher gemacht, dass im heftig umkämpften Markt der Aufmerksamkeit, auf dem jetzt potenziell alle mitspielen können, jedes Mittel recht zu sein scheint, das schnell Aufregung erzeugt: Alles, was laut, wenn nicht schrill, schnell erfassbar und zugleich polarisierend, dabei deutungsoffen für die größtmögliche Zahl ist, hat eine Chance durchzukommen. Wie in der Werbung verspricht Emotion den

2 Schwarzenegger (2022); verschiedene Versionen waren auf YouTube und anderen Plattformen zu sehen.

3 So der Titel der berühmten Habilitationsschrift von Habermas (1962), der die Herausbildung einer bürgerlichen Öffentlichkeit seit der frühen Neuzeit zum Gegenstand hat.

direktesten und stärksten Appell; „Emotionalisierung“ gilt deshalb durchgängig als eines der charakteristischen Merkmale in Sozialen Medien, deren „affective politics“ (Boler/Davis 2018) letztlich „affective publics“ (Papacharissi 2015) hervorbringe, und zwar mithilfe von „affective feedback loops“ (ebd.: 23).

Hier ist nicht der Ort, die vielfältigen Aspekte Sozialer Medien ausführlicher darzustellen, wie etwa in den Arbeiten, die unter Berufung auf Habermas und mit ihm nach dem „neuen Strukturwandel der Öffentlichkeit“ (Seeliger/Sevignani 2021) fragen und nach der aktuellen Lage eines „hybriden Mediensystems“ (Staab/Thiel 2021: 284). Dabei geht es entscheidend um die Frage, inwieweit Soziale Medien – auch angesichts des „Anpassungszwangs“, den sie auf Journalismus und klassische Medien ausüben (Jarren/Fischer 2021: 373) – durch emotional übersteigerte Irrationalität, Dauererregung und Blasenbildung Milieus von „konkurrierenden Halböffentlichkeiten“ (Habermas 2021: 497) entstehen lassen und damit eine deliberative Meinungs- und Entscheidungsbildung unterminieren. Im frei zugänglichen und doch nur „halböffentlichen“ virtuellen Raum der Sozialen Medien, der keine öffentliche Verantwortung oder gar Sanktionen für Lügen und Falschaussagen oder Widersprüche kennt (wie das Presserecht), ist „die Trennschärfe zwischen öffentlich und privat und damit *der inklusive Sinn* von Öffentlichkeit verblasst“ (ebd.: 496). Habermas folgert:

Aus der beschränkten Perspektive einer solchen *Halböffentlichkeit* kann die politische Öffentlichkeit demokratischer Verfassungsstaaten nicht mehr als inklusiver Raum für eine mögliche diskursive Klärung konkurrierender Ansprüche auf Wahrheitsgeltung und allgemeine Interessenberücksichtigung wahrgenommen werden; [...] In den USA ist die Politik in den Strudel einer anhaltenden Polarisierung der Öffentlichkeit geraten, [...]. Der – wie man nur hoffen kann: temporäre – Zerfall der politischen Öffentlichkeit hat sich darin ausgedrückt, dass für fast die Hälfte der Bevölkerung kommunikative Inhalte nicht mehr in der Währung krisierbarer Geltungsansprüche ausgetauscht werden konnten. (Habermas 2021: 497)⁴

-
- 4 Mit den gravierenden Folgen für die Struktur der Öffentlichkeit geht die Diskussion über den hier gesteckten Rahmen weit hinaus, zumal man – nach dem Maßstab, den Habermas (1962) gesetzt hat – für das Verständnis ihrer Funktionslogik dreierlei einbeziehen müsse, so Staab/Thiel (2021: 285–291): neben (1) ihrer spezifischen Medialität der „Aneignung von Daten durch überwachungskapitalistische Konzerne“ (i.S.v. Zuboff 2018) auch (2) die veränderte soziale Struktur eines „singularisierten“ Publikums (i.S.v. Reckwitz 2017), für das sie als „Kulturmaschine/Affektmaschine“ nach einer „Logik sozialer Fragmentierung“ durch kommerzielle und/oder politische Inhalte „instabile kollektive Entitäten“ erzeugen, wobei sie (3) in „eine[r] politökonomische[n] Perspektive auf veränderte Akkumulationsmuster“ letztlich „proprietäre Märkte“ bilden, „deren Regeln von mit ihnen identischen Unternehmen

Das Adressierungs- und Verschärfungspotenzial von Emotionalisierung und Ausrichtung auf Unterhaltung erscheint offensichtlich. Bucher (2020), der Beispiele aus den drei Plattformen Facebook, Twitter und YouTube analysiert, zeigt immerhin, dass selbst auf Twitter deliberative Interaktionen noch möglich sind, man müsse aber monologische von dialogischen Beteiligungsstrategien unterscheiden; allerdings seien bis zu 75 % der Beiträge dekontextualisierte Retweets, „eine grundlegend affirmative Kommunikationspraxis“ (ebd.: 133), „oft emotionale Meinungsbekundungen“ (ebd.) und damit monologisch. Für Analysen fordert er die Berücksichtigung von „Interaktionsstrukturen“ und der „Interaktionsordnung“, die wiederum von Algorithmen gesteuert ist.

Auch das Beispiel, das ich hier in einer Einzelfallstudie analysieren möchte, unterscheidet sich von den bisher betrachteten und befürchteten kommunikativen Szenarien, und zwar in einem entscheidenden Merkmal; es versteht sich nicht als Beitrag innerhalb einer demokratischen Öffentlichkeit, sondern ist ganz explizit ein Textstück „von außen“, das im autoritativen System Russlands „von innen“ nicht möglich wäre; die Textsorte ist Gegenpropaganda, bei der aber Wissen und Fakten eine zentrale Rolle spielen, weil man unterstellt, dass die andere Seite nicht nur Irrtümern oder strittigen Meinungen unterliegt, sondern lügt und täuscht, und zwar ganz grundsätzlich. Allerdings muss man dabei zunächst das Vertrauen der Adressaten (und der *bystanders*⁵) gewinnen und diese von der eigenen Aufrichtigkeit überzeugen. Es geht also – wie schon in der antiken Rhetorik – um die Herstellung von Glaubwürdigkeit und die Erzeugung von Plausibilität. Neu sind hier der mediale Stil der Inszenierung und die Multimodalität der Mittel. Wie dabei Elemente von Unterhaltsamkeit in Wort und Bild eingesetzt werden, steht im Zentrum der Analyse.

Dazu muss zunächst erläutert werden, was mit Unterhaltsamkeit gemeint ist, worin sie besteht und welches spezielle Problem hier vorliegt, geht es doch um einen Ernstfall, um Krieg. Dazu werde ich die Herleitung von Kategorien der Unterhaltsamkeit heranziehen, die Josef Klein (1997) entwickelt hat (Kap. 2) und darlegen, dass sie in einer Dimension, der der ‚Unbeschwertheit‘, verändert werden muss, will man den hybriden Verhältnissen moderner Medienkommunikation gerecht werden; dies soll die Analyse des Videos

bestimmt werden“. Hatte die Multidirektionalität der Kommunikation noch Hoffnungen geweckt – mit dieser Konstellation „zerschlägt sich die Vorstellung einer gewissermaßen eingebauten und deshalb herrschaftsfreien Kommunikation“ (Staab/Thiel 2021: 286).

5 Im Sinne von Goffman (1981: 132), der den Begriff anhand von Face-to-face-Situationen entwickelt.

zeigen, das Schwarzenegger kurz nach Beginn des Ukraine-Krieges über Social-Media-Kanäle verbreitete und das in Hollywood-Manier Unterhaltsames verschiedenster Art enthält (Kap. 3). Diese Veränderung des Kleinschen Kategorienschemas bedeutet eine Ausweitung auf Emotionen generell; diese können – so die These – im Sinne des persuasiven Ziels dieses Wissensformats zusätzlich wirkungsvoll sein, weil ihr Anfeuerungspotenzial zum Zwecke der Ermutigung genutzt werden kann.

In der Analyse werde ich den Bild-Sprech-Schrifttext einer multimodal ausgerichteten Betrachtung unterziehen, um schließlich die transkriptive Leistung der jeweiligen Sprach- und Bildkomponenten (im Sinne von Jäger 2002)⁶ in ihrer reißverschlussartigen Verzahnung (dazu Holly 2009) als Faktoren von Unterhaltsamkeit zu beschreiben.

2 Kategorien der Unterhaltsamkeit nach Klein (1997) und eine Modifikation

Wir sind gewohnt, mediale Inhalte nach Genre-Mustern zu verstehen und kennen somit als versierte Rezipienten auch Erwartungen, die man gewöhnlich mit bestimmten Genres verknüpft. Das heißt nicht unbedingt, dass bei der tatsächlichen Rezeption nur genrespezifische mediale Inhalte wahrgenommen und genutzt werden. Unsere Forschungen zur Fernsehrezeption (vgl. Holly/Püschel/Bergmann 2001), die sich auf kommunikative Aktivitäten der Nutzerinnen bei der Rezeption stützten, haben anschaulich gemacht, dass Rezipienten ziemlich unberechenbar sind. Die grundsätzliche Offenheit von Texten und das erhöhte Risiko von Missverständnissen in der Massenkommunikation können manchmal dazu führen, dass erwartbare genrespezifische Reaktionen ausbleiben, dafür aber andere unvorhergesehene Nutzungen dominieren (Holly 1996a). Insofern ist immer auch zu bedenken, dass letztlich bei der Rezipientin entschieden wird, ob etwas ernst genommen wird oder zu Erheiterung führt oder zu noch anderen Reaktionen. Ungeachtet dessen gehören zu unserem sozialen und kulturellen Wissen auch kollektive, manchmal nur gruppenspezifische Vorstellungen davon, was informativ und was unterhaltsam ist, auch wenn es dafür keine festen Grenzen und schon gar nicht starre Regeln geben kann.

