

3 Theoretische Grundlagen und Analyserahmen

3.1 Parainteraktion

In diesem Kapitel soll zunächst die Notwendigkeit eines Begriffs von Parainteraktivität in Abgrenzung zu einem Interaktivitätsbegriff deutlich gemacht werden, um im Anschluss Überlegungen hinsichtlich Parainteraktivität aus dem Bereich der gesprächsanalytischen Linguistik, aber auch aus medienwissenschaftlicher Perspektive heranzuziehen. Dabei sollen vor allem die Vorteile des Begriffs der Parainteraktion gegenüber dem der parasozialen Interaktion deutlich werden sowie der Vorschlag, die Begriffe Parainteraktion und Pseudokommunikation terminologisch für die Bezeichnung unterschiedlicher Phänomene zu unterscheiden.

3.1.1 INTERAKTIVITÄT UND PARAINTERAKTIVITÄT

Durch die technisch-medialen Spezifika von YouTube haben Nutzer*innen wie bereits beschrieben verschiedene Möglichkeiten, auf YouTube-Videos zu reagieren. Genau dazu rufen die YouTuber*innen sie in Calls to Action oftmals explizit auf. Darüber hinaus lassen sich weitere Formen der Ansprache finden, die weniger die technischen Möglichkeiten der Reaktion und Partizipation thematisieren, sondern vielmehr direkte Beteiligungsmöglichkeiten der User*innen in einer suggerierten Gesprächssituation mit YouTuber*innen realistisch erscheinen lassen. Obwohl die Videos eigentlich monologischer Struktur sind und nicht die Möglichkeit einer direkten Wechselseitigkeit bieten, lassen sich in ihnen durchaus dialogische Elemente und kommunikative Marker finden, die auf eine Interaktionssituation schließen lassen. Diese Interaktion implizierenden kommuni-

kativen Praktiken auf der Produzent*innen-Seite sollen unter dem Begriff der *Parainteraktivität* beschreibbar gemacht werden. Der Begriff findet bisher besonders in den Bereichen der Medienlinguistik, -soziologie und -wissenschaft Verwendung (vgl. Burger/Luginbühl 2014; Vorderer 1996). Um Parainteraktivität zu definieren, sind zunächst eine klare terminologische Abgrenzung zum Begriff der Interaktivität sowie eine genaue Ausdifferenzierung des Begriffs notwendig.

Orientiert wird sich deshalb an dem Interaktionsbegriff nach Imo (2013, 2016). Demnach handelt es sich nur um Interaktion, wenn die Kriterien der Situationsgebundenheit sowie der Sequenzialität erfüllt sind (vgl. Imo 2013: 55ff). Das bedeutet, dass Gesprächsteilnehmende für einen bestimmten Kontext kollaborativ eine Kommunikationssituation herstellen und sich darüber hinaus im Sinne einer sequenziellen Interaktionsstruktur wechselseitig aufeinander beziehen können (vgl. ebd.). Bezogen auf YouTube sind diese Kriterien nicht erfüllt, da durch die YouTuber*innen eine gemeinsame Gesprächssituation sowie die Möglichkeit der direkten Einflussnahme lediglich imitiert bzw. fingiert wird (so haben z.B. Fragen an die User*innen für den Gesprächs- respektive Videoverlauf keine sequenzielle Funktion). Die Möglichkeit einer Reaktionsreaktion ist nicht gegeben, weshalb sich im Fall von YouTube-Videos zwar, begründet durch das Potenzial zur Anschlusskommunikation in Form von Kommentaren o.Ä., von Dialogizität (vgl. Bachtin 2005; Linell 2003) sprechen lässt, ein Interaktionsbegriff jedoch irreführend wäre.

Einen umgekehrten Ansatz wählten Schramm, Hartmann und Klimmt, die in ihrem Beitrag von 2002 „Desiderata und Perspektiven der Forschung über parasoziale Interaktion und Beziehungen zu Medienfiguren“ aufzeigen und neben konzeptionellen Herausforderungen des Beziehungs- und Personabegriffs auf die „Unschärfe des Interaktionsbegriffs“ (ebd.: 438) hinweisen und ihn von einem Kommunikationsbegriff abgrenzen. Sie orientieren sich dabei an Definitionen (von Fischer und Wiswede 2002 sowie Blickle 1999), die Interaktion als globaler verstandener Austausch betrachten und Kommunikation als „Spezialfall der Interaktion“ (ebd.), der die Kriterien der wechselseitigen Einflussnahme im Sinne Imos erfüllt. Ausgehend von diesen Definitionen schlagen sie vor, „zur Beschreibung von PSI [= parasozialer Interaktion] eher den Begriff der ‚Interaktion‘ anzuwenden, da soziale Interaktionen auch dann gegeben sind, wenn die Interaktionen nur auf mutmaßliche Reaktionen der jeweils anderen erfolgen“ (ebd.). Während dieser Vorschlag unter dem zugrundeliegendem Interaktionsverständnis zunächst Sinn ergibt, würde die Notwendigkeit, von einer *parasozialen* Interaktion bzw. *Parainteraktion* zu sprechen, hinfällig, da unter dieser Prämisse jede Interaktion auch parainteraktiv sein könnte. Dementsprechend wird hier statt-

dessen ein engerer Interaktionsbegriff genutzt, der in Hinblick auf die Möglichkeiten des Untersuchungsgegenstands terminologisch erweitert wird.

Ein weiterer Begriff aus der klassischen Konversationsanalyse, der bei der Ansprache von Rezipient*innen in den Blick gerät, ist der des von Sacks (1992) sowie Sacks, Schegloff und Jefferson (1974) geprägten *Recipient Designs*. Darunter wird das Phänomen beschrieben, dass Sprecher*innen Überlegungen bezüglich ihrer Auswahl des Gesagten an den konkreten Hörer*innen orientieren. Der Begriff bezieht sich dabei – wie auch Imos Interaktionsbegriff – auf die gesprächsanalytische Grundannahme, dass Gespräche interaktionell und kooperativ von den Gesprächsteilnehmenden aufgebaut werden. Deppermann und Schmitt (2016) sprechen daher von Partnerorientierung und weisen somit auf Erwartungen sowie Erwartungserwartungen der Teilnehmenden untereinander hin, die in Gesprächen berücksichtigt werden und dialogische Funktionen erfüllen. Interaktive Marker, die dies indizieren, lassen sich auch in den Monologen von YouTuber*innen finden, obwohl diese, wie beschrieben, rein funktionell betrachtet für die Gesprächssituation nicht notwendig sind. Ähnlich zu den von Deppermann und Schmitt beschriebenen Massenmedien bleibt auch im Fall von YouTube „Partnerorientierung notwendig imaginär“ (ebd.: 392).

Der Begriff des *Audience Design* kommt vor dem Hintergrund des hier behandelten medialen Gegenstands ebenfalls für die Beschreibung der kommunikativen Strategien von YouTuber*innen in den Blick. Dieser Begriff stammt aus dem Bereich der Soziolinguistik und wurde von Bell (1984) anhand einer vergleichenden Untersuchung von Radio-Nachrichten in sozial-ökonomisch unterschiedlich starken Gebieten beschrieben. Seine Studie zeigte, dass Medienschaffende sich am antizipierten Publikum bzw. der antizipierten Zuhörerschaft bezüglich des Vorwissens, der inhaltlichen Erwartungen sowie des sprachlichen Stils orientieren (vgl. ebd.). Obwohl es bei den Formen der kommunikativen Ansprache der YouTuber*innen ebenfalls um eine gewisse Anpassung geht, beschreibt Audience Design eher stilistische und soziale Kategorien, deren Analyse zwar interessante Aufschlüsse bieten kann, aber nicht zwangsläufig Interaktion implizierende Mechanismen beschreibbar macht. Die Konzepte des Recipient respektive Audience Design werden also im Zuge dieser Untersuchung nicht systematisch analytisch berücksichtigt, laufen jedoch bei Parainteraktion nebenher.

Da es sich bei YouTube, zumindest im Fall von einzelnen YouTuber*innen, um *One-to-many*-Kommunikation handelt, sind Ansprachen auf YouTube mehrfachadressiert. Mit dem Begriff der *Mehrfachadressierung* beschreibt Kühn, dass

sich Sprachhandlungen, die an mehrere Adressat*innen⁵⁹ gerichtet sind, „hinsichtlich ihres Handlungsgehalts/pragmatischen Gehalt adressatenspezifisch unterschiedlich interpretieren“ (1995: 1) lassen. Dies kann sowohl absichtlich als auch unabsichtlich geschehen oder in Kauf genommen werden, wobei es Kühn nicht darum geht, diese Unterscheidung jeweils im Einzelfall zu treffen, auch da Mehrfachadressierungen unbewusst getätigt werden können und es somit zu Überschneidungen und Grenzfällen kommen kann (vgl. ebd.: 62). Diese (idealtypische) Unterscheidung von *Adressierungsarten* ist besonders dann für YouTube-Videos relevant, wenn mehrere YouTuber*innen in einem Video mitmachen und somit Adressierungen sowohl sie als auch die Zuschauenden betreffen. Auch in den Fokusgruppengesprächen kommen Mehrfachadressierungen sowohl hinsichtlich der anderen Anwesenden im Raum als auch der an die YouTuber*innen gerichteten Aussagen in den Blick. Dabei können Mehrfachadressierungen für verschiedene Adressat*innen unterschiedliche Bedeutung haben, also als *polyvalent* beschrieben werden (vgl. ebd. 51). Sie können außerdem durch eine explizite oder implizite Ausdrucksform realisiert werden. Adressat*innen werden somit entweder explizit durch verbale direkte oder indirekte Ansprache, oder averbale körpersprachliche Mittel adressiert (vgl. ebd.: 106). Eine implizite Adressierung liegt vor, wenn die adressierende Person zwar „keinerlei Kontaktsignale verwendet“ (ebd.), jedoch „z.B. Themen anspricht, Beziehungen aufbaut oder Handlungen ausführt, von denen [sie] annimmt, daß sich anwesende oder erreichbare Adressaten betroffen fühlen“ (ebd.: 49). Eine implizite Adressierung ist für den Kontext von YouTube-Videos sowohl bezogen auf textsortenspezifische Themen als auch auf die Beziehungsgestaltung zwischen YouTuber*innen und Zuschauenden relevant, obwohl erstere aufgrund der Heterogenität der Zuschauerschaft vor einem „Adressierungsdilemma“ (ebd.: 90) stehen.

Für den Kontext von Werbespots spricht bereits Ayaß von „Adressivität“ und meint damit „Phänomene direkter Hinwendungen an das Publikum“, die sich in „sprachlicher Art“, „in direkten verbalen Ansprachen in Anredeformen („Sie“, „Ihnen“, usf.)“ oder in Fragen ausdrücken können (2001: 203). Neben verbaler Adressivität geht sie bereits auf andere Varianten und Verknüpfungen dieser ein, die als multimodal beschrieben werden können: So nennt sie als „visuelle“ Form der Adressivität die Sichtbarkeit der (in diesem Fall in den einzelnen Spots) sprechenden Personen, „in meist halbnaher Einstellung, frontal an das Publi-

59 Als Adressat*innen beschreibt Kühn – in Abgrenzung zu Begriffen, die eng gefasst eine Absicht der adressierenden Person voraussetzen oder in einem weiten Sinne alle Rezipierenden als Adressat*innen fasst – denjenigen der „sinnvoll und plausibel als von der Äußerung betroffen beschrieben werden kann“ (Kühn 1995: 47f).

kum“ (ebd.). Außerdem kann die Adressierung auch durch ein *voice over* geschehen (vgl. ebd.).

Bezogen auf die realen Partizipationsmöglichkeiten für Fernsehtexte stellte Hippel (1993) heraus „daß Rezipienten von Fernsehtexten keine Möglichkeit zur Verfügung steht, direkt Einfluß auf den Fortgang der Interaktion zu nehmen“ (ebd.: 128), was selbstverständlich auch für YouTube der Fall ist. Allerdings bewertet Hippel diese Tatsache als für die jeweiligen dialogischen Potenziale wenig ausschlaggebend. Er bezieht sich dabei auf Goffmans (1981) Überlegungen zu Fernsehkommunikation, wonach für eine Interaktion „nicht eine physikalische Wechselwirkung, sondern das gegenseitige Aufeinander-Bezug-Nehmen von Personen“ (Hippel 1993: 129) notwendig ist. Diese Annahme der gegenseitigen Bezugnahme erweitert er um die Überlegungen von Horton und Strauss (1957), nach denen eine Zuschauerreaktion nicht tatsächlich realisiert werden muss, sondern lediglich durch die Produzent*innen antizipiert werden kann: „entirely in imagination“ (Horton/Strauss 1957: 579).