6 Der Jägersche Begriff von Transkriptivität (Jäger 2002) als Grundbegriff einer umfassenden semiologischen Zeichentheorie bezeichnet – anders als etwa Transkripte zur Verschriftlichung von Gesprochenem – in skripturaler Metaphorik alle Arten von Bezugnahmen von Zeichen auf Zeichen, wobei etwas „(anders) (un)lesbar“ gemacht wird.

Wie stark kommunikative Erwartungen sind, hat nicht zuletzt Grice mit seinen Konversationsmaximen gezeigt, auf deren Grundlage wir routiniert nicht-expliziten Sprachgebrauch als ironische, höfliche oder unernste (scheinbare) Abweichungen von diesen Maximen deuten können. Diese alltägliche Möglichkeit der Umdeutung vordergründiger Abweichung zeigt schon, dass es Kontexte gibt, die andere Ziele verfolgen als optimale Information, auch wenn die Griceschen Kategorien Informativität, Wahrheit, Relevanz und Klarheit eine Basis für die Kommunikation jeglicher Art von Wissen zu sein scheinen; die ersten beiden sind dabei unverzichtbar. Es ist das Verdienst von Klein (1997), diese vier Kategorien, die er als I-Kategorien (für Information) auffasst, als Kontrast- und Vergleichsgrößen für die Aufgabe zu nutzen, entsprechende U-Kategorien (für Unterhaltung) zu entwickeln; so lassen sich anstelle von vagen Beschreibungsattributen systematische Eigenschaften von Unterhaltung erschließen. Als gemeinsame Kriterien, die eine Vergleichbarkeit erst ermöglichen, legt er die Frage nach der „Dimension der mentalen Verarbeitung“ zugrunde und unterscheidet nun die I- und U-Kategorien nach ihrer „Angemessenheit“ an Ausmaß und Struktur der Verarbeitungskapazität, an das Verarbeitungsziel und an die Focuspräferenz. Das Ergebnis fasst er in folgendem Schema zusammen:

Tabelle 1: Gegenüberstellung von I- und U-Kategorien nach Dimensionen der mentalen Verarbeitung (nach Klein 1997: 182)

<i>Dimension der mentalen Verarbeitung</i>	<i>I-Kategorien</i>	<i>U-Kategorien</i>
Quantitative Angemessenheit an die Verarbeitungskapazität	Informativität	Abwechslung
Angemessenheit an das Verarbeitungsziel	Wahrheit	Unbeschwertheit
Angemessenheit an die Focus-Präferenz	Relevanz	Interessantheit
Angemessenheit an die strukturelle Verarbeitungskapazität	Klarheit	Eingängigkeit

Während es also bei Information auf kognitive Faktoren wie Neuigkeit, Wahrheit, Kontextpassung und Verständlichkeit ankommt, zählt für Unterhaltung anderes: Wechsel, Leichtigkeit, Attraktivität und mühelose Verständlichkeit. In beiden Fällen hält Klein die Angemessenheit an das Verarbeitungsziel für konstitutiv, also Wahrheit und Unbeschwertheit. Während auf Anhieb unstrittig sein dürfte, dass Wahrheit für Information zentral ist, wird Unbeschwertheit, wie Klein weiß, in vielen (nicht-prototypischen) Unterhaltungsangeboten wie Krimis oder Liebesdramen oder gar Horrorfilmen, nicht geboten, sondern stellt sich, dank der Fiktionalität des Stoffes, bestenfalls nachträglich ein, als Entspannung nach emotional fesselnder Spannung: Die Täterin ist gefasst, das Verbrechen wird bestraft, das Schmerzliche wird verarbeitet, die gestörte Ordnung wird wiederhergestellt und damit auch eine (mehr oder weniger) „unbeschwerte“ Gemütsverfassung.

Angesichts der Einschränkung dieser konstitutiven Kategorie scheint es nicht abwegig, an zentraler Stelle eine Erweiterung dessen, was für eine unterhaltsame Verarbeitung angemessen ist, vorzusehen: nicht nur Unbeschwertheit, sondern generell Emotionalität/Erregung, sei es in Form heiterer, unbeschwelter Gefühle (Freude, Liebe, Entspannung, Spaß, Geborgenheit, Stolz, Zusammenhalt) oder eben ernster „beschwelter“ Gefühle (Angst, Trauer, Wut, Empörung, Erregung, Scham, Hass, Ekel). Es ist jedenfalls zu beobachten, dass auch die „beschwerten“ Inhalte nicht nur unfreiwillig erlitten, sondern in medialen Angeboten freiwillig aufgesucht werden, und zwar aus unterschiedlichen Motiven, die von Neugier über Interesse an bestimmten Problemen (und ihrer Lösung) bis zu einer Art „Abhärtung durch Angstlust“ reichen.

Es ist zugleich offensichtlich, dass die mediale Entwicklung zu hybriden Formen immer kühnere Mischungen von Information und Unterhaltung anbietet, sei es als Doku-Drama oder Politsatire, locker präsentierte und abwechslungsreich aufbereitete Spiel- oder Quizshows zu allen möglichen Wissensgebieten, auch zu lebensnahen Fragen aus Politik, Alltag, Sport, Freizeit, Krankheiten, Ernährung usw. Auch wenn unsere Medienkompetenz uns immer noch zu ermöglichen scheint, an der Machart und Rahmung zu erkennen, ob es sich letztlich um Fakten oder Fiktionen, Information oder Unterhaltung handelt – spätestens mit sogenannten „Reality“-Formaten, in denen unklar ist, wie weit sie „geskriptet“ sind, ob es sich noch um authentische Laien handelt oder nur um schlechte Schauspielerinnen, spätestens mit Begriffen wie „alternative Fakten“ und einer neuen Qualität von digitalen Fälschungen wird unsere Medienkompetenz herausgefordert, wenn nicht überfordert. Verschwörungstheorien, die sich um diese Basisdifferenz nicht scheren, sind auf dem Vormarsch und in der Regel emotions-, nicht wissensgesteuert. In jedem Fall muss man davon ausgehen, dass mediale Formate in hohem Maße inszeniert sein können, nicht

zuletzt im Hinblick auf ihre emotional beschleunigte Eingängigkeit durch Unterhaltsamkeit.⁷

Wenn in dem folgenden Beispiel auf dem Hintergrund von Propaganda und Gegenpropaganda eine Mischung von Information und Unterhaltung als Mittel der Persuasion eingesetzt wird, dann ist jedem klar, dass es hier nicht um die Mischung von interessanten Fakten und anschaulichen Fiktionen gehen darf (wie in den meisten Dokudramen mit auch unterhaltsamem Charakter), sondern um die Differenz von Fakten, deren Wahrheit man für sich beansprucht und von Fakes, die man wohlbegründet dem Gegner unterstellt. Fiktional mag vielleicht die stilistisch unterhaltsame Aufbereitung des Stoffes anmuten, gleichzeitig läuft aber immer der Wahrheitsanspruch eines informativ ausgerichteten Formats mit, bei dem es zunächst um Wissensorientierung geht, wegen der unterstellten fehlenden oder falschen Information um „politische Aufklärung“.

3 Schwarzeneggers Video gegen Putin: Information, Appell und ‚beschwerte‘ Unterhaltung

Am 24. Februar 2022 marschierten in einer großangelegten Militäraktion russische Truppen in die Ukraine ein. Drei Wochen später, am 17. März, veröffentlichte Arnold Schwarzenegger über verschiedene Social-Media-Kanäle eine Videobotschaft an das russische Volk, die laut SPIEGEL (21.3.2022) vom kalifornischen Medienkonzern attn: produziert wurde. Das Video ist ca. 9 Minuten lang, der hörbare Sprechtext ist zusätzlich in Form von 194 englischen (und russischen) Untertiteln/UT wiedergegeben; wahrscheinlich lag für die Untertitel die geschriebene englische Fassung seiner Ansprache schon vor. Es gibt also zwei Versionen des Textes, die sich wechselseitig (im Jägerschen Sinn) transkribieren: die hörbare Stimme von Schwarzenegger und die lesbare Untertitelfassung seines Sprechtextes in Form eines nach und nach sichtbaren Schrifttextes in zwei Sprachen, dessen englisches Original wohl auch schriftlich formuliert wurde.⁸ Damit liegt auch eine sehr spezielle Form von Audiovisualität vor, da schon Sprachliches gleichzeitig in zwei Modalitäten (auditiv und visuell)

7 Die Skala reicht von relativ plumpen Verfahren von Influencerinnen bis zu den subtilsten Methoden von Polit-Talkshows (s. z.B. Holly 2012; 2018).

8 Es handelt sich also um einen typischen Fall von „sekundärer Mündlichkeit“, die in elektronischen Medien in verschiedenen Ausprägungen zu finden ist. (Holly 1994b; 1996b).

erscheint.⁹ Grundsätzlich ist also möglich, dass von Rezipienten, vor allem von solchen, die keine Englisch-Muttersprachlerinnen sind, eine Kombination des hörbaren und sichtbaren Textes wahrgenommen wird. Man kann gleichzeitig lesen, was er sagt, und hören, wie er es sagt: ein erster Aspekt von Multimodalität, zwei Lesarten oder besser Wahrnehmungsarten des Rednertexts, die sich auch wechselseitig transkribieren.¹⁰ An manchen Stellen hört man auch, dass, was er sagt, etwas anderes ist als das, was die Untertitel schreiben, allerdings nur in relativ harmlosen Details.¹¹

Aus Gründen der Lesbarkeit und der vermutbaren Textentstehung werde ich im Transkript¹² und bei der Analyse von den englischen Untertiteln ausgehen (auf die russischen Untertitel kann ich nicht eingehen), also dem Schrifttext. Später (s. dazu Kap. 3.2.1) werde ich beschreiben, wie der Ton, d. h. Schwarzeneggers Stimme den Schrifttext transkribiert, also „anders lesbar“ macht. Danach geht es um transkriptive Beziehungen zu dem vorliegenden Bildmaterial.