Diese von der realen Rezeption losgelöste Perspektivierung auf Produktseite bietet erste Beschreibungsmöglichkeiten für die im Rahmen der Videos getätigten interaktiven Formen der Ansprache. In einem nächsten Schritt werden deshalb die Überlegungen von Horton und Strauss bzw. Horton und Wohl näher betrachtet, um den Begriff der Parainteraktivität konkreter fassen und für die Analyse anwenden zu können.

3.1.2 PARASOZIALE INTERAKTION UND PARASOZIALE BEZIEHUNG

Innerhalb der Forschungsrichtung zu Medienrezeption setzte Horton mit Wohl (1956) und Strauss (1957) den Grundstein für die Betrachtung der Effekte zwischen Medienakteur*innen und Zuschauer*innen. Ihre Überlegungen wurden insbesondere im Bereich soziologisch orientierter Medien- und Fernsehwissenschaft vielfach rezipiert und weitergeführt (vgl. Hippel 1992; Schramm et al. 2002).⁶⁰ Bezogen auf das Radio und das Fernsehen beschrieben sie, inwiefern

60 Ihr Ansatz wurde außerdem, ohne dass die Autoren in ihren Beiträgen darauf Bezug genommen hatten, oftmals mit dem Symbolischen Interaktionismus (vgl. Blumer 1973) in Verbindung gebracht. Krotz argumentiert, Horton und Wohls Konzept sei „eigentlich ein symbolisch-interaktionistisches Modell zur Analyse der Rezeption von Massenkommunikation [Herv. i. O.]“ (Krotz 1996: 73), und spielt damit darauf an, dass der Symbolische Interaktionismus Menschen als grundsätzlich sozial und interaktiv Wirklichkeit erzeugend versteht und ihr Verständnis von Medienkommunikation als (para-)soziales interaktives Handeln diesem Grundsatz entspreche.

Moderator*innen Face-to-face-Kommunikation nachbilden. Bezüglich des Untersuchungsgegenstands ist hierbei anzumerken, dass es in den damals als neu geltenden Massenmedien zu einer Herausbildung personenzentrierter Formate mit dem Fokus auf einzelne Moderator*innen kam (vgl. Hartmann 2016: 77), was auch vor dem Hintergrund der hier behandelten personenzentrierten YouTube-Inhalte interessant ist.

Horton und Wohl (1956) sowie Horton und Strauss (1957) zeigen, dass Medienpersonen kommunikativ den Eindruck erwecken, mit den Rezipierenden in einer Face-to-face-Situation zu sein. Bei den Rezipient*innen wird die Illusion einer realen Möglichkeit zur Reaktion erzeugt und dadurch der intuitive Eindruck erweckt, an einer direkten sozialen Gesprächssituation teilzunehmen. Sie wählen für dieses Phänomen den Begriff der *parasozialen Interaktion*,⁶¹ also „die Erzeugung einer Illusion von face-to-face-Beziehungen zwischen Zuschauern und ‚Darstellern‘“ (Hippel 1993: 129). Die Simulation einer Interaktionssituation betrifft dabei zunächst sowohl Medienproduzent*innen als auch die Zuschauenden:

The more the performer seems to adjust his performance to the supposed response of the audience, the more the audience tends to make the response anticipated. This simulacrum of conversational give and take may be called para-social interaction (Horton/Wohl 1956: 215).

Da Produzent*innen und Rezipient*innen nicht aufeinander reagieren können, ist der kommunikative Austausch also lediglich antizipiert. Parasoziale Interaktion wird darüber hinaus beschrieben als „one-sided, nondialectical, controlled by the performer, and not susceptible of mutual development“ (Horton/Wohl 1956: 215). Die innerhalb der TV- oder Radiosendung tatsächlich stattfindende Interaktion ist folglich einseitig von der Produktseite ausgehend. Darüber hinaus wird Reziprozität beidseitig lediglich angenommen, die analog zum Rundfunk auch in Bezug auf YouTube – zumindest im engeren Sinne – nicht vorliegt. Dass obgleich der objektiv fehlenden Wechselseitigkeit bei den Zuschauer*innen eine Interaktionssituation angenommen wird, liegt vor allem an distanzverringern den sprachlichen Mechanismen, durch die das Gesprochene für die Rezipierenden informeller und privater wirkt. Dadurch kann ein Gefühl der Nähe zu den Medienakteur*innen entstehen, was Horton und Wohl als „illusion

61 Sie grenzen den Begriff der parasozialen Interaktion von denen der *orthosozialen* Interaktion, die der Face-to-Face-Interaktion entspricht, und der *vicarious*, einer Stellvertreter-Interaktion ab. Allerdings liegt ihren Überlegungen kein konkret definierter Interaktionsbegriff zugrunde (vgl. Gleich 1996: 116).

of intimacy“ (Horton/Wohl 1956: 217) bzw. als „intimacy at a distance“ (Horton/Wohl 1956) beschreiben. Somit kann das Verhältnis zu den einzelnen Produzent*innen nachhaltig verstärkt werden: „Die gefühlte Nähe und Intimität zu einer Medienperson kann auch nach der Rezeption anhalten und sich zu einer geradezu freundschaftlichen Beziehung entwickeln“ (Hartmann 2016: 75).

Horton und Wohl (1956) sprechen in diesem Zusammenhang auch von einer *parasozialen Beziehung* als Bezeichnung für die interaktiv asymmetrische Beziehung zwischen Persona und Zuschauer*in innerhalb einer parasozialen Interaktionssituation. Allerdings lässt sich der Begriff auch über die einzelne Situation hinaus verstehen als längerfristig aufgebaute Bindung der Zuschauer*innen an die Persona (vgl. Hartmann 2016: 80), die über einzelne parasoziale Interaktionen entstehen und wie in einem „Spiralprozess“ (Hartmann 2017: 51) neue parasoziale Kontakte nach sich ziehen kann. Eine rezeptionsorientierte Untersuchung zu parasozialer Beziehung in Bezug auf das Fernsehen liefert die Studie von Gleich (1997). Er konzeptioniert parasoziale Beziehungen als „*Ergebnis* parasozialer Interaktionen, die in der jeweils aktuellen Rezeptionssituation stattfinden [Herv. i. O.]“ (Gleich 1996: 114) und klärt somit die Unterscheidung zwischen parasozialer Interaktion und parasozialer Beziehung, deren terminologische Ausdifferenzierung bei Horton/Wohl und Horton/Strauss in anschließender Forschung oftmals kritisiert oder missverstanden wurde (vgl. Krotz 1996: 73; Mikos 1996). Bezüglich des möglichen Ausmaßes parasozialer Beziehungen deutet Schramm (2007) darauf hin, dass auch Rezipierende diese unter Umständen nicht von sozialen Beziehungen unterscheiden können, was ein „Rezeptionsphänomen“ (Schramm 2007: 218) beschreibt, welches „pathologische Züge“ (ebd.) annehmen kann.

Für YouTube gestalten sich diese Beziehungen aufgrund seiner informellen Ästhetik und der vielseitigeren Möglichkeiten der Teilhabe der User*innen unmittelbarer und intensiver. Parasoziale Beziehungen gehen dabei von den User*innen aus, sie sind aber genauso abhängig von den durch die YouTuber*innen formulierten Näheangeboten. Das Gefühl der Nähe oder die Intimität gegenüber den Medienpersonen wird in ihren kommunikativen Ansprachen evoziert. Dies geschieht im Sinne Horton und Wohls, indem den Zuschauenden eine bestimmte Antwortrolle, eine *answering role*, zugesprochen wird, die wiederum auf der Produktseite Berücksichtigung findet. Hartmann weist darauf hin, dass das Annehmen dieser Rolle parasoziale Interaktion bedingt: „Nur wenn das Publikum die angebotene Rolle annimmt, tritt es in das Interaktionsgefüge ein und nur dann kann sich die parasoziale Interaktion nach Horton und Wohl voll entfalten“ (Hartmann 2016: 79). Somit ist es für die Produktseite erst einmal irrelevant, inwiefern die Rezipient*innen tatsächlich reagieren, oder ob die Reaktion „en-

tirely in imagination“ (Horton/Strauss 1957: 579) ist. Die Rezeptionsseite muss jedoch die interaktiven Angebote wahrnehmen, um sich in der ‚gemeinsamen‘ Gesprächssituation wiederzufinden.

Diese Hervorhebung Hartmanns verdeutlicht seinen eher rezeptionsorientierten Fokus. Aus empirisch motivierter Perspektive machen Hartmann, Schramm und Klimmt (2004a) einen Vorschlag für ein Analysemodell, welches die von Horton und Wohl ausgeführten Überlegungen praktisch anwendbar machen soll. Dazu schlagen sie ein Zwei-Ebenen-Modell zur Beschreibung von parasozialer Interaktion bei den Zuschauer*innen vor. Sie gehen dabei von zwei „Intensitätsebenen“ (Hartmann/Schramm/Klimmt 2004a: 30) der personaorientierten parasozialen Interaktion aus. Diese weisen die prototypischen Ausprägungen „einer oberflächlich-schwachen Beschäftigung mit einer Persona (so genannte *Low-Level-PSI*) und einer intensiv-starken Beschäftigung (so genannte *High-Level-PSI*) [Herv. i. O.]“ (ebd. 31) auf. Außerdem schließen sie psychologische Kategorien bei der Beschreibung der Abläufe auf Rezeptionsseite mit ein, sodass sie *perzeptiv-kognitive*, *affektive* sowie *konative* Teilprozesse beschreiben (vgl. ebd. 30). Diese Berücksichtigung verschiedener Teilprozesse verdeutlicht den weit gefassten Interaktionsbegriff respektive ein weitgreifendes Verständnis davon, was als parasoziale Interaktion gefasst werden kann. So beschreiben die Autoren bereits die bloße Wahrnehmung und ein Begreifen des Geschehens als perzeptiv-kognitive Teilprozesse parasozialer Interaktion, wenngleich sie diese als *Low-Level-PSI* einordnen (vgl. ebd.: 33). Derartige Prozesse – obgleich sie relevante Voraussetzungen für die Aufnahme parainteraktiver Elemente sowie eigenes parainteraktives Engagement bilden – werden im Rahmen dieser Untersuchung nicht als parasoziale Interaktion bzw. Parainteraktion gehandhabt. Ebenso werden diejenigen perzeptiv-kognitiven Prozesse, die als *High-Level-PSI* klassifiziert wurden – wie eine fokussierte Aufnahme der Inhalte, Bemühungen, das Geschehen und die Motive der Persona nachvollziehen zu können, Bewertungen und Vergleiche (vgl. ebd.) – sowie affektive Teilprozesse analytisch nicht berücksichtigt. Stattdessen geraten für einen medienlinguistischen Analysezugang die konativen Aktivitäten in den Blick, zu denen die Autoren beispielsweise Zurufe, Beschimpfungen oder Gestikulieren zählen (vgl. ebd.: 35). Allerdings sei zu beachten, „dass gerade solche deutlich erkennbaren Aktivitäten oftmals nur oberflächlich an die Medienpersonen adressiert werden, sich in Wahrheit aber an andere Teilnehmer der Rezeptionssituation [...] richten“, sodass sie „nicht als *PSI mit* einer Medienperson, sondern als interpersonale Kommunikation *über* die Medienperson [Herv. i. O.]“ (ebd.: 35) rekonstruiert werden müssten. Jedoch könne auch bei der Rezeption allein nicht ausgeschlossen werden, dass konative Aktivitäten weniger an die Medienpersonen, als vielmehr an die Rezipierenden

selbst gerichtet werden und nicht-kommunikative Funktionen erfüllen (vgl. ebd.). Die Autoren verweisen daher selbst darauf, dass sich eine genaue Abgrenzung konativer PSI theoretisch und empirisch schwierig gestaltet (vgl. ebd.). Zu relevanten konativen Aktivitäten zählen sie „[m]otorische Verhaltensweisen“, die die sinnliche Wahrnehmung der Medienperson gewährleisten, „körperliche, vornehmlich mimische und gestische Verhaltensweisen, die aus echten sozialen Situationen entlehnt und auf die Persona gerichtet sind“ sowie „verbale Äußerungen zur Persona“ (ebd.: 35f). Diese Aktivitäten werden eher auf der High-Level-PSI-Ebene angeordnet, da die Autoren davon ausgehen, dass die Medienperson andernfalls „nicht für wichtig genug erachtet“ (ebd.: 36) würde, um verbale oder körperliche Handlungen hervorzurufen.⁶² Allerdings geschieht keine Systematisierung spezifischer verbaler oder körperlicher Handlungen während der Rezeption. Auf der anderen Seite wird jedoch angenommen, dass PSI-Aktivitäten der Zuschauenden von verschiedenen Faktoren der Medienpersona abhängen. Neben *Obtrusivität* (= verhältnismäßige Auffälligkeit der Persona) und *Persistenz* (= Dauer ihrer Erscheinung), die für Videos-on-Demand auf YouTube vernachlässigt werden können, nennen die Autoren zwei weitere Kategorien, die maßgeblich erscheinen: 1) die *Adressierung*, und 2) die von der Medienpersona ausgehende *Motivation zur Auseinandersetzung* (vgl. ebd.: 37ff). Unter Adressierung fassen sie direkte oder indirekte verbale und non-verbale Bezugnahmen (vgl. ebd.: 38), oder auch Blicke und Gesten in die Kamera sowie das Zoomen der Kamera auf Medienpersonae (vgl. ebd.: 39) und beschreiben damit Kategorien, die im weiteren Verlauf der Untersuchung für die Analyse als Modalitäten konzeptioniert werden (s. Kapitel 3.2). Motivierend für eine tiefere Auseinandersetzungen seien unter anderem der erste Eindruck über die physische Attraktivität (der Persona und ihres Umfelds) (vgl. ebd.).