Was auf der Bildebene zu sehen ist, kann man in 74 Bildeinheiten/BE einteilen; sie zeigen entweder Kameraeinstellungen von Schwarzenegger oder eingeblendete Fotos/Fotosequenzen bzw. Filmsequenzen. Schwarzenegger ist in gut zwei Dritteln der Zeit im Bild (insgesamt 6 min 18 sec), die restlichen 2 min und 48 sec lang sind Fotos/Fotosequenzen und Filme/Filmsequenzen zu sehen, wobei letztere leicht überwiegen. Von Schwarzenegger gibt es zwei Typen von Einstellungen (vgl. Abb. 1 und 2):¹³ die Standardeinstellung (insgesamt 5 min 8 sec) zeigt ihn nah frontal und leicht von unten, die Varianteneinstellungen (insgesamt 1 min 10 sec) sind halbnah von rechts unten, nicht immer gleich und auch Schwenks enthaltend.

9 Dazu außerdem Bilder, man muss also zwischen Bildlichkeit (kodal) und Visualität (modal) unterscheiden (vgl. Holly 2016; 2013).

10 Zur Differenz von Lesen und Hören siehe auch Jäger (2014) und (2020).

11 Solche Abweichungen von Schwarzeneggers Sprechversion werden, wenn sie über die übliche Differenz zwischen Orthografie und phonetischer Realisierung hinausgehen, im Transkript in eckigen Klammern erfasst.

12 Das Transkript ist zugänglich im Anhang, der digital auf der Website <https://www.peterlang.com/document/1392872> vorliegt zusammen mit Abbildungen der verwendeten Einstellungen und der eingeblendeten Fotos und Stills der eingeblendeten Filme.

13 Die Nummerierung der Abbildungen folgt ihrer Position im Video und der Bilder-Übersicht, die im digital verfügbaren Anhang zu finden ist, nicht der Reihenfolge ihrer Behandlung im Text.

Die transkriptiven Beziehungen zwischen Sprechtext, Untertiteln und den sichtbaren Bildern sind durch ständige Wechsel komplex, zumal Bildwechsel nicht immer mit Wechseln von Untertiteln koordiniert sind. (Im Transkript ist erfasst, wenn Untertitel überlappend auf zwei aufeinanderfolgenden Bildeinheiten zu sehen sind.)

Im Folgenden werde ich mich, auch zum Zwecke eines inhaltlichen Überblicks, zunächst mit dem Sprechtext beschäftigen, wie er in den Untertiteln als Schrifttext wiedergegeben ist; danach sollen Beispiele der vielfältigen Transkriptionsphänomene zwischen Ton, Schrifttext und Bildern beschrieben werden, jeweils im Hinblick auf Unterhaltungselemente.

3.1 Der Sprech-/Schrifttext: ein klassisch-rhetorischer Aufbau

Schwarzeneggers früher Appell an die russische Bevölkerung wurde in der Presse geradezu gefeiert (Zitate jeweils online). Die FAZ (18.3.2022) bescheinigte ihm, „dass er selten so ein gutes Drehbuch hatte“, der SPIEGEL (17.3.2022) lobte: Er „verbindet seinen Appell sehr geschickt mit seiner persönlichen Geschichte“ und beschrieb in einem zweiten Text (18.3.2022), warum seine Rede „eine Meisterwerk politischer Rhetorik ist“ (ebd.); die Süddeutsche Zeitung (18.3.2022) spricht von einem „brillantem Video“, zitiert das Urteil des „Director of Digital Strategy der Biden-Administration“ („sehr, sehr bemerkenswert“), vermutet mit einem Verweis auf Hollywood, „da waren schon Soft-Power-Experten der höchsten Kategorie am Werk“, und schließt sich dem Lob an: „Zugleich ist das Video aber rhetorisch, psychologisch und vom Aufbau her eine nahezu perfekt designte Informationskriegswaffe.“ Tatsächlich folgt der Text dem klassischen rhetorischen Schema mit vier Teilen: Exordium/Einleitung, Narratio/Erzählung des Sachverhalts, Argumentatio/Beweisführung und Conclusio-Peroratio/Schluss.¹⁴

3.1.1 Einleitung und Sachverhaltsdarstellung: Exordium und Narratio

Besonders viel Platz, mehr als ein Drittel der Zeit (ca. 3 min 32 sec), nimmt die Einleitung in Anspruch und der Versuch, die Gunst des Publikums zu gewinnen (*captatio benevolentiae*) (bis Bildeinheit/BE 32 bzw. Untertitel/UT 81).¹⁵ In dieser langen Spanne hat der Redner zunächst begrüßt (*Hello, everybody*),

14 Siehe z.B. Ueding/Steinbrink (2011: 259–277); ein historischer Überblick bei Schirren (2009). Zu narrativen Elementen in politischen Kontexten s. auch Klein (2019: 142 ff.).

15 Der Einfachheit halber verweise ich jeweils auf Bildeinheiten/BE und Untertitel/UT.

gedankt (*thank you for sharing your time with me*), adressiert (*my dear Russian friends and the Russian soldiers serving in the Ukraine*) und in Schüben das Thema angedeutet (*there are things that are going on in the world, that are being kept from you ... terrible things that you should know about*). Dann folgt eine lange Passage, in der er seine persönliche Geschichte mit diesem Thema verbindet, und zwar verknüpft mit zwei Personen, dem russischen Gewichtheber Yury Petrovich Vlasow, und seinem eigenen Vater, in der Tat eine filmreife Story, die im Weiteren zur Legitimationsfolie und Bewertungsgrundlage seiner Argumentation und seines Appells wird (s. auch Abb. 3–5 unten). Zugleich wird ein äußerst ergreifend ausgemalter semantischer Frame skizziert, von guten und schlechten Vorbildern, mehr noch: von Helden und tragisch verstrickten Täter-Opfern.¹⁶ Am Übergang von *terrible things*, die er als Thema angedeutet hat, zu einer „human-touch-story“ steht zunächst die Verklärung des russischen Helden Vlasow (UT 10–58; Textzitat 1 im digitalen Anhang, dort Abschnitt I).¹⁷

Diese anrührende Geschichte ist tatsächlich eine persuasive Mehrzweckwaffe. Nicht nur plausibilisiert sie Schwarzeneggers Verbindung zu Russland und den Menschen dort (s. Abb. 6–10), sie zeigt ihn als einen Bewunderer und Nacheiferer eines russischen Volkshelden, der ihn zugleich zum Rebellen gegen den eigenen Vater werden lässt. Man kann sich für die gute Sache sogar gegen den eigenen Vater stellen! Zwischen den Zeilen steht noch mehr: Man wird durch ein Idol selbst zu einem erfolgreichen Vorbild, das jetzt mit der ganzen Autorität eines starken Mannes auftreten kann und vermittelt: Tut es mir nach! Ganz nebenbei werden mit der Figur des Vaters schon die Themen Kriegsschrecken und -verbrechen und Nationalismus, der in die Irre führt, angrissen.

16 Zur Figur des Helden in politischen Kontexten und einer semantischen Analyse des Begriffs Thomä (2019: 25–32). Obwohl im Video Helden von Anfang bis Ende eine zentrale Rolle spielen (explizit: Vlasow, die Protestierenden, implizit Schwarzenegger selbst als Kinoheld), sind nach der Definition von Thomä eigentlich nur die Protestierenden echte Helden, weil sie sich in Gefahr begeben, einer „höheren Sache“ (ebd.) widmen und andere überragen. Die Gewichtheber überragen andere zwar, aber ohne Gefahr und kaum für eine große Sache.

17 Zur besseren Lesbarkeit sind drei längere Textstellen aus dem Transkript, auf die (wie hier) verwiesen wird, zu Blöcken zusammengefasst, im digitalen Anhang (dort Abschnitt I) wiedergegeben.



Abb. 3: Film (drei Einstellungen) Vlasow



Abb. 4: Foto Vlasow



Abb. 5: Film (drei Einstellungen) 2. Weltkrieg

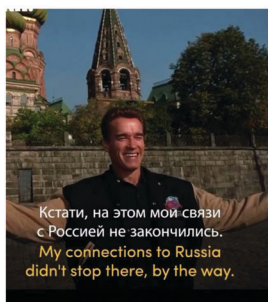


Abb. 6: Foto Schwarzenegger

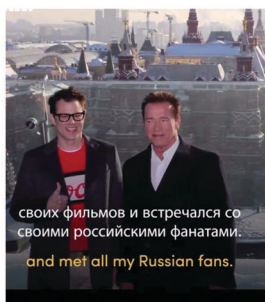


Abb. 7: Foto Schwarzenegger/Fan



Abb. 8: Film (vier Einstellungen) Red Heat

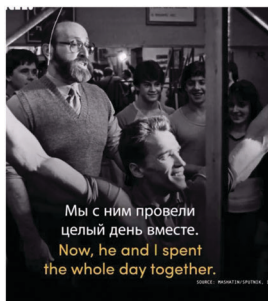


Abb. 9: Foto Vlasow/Schwarzenegger

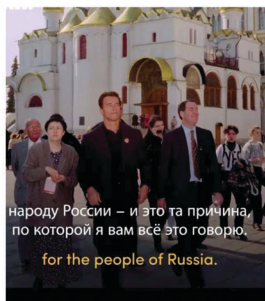


Abb. 10: Foto Schwarzenegger/Volk

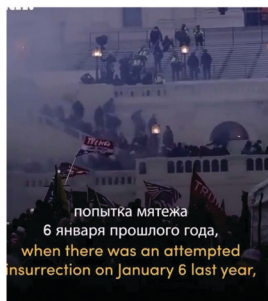


Abb. 11: Film (eine Einstellung) Capitol 1

Die Legitimation eines Eintretens für die richtige Seite und gegen falsche Autoritäten wird im Folgenden noch deutlicher expliziert, zusammen mit der ersten Verwendung des Wortes *Krieg/war*. Dazu wird die *Captatio benevolentiae* noch einmal gesteigert und explizit mit der Hoffnung auf Akzeptanz für die folgenden Informationen (*the truth about the war in Ukraine*) verknüpft, auch wenn diese die russische Führung belasten; als weiteres Argument wird wiederum Schwarzeneggers eigenes Verhalten, diesmal im eigenen Land, angeführt und zu einem Prinzip verallgemeinert: falsche Entwicklungen muss man laut benennen (s. Abb. 11) (UT 59–67; Textzitat 2 im digitalen Anhang, dort Abschnitt I).