Grundsätzlich ist dieser Analysezugang deutlich sozial- und medienpsychologisch ausgelegt, was auch durch die Beachtung von Kategorien wie Identifizierung und Identitätsbildung deutlich wird. So führen die Autoren aus, dass Zuschauer*innen bei der Rezeption durchaus Identitätsarbeit betreiben und daher Prozesse parasozialer Interaktion „am relevantesten für die Identitätsbildung von Jugendlichen“ (Schramm/Hartmann 2010: 216) sind. Diese Feststellung

62 Im Rahmen der Fokusgruppengespräche wird sich herausstellen, inwiefern die Jugendlichen während der Rezeption pseudokommunikative Aktivitäten auch bei YouTube*innen tätigen, die für sie nach eigener Angabe privat nicht relevant sind. Derartige High-Level-Aktivitäten scheinen also auch dann aufzutreten, wenn Rezipierende dem Gesehenen situativ Relevanz zusprechen. Darüber hinaus ist anzunehmen, dass sich Unterschiede bei der Rezeption von Fernsehen und YouTube-Videos zeigen, die wiederum jew. parainteraktive Elemente aufgreifen.

richtet die Aufmerksamkeit vor allem auf YouTube-Videos mit jugendlicher Zielgruppe und in ihnen beobachtbare Mechanismen parasozialer Interaktion. Auch vor dem Hintergrund möglicher werbender Elemente ist diese Feststellung wichtig, da durch besagte Mechanismen im Rahmen von Identifikationsprozessen auch die Meinungen der YouTuber*innen bezüglich der gezeigten Produkte potenziell von den User*innen übernommen werden können. Hier wird es also darum gehen, die Annahmen bezüglich parasozialer Interaktionsverfahren auf Produktseite aufzugreifen und zu beschreiben.

Aus Perspektive der Rezeptionsseite untersuchten außerdem Hartmann und Goldhoorn (2011) das Erleben von parasozialer Interaktion mit Fernsehpersonae. Dabei konzipierten sie in Anlehnung an die „Parasocial Interaction scale“ (ebd.: 1105) von Rubin/Perse/Powell (1985) eine „Experience of Parasocial Interaction (ESPI) scale“ (ebd.: 1104) um anhand derer zu veranschaulichen, unter welchen Voraussetzungen Zuschauende eine mehr oder weniger intensive parasoziale Erfahrung erleben. Im Gegensatz zu dem Ansatz der PSI scale, die wie die Autor*innen bemängeln in erster Linie die „users friendships toward media performers“ (ebd.: 1105) ‚misst‘, bei der ein Begriff von parasozialer Interaktion zugrunde liegt, der eher die parasoziale Beziehungsebene zwischen Rezipient*in und TV-Person meint, stellt ihr Ansatz in den Mittelpunkt, inwiefern sich Rezipierende in eine scheinbare Interaktion involviert fühlen (vgl. ebd.). *Experience* definieren sie dabei als „an immediate feeling or impression that results from users’ automatic mindreading activities“ (ebd.: 1107) und meinen mit letzteren Aktivitäten die ‚gegenseitige‘ aufmerksame Wahrnehmung und Anpassung von Verhaltensweisen, wie etwa der Körperhaltung, der Gestik, Gesichtsausdrücke oder auch des Sprechrhythmus (vgl. ebd.). Eigens für Ihre Untersuchung drehten sie Videoclips, in denen die Person ihre körperliche Ausrichtung von und zu der Kamera sowie das Sprachregister variiert, und zeigte diese Variationen in unterschiedlichen Gruppen, worauf in diesen anschließend die Skalen ausgefüllt wurden.⁶³ Sie stellten heraus, dass Rezipierende eine parasoziale Erfahrung als intensiver beschreiben, wenn sie sowohl verbal als auch körperlich – also analog zu körperlichen Ausdrucksressourcen, siehe Kapitel 3.2.2 – angesprochen werden (vgl. ebd.). Darüber hinaus waren die Attraktivität der Medienpersonen sowie das (selbst eingeschätzte) Vermögen der Zuschauenden, Perspektiven anderer aufzugreifen, ausschlaggebend. Die Autor*innen geben außerdem an, dass, je stärker die parasoziale „experience“ wahrgenommen wurde, auch das „com-

63 Zusammen mit der *EPSI scale* wurden verkürzte Versionen der *PSI scale* sowie der von Schramm/Hartmann (2008) entwickelten „Parasocial Process scale“ (ebd.: 1111) zum Ausfüllen gereicht, bei denen die Teilnehmenden Aussagen in sieben Punkten im Sinne einer Likert-Skala zustimmen oder widersprechen konnten.

mitment in social norms“ (ebd.) höher war, was in Bezug auf eine Analyse von Parainteraktion in Hinblick auf Gender insofern relevant ist, als dieser Befund die Annahme nahe legt, dass erfolgreiche Parainteraktion auch das Annehmen sozialer Gendernormen begünstigen kann.

Auch in Bezug auf YouTube wurden ähnliche Studien durchgeführt, die auf Fragebögen mit Skalen zurückgreifen und sich an der Parasocial-Interaction-Scale von Rubin und Kolleg*innen orientieren. Diese befassen sich vor allem mit der Beziehungsebene zwischen Zuschauenden und YouTuber*innen. So fragen etwa Rihl und Wegener (2019) danach, ob die Nutzung der Partizipationsmöglichkeiten der Plattform (Liken/Disliken, Kommentieren, das Verfolgen der Kommentare anderer sowie Reaktionen der YouTuber*innen auf eigene Kommentare) Auswirkungen auf die Stärke der parasozialen Beziehungen hat. Dabei orientieren sie sich an der bereits erwähnten Forschung von Gleich, der in seinen prägnanten Studien (1995, 1996, 1997) die PSI-Scale für den deutschsprachigen Raum abgewandelt und parasoziale Interaktion mit und vor allem parasoziale Beziehungen zu Fernsehpersonen untersuchte. Gleich hatte zuvor die ursprüngliche Skala hinsichtlich ihrer verschiedenen und nicht getrennten Analysekriterien kritisiert (vgl. Gleich 1996). Außerdem gibt er korrekterweise zu bedenken, dass die Skala zwar parasoziale Interaktion messen sollte, aber „nach meinem Dafürhalten eher Beziehungsaspekte zwischen Zuschauer und Fernsehpersonen“ (ebd.: 125) erfasse, auf die er sich in seiner Untersuchung bezieht. Unter Nutzung dieser von Gleich abgewandelten Skala verdeutlichen Rihl und Wegener zum einen im Rahmen einer Online-Befragung, inwiefern die Skala auf YouTube übertragbar ist, und beschreiben zum anderen, dass die Empfindung einer parasozialen Beziehung bei denjenigen User*innen höher ist, die öfter und intensiver Gebrauch der Partizipationsmöglichkeiten machen (vgl. Rihl/Wegener 2019: 562f), also auch von ihrer Seite aus an der Beziehung ‚arbeiten‘.

Rubin und McHugh (1987) erstellten außerdem ein Pfadmodell, mithilfe dessen der Entstehungsprozess einer parasozialen Beziehung zu Fernsehpersönlichkeiten dargestellt werden sollte. Ausgehend von diesem Modell schlugen Kurtin et al. (2018) ein eigenes vor und untersuchen, inwiefern „YouTube Exposure“ eigenständig oder in Kombination mit „Social Attraction“, „Task Attraction“ und/oder „Physical Attraction“ der YouTuber*innen zu „Parasocial Relationships“ führen und Auswirkungen auf die „Relationship Importance“ für die User*innen haben kann (vgl. ebd.: 240). In einer Online-Umfrage befragten sie Studierende und stellten heraus, dass während *Social Attraction* und *Physical Attraction* sowohl positive Auswirkungen auf die grundsätzliche *YouTube Exposure* als auch auf die parasozialen Beziehungen hatten, sie jedoch keinen Effekt auf die *Relationship Importance* zu haben schienen (vgl. ebd. 244). Die Autorinnen er-

klären dies damit, dass physische und soziale Attraktivität durchaus zu einem kurzweiligen Engagement führen kann, alleinstehend jedoch nicht ausreicht für eine beständige Beziehung. In diesem Zusammenhang weisen sie auf fehlende Präzision bei der terminologischen Abgrenzung parasozialer Beziehungen hin, da anstelle dieser teilweise eher einzelne parasoziale Interaktionen erfragt wurden (vgl. ebd.: 247). Beide Studien verweisen somit noch einmal auf die Notwendigkeit, terminologisch streng zwischen parasozialer Beziehung und parasozialer Interaktion zu unterscheiden, sodass Rihl/Wegener (2019: 556) in Anschluss an Vorderers (1998) Vorschlag und im Sinne Hartmanns (2016, 2017) betonen, dass sich parasoziale Interaktion auf den Akt der Mediennutzung beschränkt, während parasoziale Beziehungen einen über-situativen, prozesshaften Charakter aufwiesen (vgl. Krotz 1996) und durch mehrere, sich wiederholende parasoziale Interaktionen entstehen können.

Obgleich die Studien in ihren Ergebnissen hinsichtlich parasozialer Interaktion unkonkret sind, zeigen sie dennoch für diese Untersuchung relevante Tendenzen: Zum einen wird die Relevanz der Nutzung der Partizipationsmöglichkeiten auf YouTube deutlich, die sich nicht nur auf den Algorithmus bezieht, sondern auch auf die Beziehungsebene zwischen YouTuber*innen und Zuschauenden und somit Calls to Action in Videos eine größere Bedeutung zukommen lässt. Zum anderen wurde neben der sozialen auch die physische Attraktivität in den Vordergrund gerückt, die besonders bei Beauty-Videos eine gesonderte Stellung einnimmt.

Auch zu Gender und parasozialen Beziehungen im Kontext von YouTube wurde bereits geforscht. Chen (2016) analysierte beispielsweise die digitale Selbstrepräsentation von taiwanischen Amateur-YouTuber*innen. Parasoziale Beziehung wird hier aus der Perspektive der YouTuber*innen betrachtet. Grundlage der Untersuchung war ein dreifacher Datenzugang, bei dem inhaltsbasierte Analysen von Videos mit Interviews mit rund 45 YouTuber*innen sowohl face-to-face als auch online kombiniert wurden, in denen sich u.a. über die analysierten Videos unterhalten wurde: „Participants were asked questions about the purpose of their performances on YouTube videos, what they communicate, how they chose materials and stories“ (Chen 2016: 240). Die Interviews wurden inhaltsanalytisch durch eine Codierung im Sinne der Grounded Theory ausgewertet und offenbarten, dass bereits die Anzahl der Aufrufe einzelner Videos als Ausdruck von Wertschätzung empfunden wird: „When viewers form positive attitudes towards them, they gain confidence to develop parasocial relationships with viewers“ (ebd.: 248). Aber auch die Kommentare der User*innen würden die Creators freuen und sie motivieren (vgl. ebd.: 248f). Hier und da kam es auch zu beständigerem Austausch online zwischen YouTuber*in und User*in (vgl.

ebd.) Grundsätzlich zeigte sich, dass sich YouTuber*innen über die parasoziale Ebene mit User*innen anfreunden und somit eine Beziehung aufbauen. Die damit einhergehende Selbstpräsentation betrifft auch Gender, z.B. bei YouTuber*innen, die online eine „digital female identity“ (ebd.: 243) annehmen wollen und sich dafür an kulturell etablierten Indexen für Gender orientieren (z.B. lange Haare und Make-up). Dies wird unter dem Begriff der „gender-switching strategy“ (ebd.: 247) gefasst, die sowohl auf großen Anklang, als auch bei einigen YouTuber*innen auf negative Kommentare der User*innen gestoßen ist (vgl. ebd.: 247f).

Tolbert und Drogos nähern sich den Zusammenhängen zwischen parasozialen Beziehungen und Gender aus der Rezeptionssicht. Dazu erhoben sie Fragebögen bei rund 161 *Tweens* im Alter von neun bis zwölf Jahren (vgl. ebd.: 5), in denen u.a. die Faktoren parasoziale Beziehungen, *wishful identification* („the desire to be like a media personality“ (ebd.: 3)) und verschiedene Attribuierungen (*smart, successful, attractive, funny, popular, violent* (vgl. ebd.: 6)) anhand einer Likert-Skala und in Bezug auf die Lieblings-YouTuber*innen erfragt wurden (vgl. ebd.). Die Ergebnisse zeigten, dass sich parasoziale Beziehungen und „feelings of friendship“ (ebd.: 10) abzeichnen. Es wurde deutlich, dass PSB sowie auch WI zum einen mit der Rezeptionszeit der Videos der YouTuber*innen zusammenhängen, zum anderen von den wahrgenommenen Ähnlichkeiten zu ihnen (vgl. ebd.: 7f). Für *Tweens* ist YouTube außerdem eine Anlaufstelle, um mehr über (ihr) Gender herauszufinden und damit einhergehende Normen und Verhaltensweisen (vgl. Tolbert/Drogos 2019: 2, 10). Dementsprechend können YouTuber*innen eine Rolle in der Entwicklung der eigenen Genderidentität spielen (vgl. ebd.: 3), wobei gilt: „stereotypes on YouTube can limit tweens’ understanding of gender roles“ (ebd.:10). Die von den Autorinnen Befragten nannten YouTuber*innen des gleichen Genders als ihre Favoriten. Außerdem nannten sie unterschiedliche Attribute bei YouTuberinnen und YouTubern: „Female YouTubers were rated as more attractive and popular than male YouTubers, and male YouTubers were rated as more violent than female YouTubers“ (ebd.: 10). Auch das Potenzial einer *wishful identification* gestaltete sich geschlechtsspezifisch: „boys’ WI increased as YouTubers violence increased, whereas girls’ WI decreased as YouTubers violence increased“ (ebd.: 7).⁶⁴ Hinsichtlich parasozialer Beziehung demonstrieren die Autorinnen Zusammenhänge zwischen den Attributionen und dem Gender der Teilnehmenden. So würden PSB stärker ausfallen,

64 Die Autorinnen geben diesbezüglich zu bedenken: „Although these results may be cause for concern, we do not know the way violence is being portrayed in these YouTube videos, or the meaning that young viewers are making of this type of mediated violence“ (Tolbert/Drogos 2019: 10).

wenn die YouTuber*innen als *funny* angesehen werden, allerdings war sie bei den Jungen bei *successful* und *attractive* stärker und bei den Mädchen, wenn die YouTuber*innen als *popular* aufgefasst wurden (vgl. ebd.: 8).

Hinsichtlich konkreter und sprachlicher Ausprägungen parasozialer Interaktion in Online-Videos, hat Frobenius (2011) für Vlogs herausgestellt, inwiefern englischsprachige Vlogger*innen in den Eröffnungsphasen ihrer Videos ihre eigentlichen Monologe so ausgestalten, dass sie Parallelitäten zu Dialogen aufweisen. Dabei hebt sie besonders die Funktionen dieser Eröffnungsphasen heraus, die sich zum einen auf die Partizipation der Rezipierenden unter Nutzung der Möglichkeiten der Plattform, zum anderen auf die Konstruktion (sozialer) Identität der Vlogger*innen beziehen. Als besonders deutliche parasoziale Interaktion zu werten sind die von ihr beschriebenen Begrüßungsaktivitäten, die oft durch die Formulierungen mit „*hi, hey, hello*“ (ebd.: 818) zusammen mit der Adressierung „*everyone, everybody, guys, and YouTube and youtubers* [Herv. i. O.]“ realisiert werden. Neben den Begrüßungen beschreibt die Autorin auch anderen Strategien wie etwa (nervöses) Lachen und ungeplant wirkende Einstiege, die sie auf Nervosität aufgrund des fehlenden Gegenübers zurückführt (vgl. ebd.: 820).⁶⁵ Außerdem macht sie stark, inwiefern in der Nachbearbeitung Entscheidungen darüber getroffen werden, welche Aufnahme zu Beginn der Videos stattfindet, sodass sie die Relevanz einer multimodalen Betrachtung bereits andeutet (vgl. ebd.: 823).

In Bezug auf Beauty-Videos auf YouTube untersuchte Rasmussen (2018) die Zusammenhänge zwischen parasozialer Interaktion, der Reichweite von YouTuberinnen und ihrer Glaubwürdigkeit sowie der persuasiven Kraft von Produktvorstellungen. Sie fragte mithilfe von Videosichtungen und einer Erweiterung der Parasocial Process Scale danach, ob parasoziale Interaktion stattfindet, YouTuberinnen als glaubwürdig betrachtet wurden und die Teilnehmenden das Nachkaufen gezeigter Produkte in Erwägung ziehen würden (vgl. 2018: 285f). Zunächst zeichnete sich eine Korrelation ab zwischen den reichweitestärkeren YouTuberinnen und den Angaben zu „*feelings of knowing the celebrity*“, „*feeling as though the celebrity was the viewer’s friend*“ und „*enjoyment of the speaker’s personality*“ (ebd.: 288). Als „*knowledgeable*“ (ebd.) wurden auch YouTuberinnen mit weniger Reichweite eingeschätzt. Hinsichtlich der Kaufintention gaben jedoch mehrere bei weitreichendstarken YouTuberinnen an, dass sie „*would likely purchase a product reviewed by the speaker*“ (ebd.: 289) als bei geringerer Reich-

65 Unsicherheitsmarker kommen in den hier betrachteten Beauty- und FIFA-Videos teilweise ähnlich vor, inszenieren jedoch vielmehr die Nahbarkeit der YouTuber*innen, wie sich im Zuge der Analyse zeigen wird.

weite. Dass der Anteil allerdings bei etwas reichweiteschwächeren Kanälen immernoch beachtlich war, führt die Autorin vor allem auf die professionell wirkende Ausgestaltung der Videos und das Expertinnenwissen der YouTuberinnen zurück, sodass Glaubwürdigkeit in Verbindung mit Wissen als Kernbausteine des Influencermarketings beschreibt (vgl. ebd.: 289f.). Purnamaningsih und Rizkalla zeigen darüber hinaus, dass „physical attractiveness and attitude homophily have an impact on PSI and can affect viewers intention to purchase beauty products in the long term“ (2020: 13).

Für Let's Plays erforschten Kreissl, Possler und Klimmt (2021) parasoziale Beziehungen zu Let's-Player*innen. Sie fragen nach der „dimensional architecture of fans' PSR to Let's Players“ (2021: 1029) und danach, inwiefern „individual PSR dimensions related to fans' frequency of watching a Let's Player's content“ (ebd.) sind. Dafür führten sie angelehnt an vorausgehende Untersuchungen (z.B. Hartmann/Goldhoorn 2011; Hartmann/Klimmt 2005) eine Online-Umfrage zu deutschen Let's-Player*innen mit knapp 900 Teilnehmenden durch, von denen 80% männlich waren (vgl. ebd. 1030). Es stellte sich heraus, dass „entertainment, responsiveness, and liking“ (ebd.: 1033) die zentralen Merkmale für die Bildung einer parasozialen Beziehung sind, während die Let's-Player*innen weniger „as informative advisors or idols“ (ebd.) wahrgenommen wurden. Dementsprechend wurde auch das unterhaltende Potenzial relevanter gesetzt als das Können im Spiel. Im Gegensatz zu anderen Untersuchungen (vgl. Tolbert/Drogos 2019) waren hier Ähnlichkeiten zwischen Let's-Player*innen und User*innen nicht ausschlaggebend (vgl. ebd.). Stattdessen zeigte sich, dass alle Kategorien mit der Häufigkeit der Let's-Play-Rezeption in Verbindung standen, besonders *liking* und *entertainment*: „The more respondents perceived these dimensions as characteristic features of their PSR to Let's Players, the more days a week they watched videos from these YouTubers“ (ebd.: 1033f). Die Autoren geben allerdings zu bedenken, dass die Richtung der Einflussnahme nicht völlig klar ist und PSB genauso gut der Grund für eine hochfrequentierte Rezeption sein kann (vgl. ebd.: 1035). Darüber hinaus stelle sich eine gewisse Ambiguität in der Auffassung der User*innen von ihre liebsten Let's-Player*innen heraus:

While they do not consider them as far away, elevated, special, or worth idolizing, they simultaneously also do not perceive them as particularly similar to themselves. In this sense, many fans do neither strive to become more like their favorite Let's Players nor do they see these online stars as ‚one of them‘ (ebd.: 1035).

Die erwähnten Studien bilden keine erschöpfende Übersicht über die Erforschung von parasozialer Interaktion und parasozialen Beziehungen, die sich über

YouTube hinaus auch auf andere soziale Medien bezieht.⁶⁶ Stattdessen wurde deutlich, inwiefern verschiedene Ansätze sich mit dem Phänomen auseinandersetzen und dabei entweder die Produkt- oder die Rezeptionsseite in den Vordergrund rücken. Im Folgenden soll es ebenfalls um Untersuchungen gehen, die auf dem Grundgedanken Horton und Wohls aufbauen, diesen jedoch aus linguistischer Sicht problematisieren und ausbauen.

3.1.3 PARAINTERAKTION UND PSEUDOKOMMUNIKATION

Das Grundkonzept der parasozialen Interaktion von Horton/Wohl (1956) bzw. Horton/Strauss (1957) wurde in der Forschung inhaltlich weitergeführt, jedoch hinsichtlich seiner Begrifflichkeit problematisiert. Ayaß (1993: 35f) und Hepp (1998: 65) haben in diesem Zusammenhang besonders die terminologischen Unschärfen des Begriffs der parasozialen Interaktion kritisiert. Für sie fehle es dem Begriff an Präzision hinsichtlich dessen, was als ‚para‘ bezeichnet wird. Parasoziale Interaktion impliziere demzufolge, dass der soziale Anteil vorgetäuscht sei, wohingegen tatsächlich die Interaktion imitiert werde. Die Situation sei dabei jedoch durchaus als sozial zu beschreiben, weshalb der Ausdruck parasozial nicht zutreffen würde (vgl. Ayaß 1993: 35f). Ayaß kritisiert in diesem Zusammenhang, dass Rezipierende in der Medienforschung oftmals als passive Zuschauende aufgefasst worden sind, die Inhalte unbeteiligt aufnehmen, und ihnen somit ihr aktives Handeln aberkannt wurde. Stattdessen stellt sie heraus: „Was Menschen tun, die Fernsehen produzieren oder rezipieren, *ist* Handeln, und in den meisten Fällen auch soziales Handeln. Sogar der einsame Zapper ‚handelt‘ [Herv. i. O.]“ (Ayaß 1993: 36). Hepp verteidigt ebenfalls den Aspekt des Sozialen, da es durch die Rezipient*innen zu einer bestimmten Bezugnahme gegenüber der Medienakteur*innen kommen würde und derartige Beziehungskonstitutionen soziale Handlungen darstellen (vgl. Hepp 1998: 65). Ferner liege in einer möglichen Identifikation der Zuschauenden mit oder aber einer Ablehnung von verschiedenen angebotenen Rollen eine soziale Entscheidung (vgl. ebd. 65). Im Gegensatz zum sozialen Anteil sei es vielmehr der Aspekt der Interaktion, der in seiner Begrifflichkeit unscharf wäre. In diesem Sinne betont Ayaß, dass das Handeln von Medienpersonen oder Rezipierenden nicht scheinbar soziales, sondern „scheinbar interaktive[s] Handeln“ (Ayaß 1993: 36) sei, sodass die Formen der Herstellung vermeintlicher Wechselseitigkeit „soziologisch adäquat höch-

66 Z.B. wurde auch die parasoziale Beziehung zu digital kreierten ‚virtuellen Influencer*innen‘ untersucht (vgl. Stein/Breves/Anders 2022).

stens als *para-interaktiv* bezeichnet werden [Herv. i. O.]“ (ebd.) könnten. Auch Hepp versteht den Interaktionsbegriff in einer engeren Definition – nach der Reziprozität konstitutiv für Interaktion ist – und somit die Fernsehkommunikation als *para-interaktiv* (vgl. Hepp 1998: 65). Damit der soziale Aspekt nicht als trügerisch missverstanden wird, schlagen beide alternativ den Begriff der *Parainteraktion* vor, der das Phänomen parasozialer Interaktion beschreibt, jedoch terminologisch deutlicher fasst. Hartmann (2017) weist in Hinblick auf terminologische Fragen auf die u.a. von Krotz (1996) genutzte alternative Bezeichnungen der *Parakommunikation* von parasozialer Interaktion hin (obgleich sich Krotz eher auf parainteraktive Aktivitäten auf Rezeptionsseite bezieht). Er schildert, dass diese Bezeichnung eine Möglichkeit sei, den (para)interaktiven Charakter dieser Aktivitäten in der Tradition des symbolischen Interaktionismus hervorzuheben und sich gleichzeitig von der Forschung im Kontext des Uses-and-Gratification-Ansatzes abzugrenzen (vgl. ebd.: 45). Er plädiert jedoch dafür, neue Begriffe „sparsam“ zu verwenden und, da es „üblich“ sei, bei dem Begriff der parasozialen Interaktion zu bleiben (ebd.). Während dieser Rat für den englischen Sprachraum, in dem weitestgehend von *parasocial interaction* die Rede ist, eine Überlegung wert ist, wird im Kontext dieser Untersuchung der Begriff der (sozialen) Parainteraktion bevorzugt, da er nicht nur terminologisch konkreter ist, sondern „sich in der Medienlinguistik der letzten Jahre als Erweiterung des Interaktionsbegriffs unter bestimmten medialen Bedingungen durchgesetzt hat“ (Luginbühl/Meer 2022).