Diese letzte Parallele ist zwar schräg, denn Schwarzenegger trat ja für, nicht gegen die eigene Regierung auf. Wichtiger scheint aber, dass er damit einen im Alltag verankerten Argumentationsstil verwendet, bei dem es mehr auf Haltung als auf Logik ankommt. Für unseren Zusammenhang ist jedenfalls entscheidend, dass Schwarzenegger seine gesamte Einleitung mit einer persönlichen bildhaften Erzählung seines eigenen Lebensweges verbindet, die alle Elemente einer narrativen und damit auch unterhaltsamen Darstellung enthält: der Zufall, dass er zu einer Weltmeisterschaft, die in Wien stattfand, eingeladen war; die szenische, gewissermaßen bildliche Vergegenwärtigung einer Backstage-Begegnung eines Vierzehnjährigen mit einem Idol, ein magischer Moment, in dem seine Jungen-Hand von der Riesenhand eines Muskelmannes „verschlungen“ wurde, zugleich eine Szene mythischer Kraftübertragung; das Foto überm Bett, ein typischer Zug pubertärer Verehrung des im Bild gegenwärtigen Idols; der Vaterkonflikt und dessen traumatischer Hintergrund, der einen Parallel-Konflikt zwischen Nazis und Russen umkehrt: jetzt sind die Russen nicht Verteidiger, sondern Angreifer, womit unterschwellig nahegelegt wird, dass sie jetzt wie Nazis agieren, wo sie doch gleichzeitig vorgeben, man müsse die Ukraine „entnazifizieren“. Hier wird vorbereitet, dass dies ad absurdum zu führen ist, ohne dass Argumente explizit werden müssen. Die Geschichte zeigt schon in ihrer Struktur, wie Recht und Unrecht verteilt sind. Die Guten erkennt man an ihrer Leistung und ihrer Menschlichkeit, was in der Wiederbegegnung mit Vlasow wiederum in Eigenschaften und symbolischen Taten zum Ausdruck kommt. Ein Kindertraum ist wahr geworden, die Identifikation mit dem Vorbild hat ein *happy end*, das eine Flut positiver Gefühle auslöst. Der Held ist *thoughtful*, *kind*, *smart* und *giving*: Er verschenkt eine Tasse, die seither nun jeden Morgen begleitet und zum Beweis in die Kamera gehalten wird. Wir erleben einen Star ganz privat, der uns durch Selbstoffenbarung Nähe anbietet und dem wir vertrauen können.

Personalisierung, Dialogisierung und Intimisierung, wesentliche Elemente einer unterhaltsamen Geschichte mit „human touch“, schlagen sich natürlich

auch in entsprechenden Referenzausdrücken nieder: über 50 Mal *you/your*, explizite Adressierungen von Gruppen und Personen, 43 Mal Selbstreferenzen mit *I*, auch mit direkt dialogischer formelhafter Verbindung: *I'm speaking to you...things that you should know about*, überhaupt performative Formeln mit Relevanz für den Beziehungsaspekt wie *thank you, I'm speaking to you, before I talk about, let me (just) tell you (3x), the reason why I'm telling you, I hope that you will let me tell you, I understand that, I hope that you will hear what I have to say, may I remind you that I speak, I/we know that (3x), what you should know, you're also not being told the truth, I regret to tell you, you already know, I don't want you, let me just ask you, remember, I urge, I ask you to help me, let your fellow Russians know, let me close with, may God bless you all*. Sie gehören ebenso zu einem persönlichen informellen Stil wie sogenannte Kolloquialismen, bestimmte Gesprächswörter, Gliederungssignale: *hello, I mean, you see (3x), oh, now (3x), I might add, no, see*.

3.1.2 Beweisführung/Argumentatio und Schluss/Peroratio

Nachdem die Eröffnung mit Thema und Näheaufbau zu den Adressaten geklärt und dabei auch schon eine allgemeine Sachverhaltsdarstellung in Form einer veritablen Erzählung auf bewährte, unterhaltsame Weise geleistet sind, kann die eigentliche argumentative Beweisführung beginnen, die Putin und den Kreml in einer langen Kette von Details für Lüge, Gewalt, Zerstörung, Brutalität, Unmenschlichkeit und Isolation des eigenen Landes verantwortlich macht, knapp 3 Minuten lang (s. dazu Liste mit Textverweisen im digitalen Anhang, dort Abschnitt II).

Nach einer relativ kurzen, aber intensiv und reich bebilderten Passage (2 min 58 sec) ermöglicht ein Vergleich mit dem 2. Weltkrieg eine Verknüpfung mit dem Anfang und einen Übergang zur Conclusio. Wiederum greift Schwarzenegger auf die tragische Verstrickung seines irregeleiteten Vaters zurück, den der Krieg körperlich und seelisch gebrochen zurückgelassen habe; vor solchen persönlichen Folgen sollen die Adressaten bewahrt werden (UT 141–160; Textzitat 3 im digitalen Anhang, dort Abschnitt I). Am Ende der Passage dann die eindringliche Schlussfolgerung: die eindeutige Verurteilung des Kriegs als *illegal und senseless/sinnlos* und folgeschwer für Leib, Leben und Zukunft der Adressaten.

Der dritte und kürzeste Teil des Videos (2 min 22 sec) widmet sich danach eindringlichen hochpathetischen und ganz direkten Appellen an verschiedene Adressatinnen, ein regelrechtes Schlussfeuerwerk mit allem sprachlich-rhetorischen Aufputz. Standen am Beginn unterhaltsame und dramatische Szenen aus der Vergangenheit, so regiert jetzt Pathos: Die emotionale Anklage wird umgemünzt in emotionale Aufforderungen. Rhetorische Fragen richten

sich an die Machthaber im Kreml, explizite Appelle an die Soldaten erinnern an familiäre Verbindungen und fordern auf zur Verbreitung der Wahrheit.

Präsident Putin, der bisher nicht mit Namen genannt wurde, spricht er nun mit einer bündigen Dreierformel mit Klimax direkt an, gewissermaßen ein Showdown „von Mann zu Mann“: *You started this war. You are leading this war. You can stop this war.*

Mit der Erwähnung aller Russinnen, die auf den Straßen gegen die Invasion der Ukraine protestiert haben, ist der eigentliche Zweck des gesamten Videos erreicht: Er rühmt ihre Tapferkeit, würdigt das erlittene Leiden, wenn sie verhaftet, eingesperrt und geschlagen wurden und ermutigt so den Widerstand gegen Putin. In einer weiteren rhetorischen Klimax schließt er den Kreis zum Helden seiner Jugend Vlasow, wiederum in einem Dreierschritt: *You are my new heroes. You have the strength of Yury Petrovich Vlasov. You have the true heart of Russia.* Nicht weniger pathetisch seine abschließende Grußformel, eine Variante des Endes amerikanischer Präsidentenansprachen: *My dear Russian friends, may God bless you all.*

3.2 Unterhaltung durch audiovisuelle Multimodalität: der Sprech-Schrift-Bild-Reißverschluss¹⁸

Nachdem in einem ersten Schritt der Sprechtext in seinem klassisch-rhetorischen Aufbau schon einige Hinweise auf unterhaltsame Elemente gegeben hat, vor allem im ersten Drittel durch die mythisch und bildlich-szenisch angereicherte Erzählstruktur einer Geschichte von Helden (und ihren Gegenspielern) und die durch autobiographische Authentizität hergestellte Intimität der Beziehung eines Hollywood-Stars zu seinen Adressaten, soll nun die funktionale Multimodalität des Videos bei der Erzeugung einer attraktiven, weil unterhaltsamen Gestaltung deutlicher fokussiert werden. Natürlich enthält auch schon die sprachliche Formulierung auf allen Ebenen vom Laut bis zum Text eine Reihe von weiteren Mitteln, die rhetorisch-stilistisch zur Unterhaltsamkeit des Ganzen beitragen, worauf ich hier aber nicht weiter eingehen will. Stattdessen will ich vor allem zeigen, wie durch eine reißverschlussähnliche Verzahnung des *gesprochenen* und *hörbaren* Texts, auf zwei Weisen, einmal mit dem *sichtbaren* Schrifttext, dann vor allem mit den *sichtbaren* Video-Bildern (Kamera-Einstellungen

18 Der plakative Ausdruck „Reißverschluss“ (Holly 2009) soll beibehalten werden, auch wenn (wie oben gezeigt) schon auf der Sprachebene weitere Differenzierungen vorliegen; dies gilt – wie man sehen wird – auch für die Bildebene, so dass man eigentlich eher von einem ganzen Geflecht von Bezugnahmen wie bei einem Klettverschluss sprechen kann.

des Sprechers, Fotos und Filmsequenzen) Unterhaltsamkeit hergestellt bzw. gesteigert wird. Dabei ist wichtig zu beschreiben, dass diese Effekte häufig erst im Zusammenspiel von wechselseitigen Transkriptionen der Modalitäten entstehen, d.h. weil der Sprechtext das Bild bzw. den Schrifttext oder das Bild den Sprechtext „(anders) lesbar macht“, wie Jäger (2002) in seiner Theorie der Transkriptivität solche Formen der Bezugnahme erläutert.

3.2.1 *Schwarzenegger im Ton: Charisma, Nahbarkeit, Menschlichkeit*

Wie schon erwähnt dominiert auf der Bild- und Tonebene Schwarzenegger als Sprecher und Person, es ist *sein* Text, *seine* Person autorisiert und authentifiziert, was er sagt, mit seinem Bild und seiner Stimme. Damit soll unumstritten klar werden, dass er im eigenen Namen spricht,¹⁹ aus eigener Verantwortung mit dem ganzen charismatischen Gewicht seiner öffentlichen Aura als Star.²⁰ Man könnte, wenn man ihn schlecht versteht, den Sprachtext nur als Schrift wahrnehmen, und die Transkription seiner Worte durch seinen Ton damit ausblenden. Das gilt vor allem, wenn man nur die russischen Untertitel wahrnimmt und kein Englisch versteht; aber selbst dann werden wohl noch gewisse emotionale Qualitäten seiner Stimme wahrgenommen. In anderen Fällen der Doppelwahrnehmung von Schrift und seiner phonetischen Realisierung des Textes gilt, was als eine erste Ebene der Transkriptivität zu beschreiben ist, Formulierungen werden semantisch überformt und angereichert durch Lautqualitäten.