Diesem Verständnis soll sich unter der Prämisse angeschlossen werden, dass der Begriff der Parainteraktion in (medien-)linguistisch orientierter Forschung das gemeinte Phänomen genauer beschreibt, besonders in Hinblick auf den hier zugrundeliegenden Interaktionsbegriff nach Imo (2016). Allerdings wird der Begriff bei der Analyse der YouTube-Videos dahingehend ausgeweitet, dass Parainteraktivität nicht nur verbal sondern auch multimodal gefasst wird. Neben der grundsätzlichen terminologischen Schwierigkeit des ursprünglichen Begriffs ‚parasozial‘, muss zusätzlich geklärt werden, was *para* in diesem Kontext genau bedeutet: ‚neben‘ bzw. ‚nebenher‘, oder eher ‚darüber hinaus‘, ‚jenseits‘, oder gar ‚entgegen‘. Der Begriff parasozial bzw. parainteraktiv meint jedoch eine auf sich selbst bezogene Kommunikation, die ein ‚So als ob‘ fingiert. Allerdings gibt es einen Unterschied zwischen Parainteraktion auf der Produktseite und der Rezeptionsseite. Denn im Gegensatz zur Produktseite wird bei der Rezeption eine Wahrnehmung durch die andere Person, also den Medienschaffenden, gar nicht erst angenommen. Die kommunikative und mediale Voraussetzung ist eine völlig andere. Es ist daher nötig, die von den Medienpersonen getätigten Ansprachen an das Publikum terminologisch von denen zu unterscheiden, die das Publikum

an die Medienpersonen richten. Da Parainteraktion hier als Unterkategorie von Kommunikation betrachtet wird, bei der die Wahrnehmung der anderen Person grundsätzlich angenommen werden kann (sofern das Video veröffentlicht wurde), kann bei den an die Medienpersonen gerichteten Aktivitäten der Zuschauenden nicht von Kommunikation gesprochen werden. Vielmehr handeln die Zuschauenden *pseudokommunikativ* den Medienpersonen gegenüber (während sie den anderen Anwesenden gegenüber sehr wohl kommunikativ und interaktiv handeln). Hinsichtlich einer Rezeptionsanalyse stellt sich die Frage, was konkret analytisch unter Pseudokommunikation fällt, was nicht und inwiefern Pseudokommunikation situativ eingebettet ist. Im Gegensatz zu Ansätzen, die pseudokommunikative Handlungen auch über die Rezeptionssituation hinaus betrachten,⁶⁷ wird sie hier analog zu Parainteraktion auf Produktionsseite als kommunikative Handlungen, die eine Face-to-face-Situation fingieren, verstanden. Pseudokommunikation kann sicherlich kognitiv über einzelne Situationen hinauswirken, wird jedoch analytisch situativ während der Rezeption gefasst.

Der Begriff der „Pseudokommunikation“, bzw. „pseudokommunikative Äußerungen“ (Ayaß 2001: 215) wurde bereits im Kontext der eingangs erwähnten von Holly, Püschel und Bergmann geleiteten Rezeptionsstudie *Über Fernsehen Sprechen. Die kommunikative Aneignung von Fernsehen in alltäglichen Kontexten* vorgeschlagen, die untersuchte, inwiefern Zuschauer*innen auf parainteraktive Angebote von Medienpersonen reagieren. Im Rahmen des Projekts sind eine Vielzahl an Veröffentlichung und gleich mehrere Sammelbände entstanden, wobei in einem auch Ayaß' (1993) Beitrag zur terminologischen Schärfung des Parainteraktionsbegriffs zu finden ist und sich in einem anderen auf den Begriff der Pseudokommunikation „geeignet“ wurde (vgl. Ayaß 2001: 215). In diesem geht es um vielerlei Ausprägungen fernsehbegleitender Rezeptionskommunikation,⁶⁸ worunter unter anderem pseudokommunikative Äußerungen fallen. Die Autor*innen wenden ein enges Verständnis von Pseudokommunikation an, das sich in erster Linie auf die verbale direkte Ansprache von Medienpersonen beschränkt. So weisen Holly und Baldauf im Zuge der Erörterung verschiedener

67 So umfassen z.B. bei Gleich „PSI-Phänomen“ auch die „prä- und postkommunikativen Phasen“ der Rezeption, die „Prozesse der Selektion, Präferenzen und Gratifikation“ mit einschließen (1996: 125) und im Zuge dieser Untersuchung als rezeptionsvorbereitende, -begleitende und Anschlusspraktiken gefasst werden, die mitunter Hinweise auf parasoziale Beziehungen liefern.

68 Im Rahmen der Studie wird auch von fernsehbegleitendem Sprechen gesprochen, bei dem es sich um „eine sehr spezifische Form ‚empraktischen Sprechens‘ in einem so genannten ‚open state of talk‘ (Goffman 1981: 134f.) handelt, wobei in kleinen Häppchen und kurz, doch sehr kompakt und effektiv kommuniziert wird und der Kontext eine erhebliche Rolle spielt“ (Holly 2002: 360).

kommunikativer Konstellationen zwischen Medienperson/en und rezipieren Person/en darauf hin, dass Zuschauende auch die Möglichkeit haben, „eine Kommunikation zu den Fernsehakteuren zu simulieren; dies nennen wir ‚Pseudokommunikation‘“ (2001: 44). Damit verstehen sie Pseudokommunikation in einem engeren Sinne als Teil parainteraktiver Aktivitäten von Rezipierenden: „Parainteraktion kann bis zu scheinbar an den Fernsehakteur gerichteten sprachlichen Handlungen führen [...] Die Rezipienten reagieren, als könnten die Fernsehakteure sie hören und verstehen; sie inszenieren Kommunikation. Diese Sprechhandlungen bezeichnen wir als ‚Pseudokommunikation‘“ (ebd.: 59). Während die Autor*innen einiger Beiträge des Bandes den Begriff der Pseudokommunikation auch auf Aktivitäten der Medienpersonen beziehen, wird er in dieser Untersuchung ausschließlich für die Rezeptionsseite verwendet.

Die Autor*innen sprechen außerdem von der *kommunikativen Aneignung* von Fernsehen-/Medientexten (vgl. Holly 2001: 13; Faber 2001a). Damit bezeichnen sie „eine kreative Begegnung von Rezipient und Medienprodukt“, die „explizite und implizite, assoziative und intellektuell-spielerische“ Formen annehmen kann (Holly 2011: 19; vgl. Faber 2001a: 34–37).⁶⁹ Als Voraussetzung für die Aneignung nennen sie eine bestimmte Ausprägung der Beschäftigung mit dem (Fernseh-)Text, die sie in mehrere Grade und Arten unterteilen: 1) Wahrnehmen, 2) aktives Zuwenden, 3) intensive Aufmerksamkeit, mit 4) kognitivem Engagement oder 5) emotionaler Beteiligung, 6) metakommunikative Beschäftigung und 7) Evaluationen und Bewertungen (vgl. Holly/Baldauf 2001: 49f). Um Aufschlüsse über die kommunikative Aneignung von und die rezeptionsbegleitende Kommunikation beim Fernsehen zu erhalten, wurden mithilfe von Tonbandgeräten Gespräche während der Fernsehrezeption in privaten Wohnzimmern aufgezeichnet. Im Zuge der Auswertung erkennen die Autor*innen mehrere Strukturen und Formen des fernsehbegleitenden Sprechens u.a. *freistehende* Äußerungen, *Gesprächsinself*, Adressierungen, Kommentare, Bewertungen und auch Parainteraktion (vgl. Baldauf 2001: 65–72). Die Untersuchung hat gezeigt, dass die Rezipient*innen sowohl abhängig als auch unabhängig von einer direkten Aufforderung durch die Medienpersonen mit diesen (sprachlich) interagieren. Baldauf hält diesbezüglich fest, dass pseudokommunikative Äußerungen zum großen Teil im Sinne Kühns (1995) mehrfachadressiert sind, ferner, sie sich hinsichtlich ausgedrückter Bewertungen oder Inszenierungen vor allem an die anderen Rezipierenden richten, sodass sie vorschlägt, „von einer verdeckten Mehrfachadressierung, einer Art Inszenierung, zu sprechen, wobei der Adressat im Fernsehen

69 Für eine umfassende Darstellung verschiedener enger und weiter gefasster Aneignungsbegriffe vgl. Faber (2001a).

nur ein ‚Pseudo-Adressat‘ ist“ (ebd.: 71). Auch Püschel (1993) stellte bereits fest, dass Äußerungen sowohl pseudokommunikativ an die Medienpersonen adressierte Redebeiträge als auch zugleich *turns* in der Interaktion der anwesenden Rezipierenden sein können, wobei er darauf hinweist, dass eine Adressierung letzterer im Falle selbstgesprächshafter Äußerungen nicht immer beabsichtigt sein muss (1993: 124f).

Darüber hinaus beschreiben die Autor*innen, besonders (Klemm 2001a),⁷⁰ verschiedene Sprachhandlungsmuster. So stellten sich folgende Sprachhandlungsfelder rezeptions-begleitenden Sprechens heraus: 1) das Organisieren der Rezeption, 2) das kognitive Verstehen und 3) das emotionale Verarbeiten der Inhalte, 4) das aneignende Übertragen und Einordnen der Ereignisse auf die eigene Lebenswelt,⁷¹ 5) das individuelle oder gruppenbezogene Vergnügen, 6) das Deuten und 7) das Bewerten des Texts (vgl. Holly 2001: 20; Klemm 2001a). Ausgehend von diesen Beobachtungen kann fernsehbegleitender Kommunikation folgende Funktionen erfüllen: 1) Die Schaffung einer geselligen Situation, 2) gegenseitige Hilfe beim Verstehen, Deuten und Bewerten, 3) die Herstellung einer Verbindung zwischen Fernseh- und Alltagswelt und 4) die Vergemeinschaftung der Gruppe über den Abgleich von Einstellungen und Werten (vgl. Holly 2001: 20; Klemm 2001a: 112ff). Fernsehen sei in Bezug auf die Rezeption in Gruppen:

vor allem eine Orientierungsressource für die direkte Kommunikation in primären Gruppen. Dabei kommt die Orientierung nicht in erster Linie aus den Sendungen des Fernsehens, sondern von den Zuschauern selbst, die – auf diese Weise immer wieder und unauffällig kommunizierend – ihr Wissen, ihre Einschätzungen und ihre Bewertungen austauschen und gegeneinander abgleichen, wobei ihnen die Fernsehsendungen als symbolisches Ausgangsmaterial und Vorlage dienen (Holly 2002: 353f).

Die Autor*innen zeigten außerdem auf, inwiefern Rezeptionsstile in verschiedenen Gruppen je nach Anlass und Stimmung divergieren, und sich darüber hinaus verschiedene gattungsspezifische Ausprägungen⁷² beobachteten lassen: „Auch wenn nur wenige Sprachhandlungsmuster des fernsehbegleitenden Sprechens wirklich gattungsspezifisch sein werden, leuchtet doch ein, dass bestimmte

70 Siehe außerdem Klemms Dissertationsschrift (2000), Kapitel 5.

71 Sie verwenden den Begriff in Anlehnung an Schütz bzw. Schütz/Luckmann (1994 [1979]) im Sinne alltäglicher kommunikativ vermittelter Erfahrungsräume.

72 Detailliert beschrieben werden die Gattungen Nachrichten (Klemm 2001b), Ratgeber-sendungen (Habscheid 2001), Krankenhausserien (Faber 2001b) und Werbespots (Ayaß 2001).

Gattungen mehr als andere die Zuschauerkommunikation entfachen“ (Holly 2002: 345f). Im Vorausblick auf die Analyse ist zu sagen, dass sowohl Beauty- als auch FIFA-Videos dieses Merkmal zu erfüllen vermögen. Auf diese und weitere Ergebnisse aus dem Kontext der Studie wird im Laufe der Analyse an relevanten Stellen zurückgekommen wie auch auf Erkenntnisse der eingangs erwähnten Untersuchung von Hepp (1998).

In Hinblick auf jüngere linguistische Forschung zu (sozialer) Parainteraktion in anderen medialen Umgebungen ist unbedingt der Sammelband von Luginbühl und Meer (2022) zu erwähnen, in dem es unter anderem um Parainteraktion in sozialen Medien geht. Die Autor*innen zeigen neben diversen methodischen Zugängen verschiedener Teildisziplinen auf, dass erstens eine Definition dessen notwendig ist, was als Parainteraktion beschrieben wird – wobei sie sich für ein weiteres, jedoch differenziertes Verständnis aussprechen –, zweitens, dass sich Parainteraktion textsortenspezifisch gestaltet und drittens es bei Fragen der ‚Messbarkeit‘ von Parainteraktion darum gehen muss, „präferierte Lesarten der Medientexte zu rekonstruieren“ (Luginbühl/Meer 2022: 17) und daraus resultierende Annahmen im Kontext einer Rezeptionsforschung zu überprüfen. Bezüglich der Gegenstandskonstitution von Parainteraktion setzen sie die Unterscheidung zwischen interpersonalen Formen von Parainteraktion, die im Sinne des Grundkonzepts Horton und Wohls auf die Illusion einer wechselseitigen Interaktion abzielen, einerseits und kognitiv/affektiven Formen, die linguistisch schwierig zu fassen sind und sich bereits in den Überlegungen von Hartmanns Zweiebenen-Modell niederschlugen (vgl. ebd.: 10), andererseits in den Fokus. Vor diesem Hintergrund sprechen sich vor allem Meer und Weidner dafür aus, mit einem erweiterten Verständnis von Parainteraktion sowohl die „interaktiv ausgerichteten Aspekte“ als auch „darüber hinaus gehende Strategien der Herstellung von sozialer Nähe“ zu fassen (ebd.: 11).

Bezogen auf die Gegenstandsbereiche dieser Untersuchung sind zunächst die Beiträge interessant, die sich mit Werbung und Influencing beschäftigen. Zum einen wertet Stöckl Werbeposts von Influencer*innen auf Instagram und Printanzeigen hinsichtlich (parainteraktiver) „message cues“ (Stöckl 2022: 23) aus und kommt zu dem Ergebnis, dass in den Influencing-Beiträgen im Vergleich zu Printanzeigen eher weniger Parainteraktion zu beobachten ist, was er damit erklärt, dass durch das auf Instagram grundlegende Angebot des Austauschs dort weniger ‚aktivierende‘ Ansprache notwendig ist, als es bei Printanzeigen der Fall zu sein scheint (vgl. ebd.: 49). Zum anderen untersucht Meer eine Werbekampagne, die in Kollaboration mit den Teilnehmerinnen der Fernsehshow *Germany's Next Topmodel* entstanden ist und neben Werbespots auch Printanzeigen, Instagram-Beiträge und Tutorials auf YouTube umfasst. Aufbauend auf ihre

vorausgehende Forschung (s. Kapitel 2) erläutert sie, dass Parainteraktion genutzt wird, um Glaubwürdigkeit zu erzeugen und darüber hinaus, dass dies nur dann erfolgreich gelingen kann, wenn parainteraktive Elemente auch in den für die Rezipierenden relevanten Kampagnen mit teilweise osmotisch werbenden Formaten stattfinden, also „lebensweltlich integriert genutzt werden“ (Luginbühl/Meer 2022: 12). Bezüglich audiovisueller Medienangebote untersucht außerdem Weidner Kochshows im Fernsehen und beschreibt dabei unter Rückbezug auf vorausgehende Forschung spezifische parainteraktiv genutzte (mediale) Ressourcen, anhand derer sich im Rahmen dieser Untersuchung orientiert wird, weshalb sie im nächsten Kapitel vertiefend erläutert werden.

In Hinblick auf den Untersuchungsgegenstand YouTube nimmt Ebel Lernvideos und zugehörige Kommentarlisten in den Blick, wobei die Videos eine komplett andere semiotische Ausgestaltung als die hier fokussierten Beauty- und FIFA-Videos aufweisen und in sich Parainteraktion in ihnen weniger durch direkte Ansprachen als vielmehr durch eine bestimmte paraverbale Ausgestaltung ausdrückt, bei der die Komponenten des Sprechausdrucks und der Ansprechhaltung eine zentrale Rolle spielen (vgl. Ebel 2022: 163). Obgleich die im Zuge des Sprechausdrucks relevanten prosodischen und artikulatorischen Aspekte von Parainteraktion in YouTube-Videos im Rahmen dieser Untersuchung nicht systematisch berücksichtigt werden (können), stellen sie dennoch Kategorien dar, die insbesondere unter einer genderspezifischen Perspektive aufschlussreich sein können. Hinsichtlich der hier untersuchten Beauty- und FIFA-Tutorials kommt jedoch auch der Beitrag von Horst und Ladewig in den Blick, die ausgehend von filmwissenschaftlichen Überlegungen Yoga-Tutorials auf YouTube hinsichtlich parainteraktiver Elemente betrachten, die auf die Körperlichkeit der Rezipierenden abzielt. Durch medientechnische Ressourcen wie „Kamerabewegungen, Bildkomposition, Lichtgestaltung“ (Luginbühl/Meer 2022: 13) würden Rezipierende dazu angehalten, die präsentierten Bewegungen simultan auszuführen, wodurch es zu einer *körperlichen Verflechtung* zwischen ihnen und der Medienperson und somit zu einem Eindruck parainteraktiver Nähe kommen würde (vgl. Horst/Ladewig 2022: 132ff). Diese Beobachtungen sind für Beauty- und FIFA-Tutorials insofern relevant, als auch dort nahegelegt wird, die gezeigten Schminkschritte oder Spielmechanismen körperlich nachzuahmen, wenngleich eine zeitgleiche Ausführung eher in den Beauty-Videos angelegt ist. Bezogen auf die FIFA-Videos ist zudem auch der Beitrag von Dang-Anh relevant, der sich mit Livestreams beschäftigt und herausstellt, dass es in (zeitlich knapp verschobenen) Livevideos zu einem medienvermittelten Austausch kommen und somit „soziale Parainteraktion transzendiert“ (Dang-Ahn 2022: 185) werden kann.

Inwiefern die Streamer*innen in ihrer Ansprache auch zwischen Live- und VOD-Publikum divergieren, wird im Rahmen der Analyse deutlich werden.

Den Beiträgen dieses Sammelbandes ist gemein, dass sie Parainteraktion auf semiotischer Ebene untersuchen, Parainteraktion also multimodal konzeptionieren und hinsichtlich semiotischer Ressourcen analysieren (vgl. Luginbühl/Meer: 17). Dieses grundsätzliche Verständnis liegt auch dieser Untersuchung zugrunde, sodass es in dem nächsten Kapitel darum gehen wird, die grundlegenden Modalitäten und Ausdrucksressourcen, anhand derer sich die Analyse von Parainteraktion orientiert, theoretisch einzuführen und näher zu bestimmen.

3.2 Multimodalität

Als innerhalb der Linguistik etablierte Teildisziplin bietet die Multimodalitätsforschung besonders im Kontext von audiovisuellen Textsorten einen Analyseansatz, der es ermöglicht, neben den verbal-auditiven auch visuelle Kategorien zu beschreiben (vgl. Schneider/Stöckl 2011). Meer (2018a, 2018b) verdeutlichte bereits die Relevanz einer multimodalen Analyseperspektive anhand von Tutorials, die auf YouTube veröffentlicht wurden. Diese Argumentation wurde seit dem für weitere Beauty-Textsorten wie Hauls weiter ausgearbeitet (vgl. Böckmann et al. 2019; Meer/Staubach 2021; Och 2021a) und soll hier nun auch auf FIFA-Let's-Plays übertragen werden. Dafür wurden bzw. werden Begriffe von Multimodalität aus zwei verschiedenen linguistischen Forschungsbereichen miteinander verknüpft: Einerseits wird der Begriff im Sinne der semiotischen Textanalyse (Stöckl 2012, 2016) aufgegriffen, andererseits um Kategorien aus der konversationsanalytischen Gesprächsforschung (Stukenbrock 2009, 2010; Weidner 2017, 2022) ergänzt. Für die Betrachtung von Parainteraktivität in hypermedialen Texten ist es zudem notwendig, medienspezifische Ressourcen einzubeziehen. Die einzelnen Forschungsansätze werden im Folgenden kurz in Hinblick auf die Videoanalyse erläutert.

3.2.1 SEMIOTISCHE KATEGORIEN

Die Betrachtung von Multimodalität hat sich in der semiotischen Forschungstradition aus dem Anspruch entwickelt, neben sprachlichen auch bildliche Elemente semiotisch betrachten zu können (vgl. Stöckl 2004). Vor dem Hintergrund der Werbetextanalyse – um die es bei Influencing letztendlich auch geht – stellt sich nach Stöckl somit die Grundfrage „nach welcher funktional-semantischen Logik die einzelnen Codes miteinander verknüpft sind und intermodal ko-

härennten Sinn ergeben“ (Stöckl 2012: 248). ‚Kode‘ als linguistische Kategorie bezieht sich hier auf die Bedeutungsebenen von Sprache und Bild. Darüber hinaus definiert Stöckl auch Musik und Geräusch als Kodes (vgl. Stöckl 2012: 248f). Im Anschluss an die konkrete Realisierung der Kodes geht Stöckl von verschiedenen ‚Modes‘ aus, die immer an die jeweiligen Kodes zurückgebunden sind: „Dabei beschränke ich mich auf Sprache, Bild, Musik und Geräusch, weil sie im Alltagsverständnis der Zeichenverwender zweifelsfrei als basale Modalitäten gelten können“ (Stöckl 2016: 6). Der Mode ‚Sprache‘ umfasst dabei gesprochene wie geschriebene Sprache, ‚Bild‘ indiziert sowohl statische als auch bewegte Bildrealisierungen und Musik und Geräusch sind in der Modalität des Tons zusammengefasst. Eine Kombination zweier oder mehrerer Modalitäten wird folglich ‚multimodal‘ genannt.

Dieses Verständnis von Modalitäten und Multimodalität bietet mit Blick auf YouTube-Videos ein breites Spektrum von Analysekatégorien, das viele in den Videos auftretende Aspekte zu beschreiben vermag. So ist zunächst unabhängig von der Kategorie der Parainteraktion festzuhalten, dass YouTube-Videos aufgrund ihrer Audiovisualität grundsätzlich multimodal sind. Bezogen auf parainteraktive Ansprache erweist sich unmittelbar die Modalität der gesprochenen Sprache und somit das Gesagte der YouTuber*innen als zentral. Da eine Fokussierung auf rein verbale Äußerungen für eine ganzheitliche Betrachtung parainteraktiver Anteile jedoch – wie bereits im Rahmen der Erläuterungen zu Parainteraktion deutlich wurde – nicht ausreicht, geraten auf der visuellen Ebene auch die Mimik und Gestik der YouTuber*innen in den Blick. Stöckl klammert diese Kategorien zunächst aus, beziehungsweise versteht sie den Modalitäten Sprache, Bild und Ton untergeordnet:

Zum anderen lässt sie [die theoretische Engführung, Anm. d. Verf.] solche Zeichenmodalitäten wie Gestik und Mimik, aber auch Typographie oder Intonation/Stimmgestaltung unberücksichtigt, weil diese in dem Sinne peripher und kaum autonom sind, dass sie in andere Modalitäten integriert werden und nur bei der gleichzeitigen Realisierung von Rede, Schrift oder Bild ins Spiel kommen (Stöckl 2016: 6).

Allerdings sind insbesondere Gestik und Mimik in ihrer interaktiven Relevanz nicht peripher, sondern bilden wichtige, das Verbale unterstützende Mittel und können darüber hinaus auch autonome Funktionen erfüllen, wie Fricke (2009) bereits für die Gestik in Face-to-face-Kommunikation herausgestellt hat. Auch bezogen auf Beautyvideos und FIFA-Let’s-Plays lassen sich bereits bei einer ersten Betrachtung eine Vielzahl körperlicher Handlungen bzw. Gestikulationen der YouTuber*innen erkennen. Körpersprachliche interaktive Elemente sollen bei

der Betrachtung deshalb nicht vernachlässigt werden, sondern bilden einen wichtigen Teil der Analyse. So wird deutlich, dass die genannten Hauptmodalitäten für den hier betrachteten Gegenstand hypertextueller Videos zu kurz greifen. Da Gestik und Mimik aus linguistischer Sicht bisher vor allem im Rahmen konversationsanalytischer Forschung behandelt wurden, wird der semiotische Ansatz nach Stöckl durch interaktionstheoretische Begriffe von Modalität ergänzt.