Zu seinem Charisma-Typus gehört nämlich die (auch hörbare) Eigenschaft, ein starker Mann und doch nicht rüde und unnahbar zu sein, eben ein Mann aus dem Volke,²¹ dessen (steirischer) Dialekteinschlag für alle unüberhörbar ist, die einer englischen Varietät als Erst-, Zweit- oder Fremdsprache mächtig sind. Dies allein

19 Darauf verwies auch sein Pressesprecher, der nach dem Lob der Biden-Administration ausdrücklich gegenüber der New York Times bekräftigte, die Botschaft komme „aus Arnolds Herzen, nicht von der Regierung“ (Süddeutsche Zeitung, 18.3.2022). In den Termini von Goffman (1981) spricht er – wenn nicht unbedingt als formulierender *author* – dann doch als verantwortlicher *principal* des Textes.

20 Die Bedeutung der Stimme für diese Textsorte war auch schon Thomas Mann klar, der, nachdem seine Texte zunächst nur verlesen worden waren, anregte, trotz eines aufwendigeren Übertragungsverfahrens selbst zu sprechen: „Auf diese Weise hören diejenigen, die drüben zu lauschen wagen, nicht nur meine Worte, sondern auch meine eigene Stimme“ (Mann 1975: 5). Er sprach in Los Angeles auf eine Platte, die nach New York geflogen wurde, dann über Telefon auf eine andere Platte in London übertragen wurde, die man vor einem Mikrofon laufen ließ.

21 Zu Charisma-Typen in der Politik (Rollen im politischen Starsystem) Schwarzenberg (1980: 9–107).

schon wegen seiner gesamten Prosodie; deutlicher noch sind markierte Lautungs-Varianten wie etwa bei *war*, das er nicht amerikanisch [ˈwɔr] ausspricht, auch nicht britisch [ˈwɑː], sondern in auffälliger Weise mit einem langen geschlossenen o und einer zu a vokalisierten r-Variante [ˈwoːɑ], beides systematische Abweichungen von Standardrealisierungen der entsprechenden Laute, die auch an anderen Stellen zu hören sind. Solche fast ein wenig drollig wirkenden Sprechweisen geben der gesamten Präsentation einen anheimelnden, zutiefst menschlich erscheinenden Zug, allein schon, weil er in dem schwer zu kontrollierenden, meist vollständig habitualisierten Sprechverhalten unverstellt, durch und durch echt wirkt. Dazu gehört auch, dass er die Geschichte seiner Begegnungen mit Vlasow in einem nahezu kindgerechten Erzählton vorträgt, fast übertreibend, wie man Kindern Märchen ausgestaltet: die Unerhörtheit der Begegnung des Vierzehnjährigen mit dem Weltmeister führt zu einem merkwürdigen Doppelakzent auf *thére wás*, seine Bewunderung drückt sich aus in einem retardierten Tempo in *the strongest man*; und wenn er ihn bei der zweiten Begegnung als *so thoughtful, so kind, so smart* beschreibt, wird seine Stimme so sanft, wie der große Vlasow selbst zu ihm war: Fälle von lautlicher Ikonizität. So werden seine Formulierungen durch die stimmliche Realisierung in markierter Weise transkribiert, also noch zusätzlich anders verständlich, nämlich als Belege für Authentizität, sympathische Menschlichkeit, geradezu kindliche Unverstelltheit und Lauterkeit.

Natürlich umfasst das stimmliche Ausdrucksrepertoire des Schauspielers Schwarzenegger auch andere Farben und Qualitäten, ohne dass er je das Niveau professionell ausgebildeter Sprecherinnen erreicht. Er liest zweifellos von einem Teleprompter ab, was ihm eine flüssige und manchmal zügige Textwiedergabe erlaubt. Dabei kann er doch argumentative Linien herausarbeiten, mit Nachdruck, mit Überzeugung, mit Zweifel, mit Bedauern und mit Leidenschaft sprechen, was hier nicht im Einzelnen belegt werden soll. Insgesamt bleibt er bei seinem Image als starker, aber letztlich gütiger und aufrichtiger Mann, der gerade auch an die Adresse eines virilen Männertypus glaubwürdig und respekttheischend kommunizieren kann. Zugleich ist sein Ton, der an eine wohlwollende, aber durchsetzungsstarke Alltagsrhetorik ohne Spitzfindigkeiten anschließen kann, klar und eingängig und erfüllt damit sowohl das Kriterium verständlicher Information wie das niedrigschwelliger populärer Unterhaltung.

3.2.2 Schwarzenegger im Bild: Abwechslung, Körpersprache, Glaubwürdigkeit

Was die bildliche Präsentation Schwarzeneggers angeht, ist das auffälligste Merkmal von Unterhaltungsorientierung wohl die sorgfältige, aber dennoch

unauffällige Bemühung um Variation. Abwechslung ist ja auch bei Klein eine zentrale Unterhaltungskategorie. Die Standardeinstellung (s. Abb. 1) zeigt ihn nah, frontal, leicht von unten, was einerseits den Zuschauer einbezieht (involvement), ihn zugleich leicht dominiert, d.h. den Sprecher eher eindrucksvoll und mächtig erscheinen lässt (impressive and powerful).²² Alternativ zu dieser Standardeinstellung, die insgesamt 5 min 8 sec zu sehen ist und die 33 Mal im Bild ist, gibt es eine Variante (s. Abb. 2), die den Sprecher halbnah von rechts unten zeigt, also wiederum leicht erhöht und mit Bewunderung, aber eher aus einer Beobachterperspektive und aus einer gewissen Distanz; die zweite Einstellung (die selbst in geringfügigen Varianten mit leichten Schwenks nach rechts und links vorliegt) ist insgesamt nur 1 min 10 sec zu sehen, also nur ca. ein Fünftel der Standardeinstellung nach Dauer, aber doch noch 20 Mal nach Frequenz. D.h., trotz der viel geringeren Gesamtzeit kommt diese Einstellung relativ häufig, wenn auch z.T. extrem kurz ins Bild. Das zeigt, dass sie in erster Linie für Abwechslung durch Reizerneuerung sorgen soll, schon allein, weil starres Blickverhalten nicht unserer Natur entspricht, aber natürlich auch, weil man so Langeweile vermeidet.²³

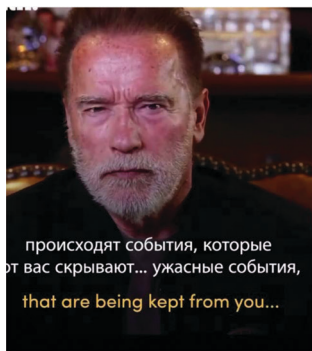


Abbildung 1: Standard, nah frontal



Abbildung 2: Varianten, halbnah von rechts

²² Zu diesen Kategorien der sozialsemiotischen Bildanalyse Jewitt/Oyama (2001: 135).

²³ Zum Wechsel von Bildern und ihrer Funktion in Talkshows Holly (2010: 376 ff.; 2012: 176 ff.; 2015: 131 ff.).

Darüber hinaus haben die beiden Einstellungen zusätzliche unterhaltungsrelevante Funktionen. Die Variante zeigt uns den Sprecher in seiner Umgebung, was uns immer wieder Gelegenheit gibt, unser Bild von ihm anzureichern. Wir sehen ihn an einem massiven repräsentativen Schreibtisch in einem Ledersessel thronen, die Hochglanz-Tischplatte spiegelt ihn, ebenso wie Gegenstände, von denen er umgeben ist; man sieht die ominöse Tasse, vor allem leuchtet hinter ihm ein großformatiges Foto einer Rückansicht des jungen Arnold-Oberkörpers in einer typischen Bodybuilder-Pose mit eindrucksvoll gespannten Armmuskeln, eine Bild-Ikone, die uns immer wieder vergegenwärtigt, wer da zu uns spricht: ein Weltstar in einer durchaus luxuriösen Umgebung, die zu ihm passt, in einer Wohlfühl-Atmosphäre in warmen Tönen, die Kleidung eher sportlich. Die Hände liegen meist auf dem Tisch, die Fingerspitzen berühren sich in Ruhe locker, werden nur gelegentlich zu Gesten geöffnet und hochgenommen, was in der seitlichen halbnahe Variante deutlicher wahrzunehmen ist.

Die Standard-Einstellung zeigt ihn näher, so dass man Kopfbewegungen, Mimik und Augen gut verfolgen kann, während die Hände nur bei stärkeren Bewegungen und nur zum Teil sichtbar werden. Zusammenfassend kann man sagen, dass er sich bemüht, seine Ansprache mimisch und gestisch so abwechslungsreich zu machen wie im gegebenen Rahmen möglich. Man erkennt medienereprobte Bildakteure meist daran, dass es ihnen gelingt, sogenannte Adaptoren zu kontrollieren, die als Anzeichen von Stress gelten, andererseits funktionale Mimik- und Gestik-Elemente zu effektivieren. In diesem Sinne zeigt sich Schwarzeneggers Körpersprache darauf ausgerichtet, in angemessener Weise den Sprechtext zu unterstützen²⁴ und ihn auf diese Weise intensiver und zugleich interessanter, d.h. auch unterhaltsamer zu machen. Ohne dass ich dies hier im Einzelnen darlegen kann, lässt sich doch festhalten, dass Schwarzenegger kontinuierlich sprachbegleitend durch Gesichtsausdruck, Blickbewegung, Handgesten, Kopfhaltung und -bewegung zusammen mit seinem Stimmausdruck eine Reihe von Verstärkungen, Modifikationen, Rhythmisierungen, Gliederungen und Veranschaulichungen des Sprechtexts vornimmt und diesen dadurch noch anders oder besser verständlich macht.

So indiziert er beispielsweise durch Kopfnicken Betonungen, durch Kopfschütteln Negationen, Bedauern oder Ungläubigkeit, durch Hin- und Herbewegung des Kopfes wird das Hin und Her im Streit mit dem Vater ikonisiert, Konzessives durch Kopfneigung, Adversatives durch Kopfheben begleitet. Immer wieder wird mit dem Zeigefinger akzentuiert, es wird durch Öffnen

24 Dazu z.B. Ricci Bitti (2014).

der Hände plausibilisiert und Fragliches angezeigt oder durch Aufstellen der Hände die Abwehr möglicher Einwände angedeutet. Die Augen, die meist leicht zugekniffen erscheinen, können zur Rhythmisierung oder Betonung geöffnet werden. Er ist also ein lebendiger Redner, der mit Körpereinsatz seine Worte verstärkt und beglaubigt, ganz im Sinne der klassisch-rhetorischen Actio-Lehre, nach der das Verhalten des Redners einen wesentlichen Beitrag zu dessen Glaubwürdigkeit und Sympathie leistet.

3.2.3 Foto- und Filmeinblendungen: Personalisieren, Bewerten, Authentisieren, Emotionalisieren

Natürlich dienen die 21 Bildeinheiten,²⁵ die als Unterbrechungen der beiden Sprecheransichten eingeschoben werden, wiederum dem Ziel, neue visuelle Anreize zu schaffen. Wichtiger, weil semantisch markiert, sind Sprach-Bild-Bezüge, die scheinbar gegen die prototypischen Potenziale der jeweiligen Kodalität verstoßen. Wenn Sprache sich nicht auf sachlich-beschreibende Sachverhaltsdarstellungen beschränkt, sondern erzählend in Bildern vorgeht wie Schwarzenegger in seiner anrührenden Helden-Geschichte, ist die Orientierung nicht ausschließlich an die I-Kategorien Informativität und Wahrheit gerichtet, sondern auf U-Kategorien wie Interessantheit und Eingängigkeit. Bilder hingegen scheinen wegen ihrer semantischen Unbestimmtheit (vgl. Sachs-Hombach 2003: 174 ff.) nur unter bestimmten Bedingungen zweifelsfreies Referieren zu gewährleisten, also I-Kategorien der Informativität und Klarheit; wo es dennoch der Fall ist, ermöglichen sie es, Referenzen unauffällig semantisch aufzuladen, mit entsprechendem persuasiven Potenzial.

Dies zeigt sich im Falle zweier Bezugnahmen auf Putin, der zentralen Figur, gegen die sich das Video richtet, dessen Name aber erst in den letzten 30 Sekunden in dem erwähnten frontalen Appell fällt. Bis dahin wurde als Urheber von Propaganda und Lüge, gegen die man die Stimme erheben müsse, zwei Mal *your government* bzw. *the Russian government* explizit gemacht (UT 81–82; 126); im Bild ist aber jeweils Putin zu sehen (Abb. 14; 24), und zwar mit typischen transkriptiven Leistungen: erstens mit Personalisierung der Referenz, die nun nicht eine Institution, sondern einen Menschen mit Verantwortung zeigt, der noch dazu mit negativen Eigenschaften bzw. Handlungen charakterisiert

25 Die Bilder sind auch im digitalen Anhang zu finden (dort Abschnitt III), der auf der Verlagswebsite vorliegt; die Filme können natürlich nur durch Stills repräsentiert werden; in einigen Fällen sind, wenn sie aus mehreren Einstellungen bestehen, auch mehrere Stills unter einer Abbildungsziffer zu sehen.

wird, im einen Fall durch Farben (vor blauem Hintergrund an einem grauen Tisch) als kalt und eigentlich unmenschlich, im anderen Fall zusätzlich noch nach dem Muster „Mit Worten sehen“ (Holly 2006) als Lügner, denn das Bild wird wiederum vom Sprachtext transkribiert, der das Prädikat *has lied* enthält, so dass wir damit angehalten sind, das irgendwie verdruckste, aber immer noch semantisch unbestimmte Gesicht auf dem Putin-Foto als das eines Lügners zu deuten, jetzt ausgestattet mit der stärker authentisierenden Überzeugungskraft eines realen Fotos, stärker als eine bloße Behauptung durch Worte. So muss man den Bösen gar nicht einmal nennen, es genügt, ihn im Bild gewissermaßen zu erwischen, wie er „wirklich“ ist; eine gängige Praxis der bildlichen „Entlarzung“ in Medien, durchaus mit Unterhaltungswert.

Auch Putins Kontrahent Selensky (Abb. 15, vor einem religiösen Symbol) wird nicht mit Namen genannt, gefolgt von Dritten, die für ein Meinungsbild in der Welt stehen: die UN als Institution und weltweit protestierende Menschen (Abb. 16–19). Danach (Abb. 20–27) folgen Schreckensbilder von Bomben, Zerstörung, Flucht, Waffen im Einsatz mit sichtbaren Detonationen und Rettungsaktionen und Toten. Solche Bilder beglaubigen das Behauptete, wie sie es zugleich emotional aufladen.

Weitere transkriptive Beispiele liefert die bildliche Verschiebung der Referenz weg vom Vater, der nicht gezeigt wird; stattdessen sehen wir kaum entzifferbare historische Filmsequenzen in schwarz-weiß (Abb. 5/BE 11), die wir erst durch den Sprachtext (UT 37–40) als Bilder von der Schlacht um Leningrad deuten, zugleich als Charakterisierung eines Täter-Opfer-Antihelden (*did vicious harm to the great city / was injured*), eine bildliche Entpersonalisierung, die aber seine spätere argumentative Rolle als zu vermeidendes Beispiel (UT 141–153) vorbereitet, am Übergang zu den eindringlichen Appellen: *I don't want you to be broken like my father*.

Im letzten Drittel dienen die Bilder vordringlich sozialsemiotisch der Charakterisierung und Emotionalisierung als persuasive Aufladungen: Putin als unnahbarer Machtmensch an seinem überlangen Tisch (Abb. 28), mit seinem Vasallen Lavrov die Köpfe zusammensteckend (Abb. 29), mit einer Waffe in der Hand neben einem General (Abb. 30); sie transkribieren die Appelle an die wiederum sprachlich anonymen Kremlmachthaber (UT 161–164), indem sie Kälte, Schändlichkeit und Gefährlichkeit der Adressaten ikonisieren.