3.2.2 INTERAKTIONALE AUSDRUCKSRESSOURCEN

Im Bereich der Konversationsanalyse haben sich vor allem Stukenbrock (2009, 2013), Fricke (2012) sowie Deppermann und Schmitt (2007) mit Begriffen von Multimodalität in Hinblick auf körperliche interaktive Ausdrucksweisen befasst. Fricke plädiert für ein multimodales Verständnis bei der Beschreibung von Sprache und Gestik, während Deppermann und Schmitt betonen, dass die von Interagierenden genutzten Ausdrucksmodi meist als „Ressourcenbündel“ (ebd.: 25) auftreten. Im Rahmen dieser Untersuchung wird sich jedoch an den von Anja Stukenbrock herausgestellten Beschreibungskriterien sowie ihrer Definition von Multimodalität orientiert:

Der Begriff Multimodalität (i.S.v. Modalität = Ausdrucksressource) begründet ein neues Forschungsparadigma, das sich im Kontext der Konversationsanalyse und der interaktionalen Linguistik herausgebildet hat und das neben den verbalen auch die non-verbalen – körperlich-visuellen – Ausdrucksressourcen (Blick, Gestikulation, Körperpositur, Körperorientierung, Bewegung, Manipulation von Objekten etc.) als gleichrangige Bestandteile der face-to-face-Interaktion untersucht (2013: 253).

Stukenbrock verwendet den Begriff ‚Mode‘ hier nicht im Sinne Stöckls, sondern fokussiert sich auf die von Personen ausgehenden Ausdrucksressourcen, wie unter anderem Gestik und Blick. In Bezug auf die Gestik machte sie in vorherigen Veröffentlichungen vor allem anhand deiktischer Ausdrücke und Gesten deutlich, dass verbale Äußerungen unter einer multimodalen Perspektive betrachtet werden müssen, die körperliche (Zeige-)Gesten inkludiert (vgl. Stukenbrock 2008, 2010, 2015). Als bisher vernachlässigten Mode, der körperliche (Para-)Interaktion betrifft, nennt Stukenbrock den Blick. Dies erscheint bezogen auf ihren Forschungsgegenstand des Zeigens in Face-to-face-Konversationen einleuchtend, bei denen es immer einen bestimmten „Verweis- bzw. Suchraum“ (Stukenbrock 2008: 4) sowie ein „Zeigeziel“ (ebd.) gibt. Der Blick tritt dabei mit koordinierenden verbalen Deiktika und/oder Gesten auf (vgl. ebd.: 28). Obwohl Stuken-

brock sich in ihren Analysen auf authentische und tatsächliche Interaktionen (im Gegensatz zu Parainteraktion) bezieht, ist die Kategorie des Blicks als Ausdrucksressource auch für den Gegenstand der YouTube-Videos zentral. Dies ist insofern naheliegend, als Hartmann/Goldhoorn bereits für das Fernsehen festgestellt hatten, dass „if a TV performer gazes directly into the camera, users may automatically acquire the feeling that the performer would look at them, personally“ (2011: 1106) und somit die Relevanz eines scheinbaren Blickkontakts herausstellten.

Ein weiterer relevanter Aspekt – besonders im Hinblick auf Beautyvideos – ist der von Stukenbrock genannte Punkt der „Manipulation von Objekten“ (s. o. Stukenbrock 2013: 253). So werden in Hauls, Tutorials oder First Impressions Produkte und Kleidungsstücke, also Objekte, präsentiert, aber darüber hinaus ausprobiert bzw. vorgeführt. Für die Analyse parainteraktiver Verfahren ist es daher sinnvoll, auch für derartige Handlungen eine Beschreibungskategorie zu finden. Diese bei Stukenbrock lediglich angerissene Arbeit an Objekten wird von Weidner (2017, 2022) weiter ausgeführt. Sie bezieht sich dabei im Gegensatz zu Stukenbrock auf einen audiovisuellen Gegenstand, indem sie sich die multimodale Ausprägung von Bewertungen (2017) respektive die Inszenierung einer geteilten Situation (2022) in Koch-Shows ansieht. Neben textsortenspezifischen Fragestellungen geht sie darauf ein, welche kommunikativen Mittel die Akteur*innen bei der Zubereitung und der Bewertung des Essens verwenden. Dabei liegt ihr Fokus neben dem Verbalen ebenfalls auf körperlich-visuellen Aspekten, die die Kategorien der Gestik, der Mimik und des Blicks, aber auch die Präsentation von Objekten einschließen (vgl. Weidner 2017: 6). Darüber hinaus prägt sie den Begriff der Objektmanipulation als Ausdrucksressource, unter dem das bearbeitende Zeigen von Gegenständen verstanden wird (vgl. ebd.: 11). Diese interaktiv relevanten Ausdrucksressourcen – Gestik, Blick, Objektmanipulation und Objektpräsentation – werden bei der Analyse der You-Tube-Videos, besonders hinsichtlich ihres parainteraktiven Einsatzes berücksichtigt.

Weidner macht allerdings darüber hinaus mit Blick auf Fernsehsendungen auf einen weiteren für die Analyse wichtigen Aspekt aufmerksam. So weist sie darauf hin, dass in ihrem Beispiel der Koch-Shows manche Objektpräsentationen nur aufgrund einer bestimmten Kameraauswahl (da mehrere Kameras das Geschehen filmen) sowie bestimmter Kameraeinstellungen für die Rezipient*innen zugänglich ist: „Nur durch die *Close ups* können die Zuschauer die Handlungsschritte, die der prospektiv Bewertende am Fischfilet durchführt, mitverfolgen [Herv. i. O.]“ (Weidner 2017: 18). Zwar nennt sie hier den Einsatz der Kamera lediglich in Zusammenhang mit der Objektmanipulation und versteht ihn daher als Mittel, diese für die Zuschauer*innen deutlicher erkennbar zu machen, lenkt

den Blick damit jedoch auf ausschlaggebende medienspezifische Gegebenheiten. Durch ihre Berücksichtigung von Einstellungsgrößen, Zoom und Schnitt (vgl. Weidner 2022: 74) lenkt sie die Aufmerksamkeit auf die Kameraarbeit live, wobei sich gerade für den Schnitt im Hinblick auf YouTube-Videos in der Nachbearbeitung noch einmal andere Möglichkeiten ergeben. Die Konstituenten des Schnitts, aber auch die der Kameraperspektive, das Verhalten der Medienakteur*innen zu der Kamera sowie auch der von der Rückbindung an einzelne Ausdrucksressourcen unabhängige Kameraeinsatz sollen daher im nächsten Abschnitt methodisch näher betrachtet werden.

3.2.3 MEDIENSPEZIFISCHE AUSDRUCKSRESSOURCEN

Bisher wurden bedingt durch den semiotisch-textlinguistischen Ansatz einerseits und Überlegungen aus der Konversationstheorie andererseits Elemente wie Bild oder Gestik als Modalität, beziehungsweise als Ausdrucksressource kategorisiert. Vor dem Hintergrund, dass diese Elemente in anderen Forschungstraditionen mit einem eher weit gefassten Medienbegriff als eigenständige Medien begriffen werden können und darüber hinaus in Anbetracht des hier behandelten Gegenstands, der sich den ‚neuen Medien‘ zuordnen lässt, ist es nötig, den hier verwendeten Medienbegriff näher zu definieren.

Unter Zuhilfenahme eines weiten Medienverständnisses, in dem auch „Face-to-Face-Kommunikation als basales Medium menschlicher Kommunikation“ (Schneider 2018: 277) gelten kann, wird ersichtlich, dass die oben genannten Modalitäten der Sprache, des Bildes, der Musik usw. jeweils auch fruchtbar als Medien beschrieben werden können. Dass sie auch in der Forschung teilweise synonym verwendet werden, stellt Schneider in der Gegenüberstellung der Konzepte von (Multi-)Medialität vs. (Multi-)Modalität kritisch fest (ebd.: 278). Während sich hier dem grundsätzlichen Verständnis darüber, dass Medien „konstitutiv für das darin Mediatisierte“ (ebd.: 277) sind, angeschlossen wird, wird bei der Untersuchung dennoch zwischen Medialität und Modalität unterschieden, mit dem Hintergrund, semiotische Modalitäten und interaktive Ausdrucksressourcen sowohl terminologisch in der jeweiligen Forschungstradition verortend als auch analytisch zu unterscheiden. Im Sinne Jägers⁷³ – abgeleitet von dem von Schneider zitierten Vorschlag, „Medialität als einen ‚heuristischen Suchbegriff‘“ (ebd.) zu verstehen – wird im Zuge der Berücksichtigung der für YouTube speziellen medientechnischen Ressourcen, deren Notwendigkeit sich bereits bei Weidner (2017) als auch bei Luginbühl/Meer (2022) abzeichnete, hier ein

73 Schneider bezieht sich hier auf Jäger (2015).

technisch ausgerichteter Medienbegriff verwendet, der die mediale Einbettung und Ausgestaltung fasst.

Medien werden daher im Anschluss an Habscheid (2000) als „materiale, vom Menschen hergestellte Apparate zur Herstellung/Modifikation, Speicherung, Übertragung oder Verteilung von sprachlichen (und nichtsprachlichen) Zeichen (im Sinne musterhafter Äußerungen)“ (ebd.: 138) verstanden. Dieser Medienbegriff ist folglich eher technisch orientiert und soll in Abgrenzung zum hier verwendeten Modalitätsbegriff die einzelnen Beschreibungskategorien differenzierter unterscheiden, um einzelne Modalitäten, Ausdrucksressourcen sowie medienspezifische Aspekte genauer erfassen zu können. Dies ist also einerseits im Sinne einer unmissverständlichen Beschreibung eine pragmatische Entscheidung, andererseits erscheint es aus analytischer Perspektive notwendig, die Modalitäten Sprache, Bild und Ton von den Ausdrucksressourcen zu unterscheiden, da sie unterschiedlich funktionieren und eine einheitliche Benennung aller als Medien zu kurz greifen respektive für jeden einzelnen Fall Erklärungsbedarf bedeutet würde. Ausgehend von der Klassifikation des Gebrauchs der Kamera, des Settings und des Schnitts als medientechnische Komponenten von YouTube-Videos, werden sie als *medienspezifische* Ausdrucksressourcen bzw. Modalitäten gehandhabt. Der zugrundeliegende technische Medienbegriff ist außerdem abzugrenzen von und dabei gleichzeitig vereinbar mit dem Begriff der Kommunikationsform, der vor dem Hintergrund eines weiten Medienverständnisses einerseits und dem Konzept kommunikativer Praktiken andererseits in der Forschung kritisiert und als obsolet gehandhabt wurde (vgl. Habscheid 2020: 151). Durch die „Revision“ des Begriffs durch Brock und Schildhauer gelang dieser innerhalb der Trias aus Medium, Kommunikationsform und Genre jedoch an Trennschärfe (vgl. 2017: 27ff).⁷⁴ Sie fassen „the technical medium as one of its [communication form, Anm. der Verf.] most basic components“ (ebd.: 28). Kommunikationsformen werden dabei definiert als aus Praxis resultierende „clusters of technical and communicative constellations with certain potentials, which are commonly restricted by conventions“ (ebd.: 32). Ausgehend von Kommunikationsformen würden dann „genre patterns“ (ebd.: 29) ermöglicht werden, bzw. adaptiert für diesen Kontext: mögliche Textsortenmuster, die sich dann je nach Konvention und Situation ausprägen. Der semiotische Gehalt vom Kommunikationsformen wird in diesem Zusammenhang ebenfalls bedacht, da die Autoren davon ausgehen, Ressourcen würden kulturell „semiotised“ (ebd.: 28), um kommunikativen Praktiken zu dienen.

74 Siehe auch die weiteren Beiträge des Sammelbands (vgl. Brock/Schildhauer (2017) (Hrsg.)), besonders den ebenfalls theoretischen Beitrag von Meiler (2017).