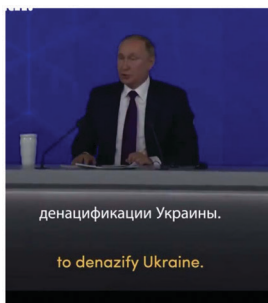


Abb. 14: Film (eine Einstellung) Putin 1

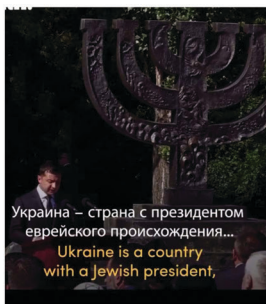


Abb. 15: Film (eine Einstellung) Selensky



Abb. 16: Foto UN

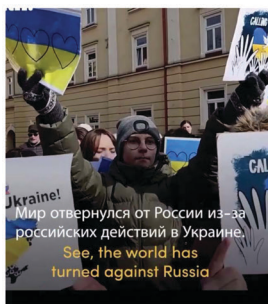


Abb. 17: Film (eine Einstellung) Protest

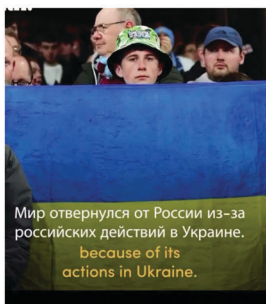


Abb. 18: Foto Protest 2



Abb. 19: Foto Protest 3

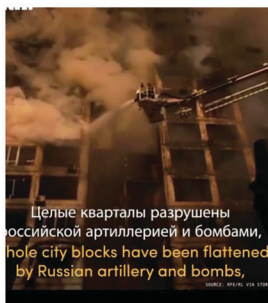


Abb. 20: Film (zwei Einstellungen) Bomben



Abb. 21: Film (eine Einstellung) Flucht

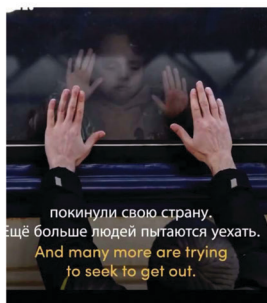


Abb. 22: Foto Flucht

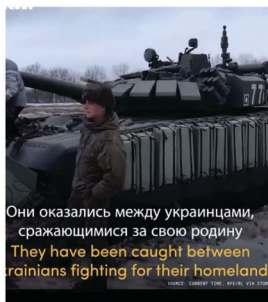


Abb. 23a: Film 1. Einstellung

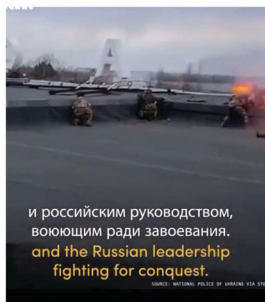


Abb. 23b: 2. Einstellung



Abb. 23c: 3. Einstellung



Abb. 23d: 4. Einstellung



Abb. 23e: 5. Einstellung

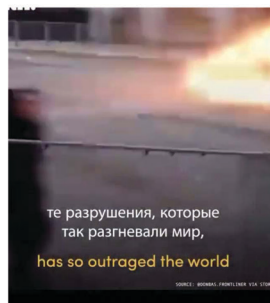


Abb. 23f: 6. Einstellung

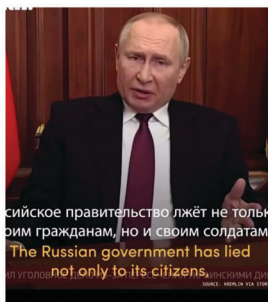


Abb. 24: Film (eine Einstellung) Putin 2



Abb. 25: Foto Kriegsgruel 1

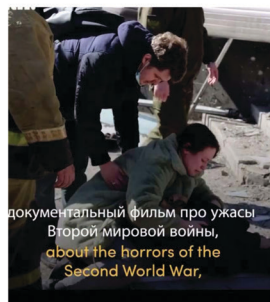


Abb. 26: Foto Kriegsgruel 2

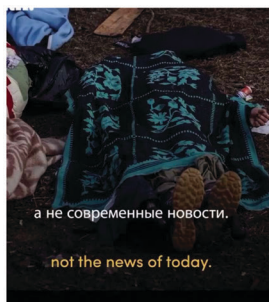


Abb. 27: Foto Kriegsgreuel 3

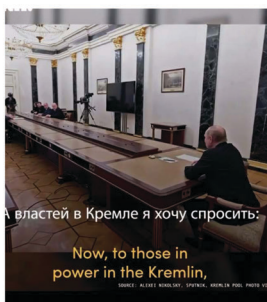


Abb. 28: Foto Putin 3



Abb. 29: Foto Putin/Lavrov



Abb. 30: Foto Putin/Waffe



Abb. 31: Film (eine Einstellung) Opfer

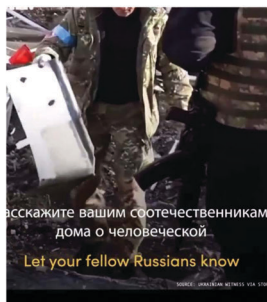


Abb. 32: Film (zwei Einstellungen) Krieg



Abb. 33: Film (eine Einstellung) Proteste 2

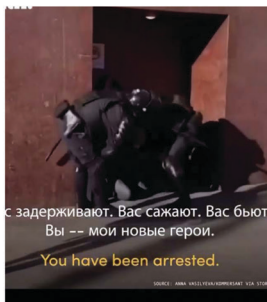


Abb. 34a: Film 1. Einstellung

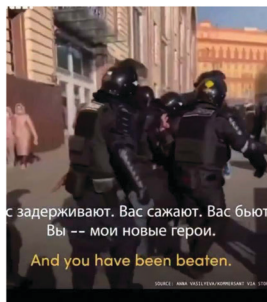


Abb. 34b: Film 2. Einstellung

Während des Appells an die Soldaten sehen wir bewegte und verwackelte Bilder von Zivilisten als Opfer, die von einem Angriff überrascht werden (Abb. 31); hier ist offensichtlich, dass es nicht informativ um das Geschehen geht, sondern vordringlich um eine Stimmung, um die Unsicherheit und Gefahr, an der das Bild uns durch Sensualisierung körpernah teilnehmen lässt (vgl. Holly 2009). Schließlich sehen wir, während Schwarzenegger sie ermutigt, als vorbildhafte „neue“ Helden Demonstranten in Russland (Abb. 33), gefolgt von Bildern der sie knüppelnden Polizisten (Abb. 34a–c); die Bilder belegen, was uns zugleich empören soll.

Starke Erregungen, die in fiktionalen Kontexten für Interesse auch an „negativen“ Themen und Gefühlen sorgen, können auch nicht-fiktionale Texte interessant und auf eine „beschwerte“ Weise attraktiv machen; dann allerdings steht am Ende nicht die reinigende und moralisch festigende Katharsis wie in der Tragödie, nicht die entspannende Einsicht, dass es sich ja nur um eine lehrreiche Fiktion, einen Mythos längst vergangener Geschehnisse handelt. Wie der rhetorische Einsatz persuasiver Mittel schon seit der Antike belegt, sind es nicht nur sachliche Argumente und die Überzeugungskraft eines ethisch agierenden Redners, sondern es sind vor allem die Wirkungen starker Emotionen, also Pathos-Elemente, die den Zuhörer binden. In persuasiven Kontexten wird Erregung für den eigentlichen Zweck der Rede, den Appell gebraucht. Der Zuhörer soll erschüttert dem Redner in dessen Meinung folgen und damit auch seiner Handlungsaufforderung.

4 Fazit: Unterhaltsame Elemente in multimodaler Transkriptivität

Die Videoanalyse hat gezeigt, wie im transkriptiven Zusammenspiel von Textstrukturen, Sprecherqualitäten in Ton und Bild, und eingebetteten (Bewegt-)Bildern eine logisch-argumentativ fundierte Sachverhaltsdarstellung entfaltet wird, die als Gegenpropaganda wirken soll und die dazu in einen informativen Zusammenhang unterhaltsam-erregende Elemente zu persuasiven Zwecken einsetzt. Schwarzeneggers autobiografisch eingebettete und politisch-historisch grundierte Heldenerzählung soll Interesse wecken und eine positive Adressatenbeziehung herstellen, ohne dass ihre Komplexität ganz geklärt werden muss. Ganz offen explizit gemacht wird dagegen die Verknüpfung von Sympathiewerbung und Überzeugungsabsicht. Am Ende mündet die Adressierung verschiedener Gruppen und Personen in gefühlsgeladene Appelle im Sinne der Ermutigung. Es entspricht der heute medienimmanenten Ubiquität des Unterhaltungsmodus, dass nicht einmal das „beschwerte“ Thema

Krieg ohne unterhaltsame Elemente inszeniert wird. Diese Art von Unterhaltung ist natürlich – so wenig wie Kriminal- oder Horrorfilme – nicht mehr ‚unbeschwert‘, ihre Emotionalität kann aber den persuasiven Appell durch Pathos intensivieren.

Auf der Basis von traditioneller Rhetoriklehre, Gricescher Kommunikationstheorie mit Kleins Kategorien der Unterhaltsamkeit, semiologisch-multimodaler Transkriptionstheorie, sozialsemiotischer und philosophischer Bildanalyse und kritischen Ansätzen von Theorien der Medienkultur sollte die medienlinguistisch-hermeneutische Analyse eines Einzelfalles auf den Umstand verweisen, dass Unterhaltungselemente auch in ernsthaften Wissensformaten keine geringe Rolle spielen. Die zugrundeliegende These war, dass im hybriden Stil des durch Social Media zusätzlich emotionalisierten und intimisierten Raums einer „Halböffentlichkeit“ die persuasive Behandlung hochpolitischer und tiefersten Themen wie Krieg und Unmenschlichkeit, deren angemessene Behandlung auf eine gemeinsame Basis von Wahrheit und faktischem Wissen angewiesen ist, auf unterhaltsame Elemente nicht verzichten kann. Wie zu zeigen war, greift Schwarzeneggers Hollywood-Stil auf tradierte Verfahren der Rhetorik zurück, die zu persuasiven Zwecken lange schon die Dreieit von Logos (sachlicher Richtigkeit), Ethos (moralischer Wertbasis) und Pathos (starker Gefühle) ins Spiel gebracht hat. Gleichzeitig zeigt sich, dass die strikte Gegenüberstellung von Information und Unterhaltung, die traditionelle Genre-Konventionen bis heute prägt, längst nicht mehr durchgängig wirksam ist. Auch wenn die Situation nicht unbeschwert ist, können angenehme Unterhaltungselemente wie Bewunderung, Heldenverehrung oder Erfolg und die emotional schockierende Tragik von Krieg und Unterdrückung in einem Stück Gegenpropaganda miteinander kombiniert werden. Die multimodale Gestaltung durch akustische und visuelle Sprachlichkeit und bewegte und nicht-bewegte Bildlichkeit, die sich allesamt reißverschlussartig wechselseitig transkribieren, sorgen nicht nur für niedrigschwellige und abwechslungsreiche Präsentation, sondern auch für Emotionalisierungen verschiedenster Art.

Literatur

Das analysierte Video findet sich z.B. auf dem YouTube-Kanal der britischen Zeitung *The Telegraph* unter <https://www.youtube.com/watch?v=fWCIXZd9c78> (29.8.2023). Das Transkript, ebenso Abbildungen der verwendeten Einstellungen und der eingeblendeten Fotos und Stills der eingeblendeten Filme sind im digitalen Anhang zu finden unter <https://www.peterlang.com/document/1392872>.

- Boler, Megan & Davis, Elisabeth (2018): The affective politics of the ‹post-truth› era: Feeling rules and networked subjectivity. *Emotion, Space and Society* 27, 75–85.
- Bucher, Hans-Jürgen (2020): Zwischen Deliberation und Emotionalisierung. Interaktionsstrukturen in Sozialen Medien. In: Marx, Konstanze; Lobin, Henning & Schmidt, Axel (Hrsg.): *Deutsch in sozialen Medien. Interaktiv – multimodal – vielfältig*. Berlin/Boston: De Gruyter, 123–145.
- Dörner, Andreas (2001): *Politainment. Politik in der medialen Erlebnisgesellschaft*. Frankfurt a.M.: Suhrkamp.
- Goffman, Erving (1981): Footing. In: Goffman, Erving: *Forms of Talk*. Philadelphia: University of Pennsylvania Press, 124–159.
- Habermas, Jürgen (1962): *Strukturwandel der Öffentlichkeit. Untersuchungen zu einer Kategorie der bürgerlichen Gesellschaft*. Darmstadt/Neuwied: Luchterhand.
- Habermas, Jürgen (2021): Überlegungen und Hypothesen zu einem erneuten Strukturwandel der politischen Öffentlichkeit. In: Seeliger & Seignani (Hrsg.): 470–500.
- Holly, Werner (1990): Politik als Fernsehunterhaltung. Ein Selbstdarstellungsinterview mit Helmut Kohl. *Diskussion Deutsch* 21, 508–528.
- Holly, Werner (1991): „Wir sind Europa“. Der Fernsehwerbespot der SPD zur Europawahl 1989. In: Liedtke, Frank; Wengeler, Martin & Böke, Karin (Hrsg.): *Begriffe besetzen. Strategien des Sprachgebrauchs in der Politik*. Wiesbaden: Westdeutscher Verlag, 258–275.
- Holly, Werner (1994a): Confrontainment. Politik als Schaukampf im Fernsehen. In: Bosshart, Louis & Hoffmann-Riem, Wolfgang (Hrsg.): *Medienlust und Mediennutz. Unterhaltung als öffentliche Kommunikation*. München: Ölschläger, 422–434.
- Holly, Werner (1994b): Secondary Orality in the Electronic Media. In: Quasthoff, Uta (Hrsg.): *Aspects of Oral Communication*. Berlin, New York: De Gruyter, 340–363.
- Holly, Werner (1996a): Hier spricht der Zuschauer. Ein neuer methodischer Ansatz in der sprachwissenschaftlichen Erforschung politischer Fernsehkommunikation. In: Klein, Josef & Diekmannshenke, Hajo (Hrsg.): *Sprachstrategien und Dialogblockaden. Linguistische und politikwissenschaftliche Studien zur politischen Kommunikation*. Berlin/New York: De Gruyter, 101–121.
- Holly, Werner (1996b): Mündlichkeit im Fernsehen. In: Biere, Bernd U. & Hoberg, Rudolf (Hrsg.): *Mündlichkeit und Schriftlichkeit im Fernsehen*. Tübingen: Narr, 29–40.

- Holly, Werner (2005): Audiovisualität und Politikvermittlung in der Demokratie. In: Kilian, Jörg (Hrsg.): *Sprache und Politik. Deutsch im demokratischen Staat*. Mannheim: Dudenverlag, 278–293.
- Holly, Werner (2006): Mit Worten sehen. Audiovisuelle Bedeutungskonstitution und Muster ‚transkriptiver Logik‘ in der Fernsehberichterstattung. *Deutsche Sprache* 34, 135–150.
- Holly, Werner (2008): Tabloidisation of political communication in the public sphere. In: Wodak, Ruth & Koller, Veronika (Hrsg.): *Handbook of Applied Linguistics. Band 4: Handbook of Communication in the Public Sphere*. Berlin/New York: De Gruyter, 317–341.
- Holly, Werner (2009): Der Wort-Bild-Reißverschluss. Über die performative Dynamik der audiovisuellen Transkriptivität. In: Feilke, Helmuth & Linke, Angelika (Hrsg.): *Oberfläche und Performanz*. Tübingen: Niemeyer, 389–406.
- Holly, Werner (2010): Besprochene Bilder – bebildertes Sprechen. Audiovisuelle Transkriptivität in Nachrichtenfilmen und Polit-Talkshows. In: Depermann, Arnulf & Linke, Angela (Hrsg.): *Sprache intermedial: Stimme und Schrift, Bild und Ton*. Berlin/New York: De Gruyter, 359–382.
- Holly, Werner (2012): Transkriptiv kontrollgemindert: Automatismen und Sprach-Bild-Überschreibungen in Polit-Talkshows. In: Conradi, Tobias; Ecker, Gisela; Eke, Norbert O. & Muhle, Florian (Hrsg.): *Schemata und Praktiken*. München: Fink, 161–189.
- Holly, Werner (Hrsg.) (2013): *Themenheft Textualität – Visualität. Zeitschrift für germanistische Linguistik* 41.
- Holly, Werner (2015): Bildinszenierungen in Talkshows. Medienlinguistische Anmerkungen zu einer Form von „Bild-Sprach-Transkription“. In: Girnth, Heiko & Michel, Sascha (Hrsg.): *Polit-Talkshow. Interdisziplinäre Perspektiven auf ein multimodales Format*. Stuttgart: ibidem, 123–144.
- Holly, Werner (2016): Mediale Kulturen: Audiovisualität. In: Jäger, Ludwig; Holly, Werner; Krapp, Peter; Weber, Samuel & Heekeren, Simone (Hrsg.): *Sprache – Kultur – Kommunikation. Language – Culture – Communication*. Berlin/Boston: De Gruyter, 448–458.
- Holly, Werner (2018): Intermedialität von Frames in einer Polit-Talkshow. In: Ziem, Alexander; Inderelst, Lars & Wolf, Detmer (Hrsg.): *Frames interdisziplinär: Modelle, Anwendungsfelder, Methoden*. Düsseldorf: University Press, 331–352.
- Holly, Werner; Kühn, Peter & Püschel, Ulrich (1986): *Politische Fernsehdiskussionen. Zur medienspezifischen Inszenierung von Propaganda als Diskussion*. Tübingen: Niemeyer.

- Holly, Werner; Kühn, Peter & Püschel, Ulrich (Hrsg.) (1989): *Redeshows. Fernsehdiskussionen in der Diskussion*. Tübingen: Niemeyer.
- Holly, Werner; Püschel, Ulrich & Bergmann, Jörg (Hrsg.) (2001): *Der sprechende Zuschauer. Wie wir uns Fernsehen kommunikativ aneignen*. Wiesbaden: Westdeutscher Verlag.
- Holly, Werner & Schwitalla, Johannes (1995): „Explosiv – Der heiße Stuhl“. Zur Inszenierung von ‚Streitkultur‘ im kommerziellen Fernsehen. In: Neumann-Braun, Klaus & Müller-Doohm, Stefan (Hrsg.): *Kulturinszenierung – Kultureffekte*. Frankfurt am Main: Suhrkamp, 59–88.
- Jäger, Ludwig (2002): Transkriptivität. Zur medialen Logik der kulturellen Semantik. In: Jäger, Ludwig & Stanitzek, Georg (Hrsg.): *Transkribieren. Medien/Lektüre*. München: Fink, 19–41.
- Jäger, Ludwig (2014): Audioliteralität. Eine Skizze zur Transkription des Hörbuchs. In: Binczek, Natalie & Epping-Jäger, Cornelia (Hrsg.): *Das Hörbuch. Praktiken audioliteralen Schreibens und Verstehens*. München: Fink, 231–253.
- Jäger, Ludwig (2020): Audioliteralität: zur akroamatischen Dimension des Literalen. In: Binczek, Natalie & Wirth, Uwe (Hrsg.): *Handbuch Literatur und Audiokultur*. Berlin/Boston: De Gruyter, 61–84.
- Jarren, Otfried & Fischer, Renate (2021): Die Plattformisierung von Öffentlichkeit und der Relevanzverlust des Journalismus als demokratische Herausforderung. In: Seeliger & Seignani (Hrsg.): 365–382.
- Jewitt, Carey & Oyama, Rumiko (2001): Visual Meaning: A Social Semiotic Approach. In: van Leeuwen, Theo & Jewitt, Carey (Hrsg.): *Handbook of Visual Analysis*. London u.a.: Sage, 134–156.
- Klein, Josef (1990): *Elefantenrunden ‚Drei Tage vor der Wahl‘. Die ARD-ZDF-Gemeinschaftssendung 1972–1987*. Baden-Baden: Nomos.
- Klein, Josef (1997): Kategorien der Unterhaltsamkeit. Grundlagen einer Theorie der Unterhaltung mit kritischem Rückgriff auf Grice. In: Rolf, Eckard (Hrsg.): *Pragmatik. Linguistische Berichte. Sonderheft 8*. Wiesbaden: Westdeutscher Verlag, 176–188.
- Klein, Josef (2019): *Politik und Rhetorik. Eine Einführung*. Wiesbaden: Springer VS.
- Mann, Thomas (1975): *Deutsche Hörer! Fünfundfünfzig Radiosendungen nach Deutschland*. Leipzig: Insel.
- Mann, Thomas (1988/2003): *Deutsche Hörer: BBC-Reden 1941 bis 1945*. München: der Hörverlag (CD).
- Meyer, Thomas (2001): *Mediokratie. Die Kolonisierung der Politik durch die Medien*. Frankfurt: Suhrkamp.

- Meyer, Thomas; Ontrup, Rüdiger & Schicha, Christian (2000): *Die Inszenierung des Politischen. Zur Theatralität von Mediendiskursen*. Wiesbaden: Westdeutscher Verlag.
- Papacharissi, Zizi (2015): *Affective Publics. Sentiment, Technology, and Politics*. Oxford: Oxford. University Press.
- Postman, Neil (1985): *Amusing Ourselves to Death. Public Discourse in the Age of Show Business*. New York: Viking Penguin.
- Reckwitz, Andreas (2017): *Die Gesellschaft der Singularitäten. Zum Strukturwandel der Moderne*. Berlin: Suhrkamp.
- Ricci Bitti, Pio E. (2014): Facial Expression and social interaction. In: Müller, Cornelia; Cienki, Allen; Fricke, Ellen; Ladewig, Silvia H.; McNeill, David & Bressemer, Jana (Hrsg.): *Body – Language – Communication. An International Handbook on Multimodality in Human Interaction*. Band 2. Berlin u.a.: De Gruyter Mouton, 1342–1349.
- Sachs-Hombach, Klaus (2003): *Das Bild als kommunikatives Medium. Elemente einer allgemeinen Bildwissenschaft*. Köln: Halem.
- Schirren, Thomas (2009): Textaufbau und Redeteilschema (partes orationes). In: Fix, Ulla; Gardt, Andreas & Knape, Joachim (Hrsg.): *Rhetorik und Stilistik*. 2. Halbband. Berlin/New York: De Gruyter, 1515–1528.
- Schwarzenegger, Roger-Gérard (1980): *Politik als Schaugeschäft. Moderne Strategien im Kampf um die Macht*. Düsseldorf/Wien: Econ.
- Schwarzenegger, Arnold (2022): Schwarzenegger tells Putin „to stop this war“. Video unter <https://www.youtube.com/watch?v=fWCIXZd9c78> (29.08.2023).
- Seeliger, Martin & Sevnigani, Sebastian (Hrsg.) (2021): *Ein neuer Strukturwandel der Öffentlichkeit. Leviathan*, Sonderband 37.
- Staab, Philipp & Thiel, Thorsten (2021): Privatisierung ohne Privatismus. Soziale Medien im digitalen Strukturwandel der Öffentlichkeit. In: Seeliger & Sevnigani (Hrsg.): 277–297.
- Thomä, Dieter (2019): *Warum Demokratien Helden brauchen. Plädoyer für einen zeitgemäßen Heroismus*. Berlin: Ullstein.
- Ueding, Gert & Steinbrink, Bernd (2011): *Grundriss der Rhetorik. Geschichte, Technik, Methode*. [5. Aufl.]. Stuttgart/Weimar: Metzler.
- Zuboff, Shoshana (2018): *Das Zeitalter des Überwachungskapitalismus*. Frankfurt a.M./New York: Campus.