In Bezug auf parainteraktive Praktiken ist es zunächst auch die Kamera, welche durch ihre Bildgebung den Interaktionsraum erst medial konstituiert (vgl. Weidner 2017: 15) und somit festlegt, was die Zuschauer*innen zu sehen bekommen. Zudem ist der Umgang mit ihr auf Produktionsseite entscheidend. Allerdings stellten bereits Horton und Wohl (1956) unter Bezugnahme auf das Fernsehen die Kamera als einen wichtigen Faktor heraus. Nach ihnen stellt sie bezogen auf den Blickwinkel in der Aufnahmesituation die Stellvertretung für die Zuschauer*innen dar: „making the camera be the eyes of the audience“ (Horton/Wohl 1956: 218). Auf der anderen Seite orientieren sich die Akteur*innen vor der Kamera an ihr, wie sie es in Face-to-face-Situationen an ihrem Gegenüber machen würden. Hartmann veranschaulicht in diesem Zusammenhang das spezifische Verhalten der Akteur*innen wie folgt: „Sie blicken in die Kamera, um die Zuschauer direkt zu adressieren. Wenn sie andere Gäste interviewen, nimmt die Kamera einen Platz in der Mitte ein, um dem Gespräch folgen zu können, und die Personae drehen sich niemals mit ihrem Rücken zur Kamera“ (Hartmann 2016: 79). Damit verdeutlicht er im Anschluss an Horton und Wohl, inwiefern das Verhältnis zwischen Akteur*innen und Kamera aus parainteraktiver Perspektive relevant ist.

Neben der Kamera als Komponente, die den Interaktionsraum festlegt, sind unter dem Aspekt der Parainteraktion auch die filmischen Kategorien der Einstellung und des Schnitts zu berücksichtigen. So kann beispielsweise, wie Weidner (2017) zeigte, ein Close-Up persönliche Nähe suggerieren. Im Hinblick auf die Analyse von YouTube-Videos bedeutet dies, dass die Begriffe von Modalität und interaktiver Ausdrucksressourcen mit medialen Ausdrucksressourcen ergänzt werden müssen. Vor diesem Hintergrund stellt Meer im Kontext der Betrachtung von Tutorials folgendes Multimodalitätsverständnis vor:

Als „multimodal“ fasse ich im Anschluss an Stöckl (2016) einerseits die semiotischen Modalitäten Sprache, Bild und Ton, die im Hinblick auf die hier untersuchten Tutorials alle drei relevant sind. Darüber hinaus sollen im Anschluss an Stukenbrock (2013) und Weidner (2017) aber auch die interaktiv relevant gesetzten Ausdrucksressourcen Blick, Gestik, Objektpräsentation/-manipulation und medial konstituierte Aspekte des Kommunikationsraums als interaktive (Teil-)Modalitäten verstanden werden (2018b: 300).

Dieser Multimodalitätsbegriff wurde bereits an anderer Stelle um „medial konstituierte Aspekte des Kommunikationsraums als (Teil-)Modalitäten“ erweitert, sodass auch „[m]ediale Aspekte wie die *Nutzung der Kamera* oder auch die Bearbeitung der Videos mit den *Mitteln des Schnitts* [Herv. i. O.]“ berücksichtigt werden (Böckmann et al. 2019: 142; vgl. auch Och 2021a).

Darüber hinaus gerät in Bezug auf die Kamera auch deren Positionierung im Raum der Aufnahme in den Blick. Dass diese auch hinsichtlich Parainteraktion relevant ist, stellt Weidner heraus: „Der Eindruck von Intimität kann auch durch Verfahren hervorgerufen werden, die – ohne direkt zu einem konversationellen Gegenzug aufzufordern – eine gesellige Interaktionssituation herstellen“ und da Geselligkeit „auch auf gegenseitiger Wahrnehmbarkeit basiert, ist sie auch interaktional ausgerichtet“ (2022: 57). Vor diesem Hintergrund erläutert sie, inwiefern eine bestimmte Proxemik den Zuschauenden den „Eindruck einer unmittelbaren Anwesenheit im Raum ermöglicht“ (2022: 72). Während sie sich anhand ihres Beispiels auf einen für die Zuschauenden scheinbar geöffneten Gesprächskreis bezieht, ist auch für YouTube-Videos, in denen die YouTuber*innen allein vorkommen, zu beobachten, inwiefern sie die Zuschauenden als räumlich anwesende Gesprächspartner*innen inszenieren. Das Gleiche gilt auch für Videos mit mehreren Personen, die sich jedoch anders als bei Weidner teilweise nicht in den gleichen Aufnahmelokalitäten befinden.

In Hinblick auf Face-to-Face-Interaktionen machten bereits Hausendorf, Mondada und Schmitt auf dabei genutzte multimodale (verbale und körperliche) Ressourcen aufmerksam (wodurch sie auch die Notwendigkeit einer multimodal erweiterten Gesprächsanalyse demonstrieren, vgl. 2012: 11). Vor dem Hintergrund dieser Tatsache stelle sich laut der Autor*innen die Frage „nach der Verortung und Verteilung dieser Ressourcen im Raum“ sowie „nach dem Einsatz und dem Nutzen materieller Bestandteile der lokalen Umgebung“ (2012: 9), so dass sie „*Raum als interaktive Ressource* [Herv. i. O.]“ (ebd.: 11) konzeptionisieren.⁷⁵ Dabei beziehen sie sich auch auf „medial vermittelte Räume“, in denen die Interaktion für die Beteiligten/Medienpersonen „im Dienste der Medialisierung ‚beeinträchtigt‘“ sei (ebd.: 29). Ausgehend von dieser Überlegung gerät somit auch für Parainteraktion der Raum als relevante Ressource in den Blick. Hinsichtlich der Ausgestaltung der Aufnahmeräume meint Holly in Bezug auf das Fernsehen, es erstaune nicht,

dass viele Talkshows in ihren Ausstattungen die Wohnzimmersituation geradezu abzubilden scheinen, wenn sie nicht eine andere vertraute kommunikative Szenerie, die von Lokalen, simulieren. Beide Settings, das pseudohäusliche und das pseudohalböffentliche mit Kneipen- oder Bistrotischen, siedeln die medialen Gespräche in bekanntem

75 Vgl. außerdem Hausendorf/Schmitt/Kesselheim (2016) und Hausendorf/Schmitt (2018), die hervorheben, dass Raum sowohl als interaktive Ressource begriffen werden muss als auch von den Beteiligten interaktiv hervorgebracht wird (vgl. Hausendorf/Schmitt 2018: 92).

Gelände an, das die Fortsetzung der Kommunikation beim Zuschauer fast suggestiv erwirkt (2002: 356).

Die Aufnahmesets würden folglich nicht nur an die Lebenswelt Zuschauender angelehnt sein, sondern darüber hinaus Anschluss- oder rezeptionsbegleitende Kommunikation nahelegen. Holly führt weiter aus:

Die zunehmende Inszenierung medialer Kommunikation als Gesprächsangebot, wenn schon nicht direkt mit dem Zuschauer, dann doch so nah und vertraut vor seinen Augen, dass er gleichsam auf selber Höhe und im möglichst gleichen Ambiente zu Hause weitermachen kann, gehört heute zu den verbreitetsten Tricks, mit denen Medien ihre Kunden einzufangen und zu ‚binden‘ versuchen. So kann man die angepeilte Zielgruppe schon fast ebenso gut am Mobiliar und Design der Studioausstattung wie an der Thematik der Sendung oder am sozialen Sprachstil der Beteiligten erkennen (ebd).

Somit würde die Einrichtung und Ausgestaltung der Sets unmittelbar Rückschlüsse auf die Ziel- und Rezeptionsgruppen zulassen, und darüber hinaus die Zuschauenden, bzw. die auch für den Kontext Influencing durchaus korrekt bezeichneten Kund*innen, an die Videoinhalte binden. Auch in Bezug auf YouTube wurde bereits darauf hingewiesen, dass die individuellen Einrichtungen von Raum und Setdesign relevante Kategorien bei einer ganzheitlichen multi-modalen Analyse von YouTube-Videos darstellen und Aufschlüsse über die parainteraktive Ausgestaltung liefern können (vgl. Och 2021a: 234, 2021b: 12). Somit werden auch die Positionierung der YouTuber*innen im Raum und die Setgestaltung in die Videoanalyse einfließen.

Mithilfe dieser Modalitäten und Ausdrucksressourcen als Beschreibungskategorien sollen im weiteren Verlauf die Videos hinsichtlich ihrer Parainteraktivität sowie ausgedrückter Genderkonstruktion analysiert werden. Daher wird in einem nächsten Schritt detaillierter auf die Kategorie Gender und dem der Analyse zugrundeliegenden Genderbegriff eingegangen.

3.3 Gender

3.3.1 THEORETISCHE GRUNDANNAHMEN EINES GENDERBEGRIFFS

Während noch zu Beginn des 20. Jh. Einigkeit darüber herrschte, dass Geschlecht eine biologische Tatsache ist, hat sich diese Sicht im Laufe des Jahr-

hundreds verändert. Zur Mitte des letzten Jahrhunderts hin etablierte sich somit eine Unterscheidung zwischen biologischem und sozialem, konstruiertem Geschlecht respektive *Sex* und *Gender* (vgl. Stoller 1968). In den 1980er Jahren rückte schließlich die Idee der *Performanz* von Gender in den Mittelpunkt. Dieser konstruktivistische Ansatz wurde insbesondere von West und Zimmerman (1987) in den Diskurs gebracht, die – in Anlehnung an Harvey Sacks' ethnomethodologischen Begriff des *doing* – auch den Begriff des *Doing Gender* prägten. Dieser meint, dass Gender von Interagierenden in Interaktion immer wieder hergestellt wird. Auf der Basis von Vorstellungen darüber, was als ‚männlich‘ und ‚weiblich‘ gilt, würden sie sich auf eine Art und Weise verhalten, die wiederum genau diese Vorstellungen festigt. Bekannt wurde das Prinzip des Doing Gender vor allem durch Judith Butler, die mit ihrem philosophischen Aufsatz „Gender Trouble“ von 1990, in dem sie ebenfalls Gender, aber auch Sex, als Konstrukte der Gesellschaft beschreibt, dem Thema große Aufmerksamkeit verschaffte.⁷⁶ Allerdings gibt es zwischen diesen beiden Ansätzen – neben ihrer grundsätzlichen Fundierung in unterschiedlichen Forschungstraditionen – Unterschiede: Bei Butler wird Gender „diskursiv, sprachlich hergestellt“, bei West und Zimmerman „sozial interaktiv produziert“ (Frey Steffen 2006: 25). Wesely fasst im Anschluss an West und Zimmerman den Prozess des Doing Gender folgendermaßen zusammen:

Der Begriff doing gender beschreibt die wiederkehrende, routinisierte, methodische Reproduktion von gender. Gender entsteht als eine routinisierte Fähigkeit, die in die alltäglichen Interaktionen eingebettet ist. Die Geschlechtsidentität entsteht also nicht nur durch Erziehung in den ersten Lebensjahren, sondern wird täglich neu erlebt und gelebt (2000: 38).

Herausgestellt werden muss an dieser Stelle insbesondere die Alltäglichkeit der Genderperformanz, die in alle möglichen Praktiken integriert auftritt. Durch diese „Praktiken der Geschlechtsdarstellung“, die aus „kulturell und historisch verankerten variablen Kleidungs-, Ornamentierungs-, Konsum-, Betätigungs-, Verhaltens- und auch Sprechweisen“⁷⁷ (Kotthoff/Nübling 2018: 14) bestehen, vermitteln die Mitglieder einer Gesellschaft ihre eigene Geschlechtsidentität und reproduzieren dabei meist Vorstellungen, die an ein Gender geknüpft sind.

76 Butlers Ausführungen, die dem diskursanalytischen Dekonstruktivismus zugeordnet werden, beziehen sich allerdings nicht auf authentische (Gesprächs-)Daten, sodass sie „sehr empiriefrem“ (Kotthoff/Nübling 2018: 48) bleiben.

77 Zu Sprechweisen vgl. auch Gottburgsen (2000: 52) in Bezug auf Kontaktanzeigen.

Daraus ergibt sich wiederum, dass stereotype Vorstellungen⁷⁸ – insbesondere der Geschlechter ‚männlich‘ und ‚weiblich‘ – noch immer in (Teilen) der Gesellschaft fest verankert sind. Durch diese Verankerung wiederum verhalten sich die Menschen dementsprechend und erhalten die Stereotype im Zuge dessen aufrecht. Diesen Kreislauf machen auch Eckert und McConnell-Ginet (2003: 33) deutlich, indem sie herausstellen: „Our smallest interactions can be imbued with gender, and our continual performance in those interactions strengthens their role in supporting gender“. Daher würde vor allem in der Beziehung und Kommunikation mit anderen die Performanz von Geschlecht deutlich zutage treten.

Für die Genderforschung in der Linguistik bietet Ayaß (2008) in ihrer Einführung „Kommunikation und Geschlecht“ einen Überblick über die zu dem Zeitpunkt aktuelle zentrale Forschung, die sich neben sprachwissenschaftlichen auch auf soziologische Arbeiten erstreckt. Sie definiert Gender wie folgt:

Soziales Geschlecht umfasst die kulturell mehr oder weniger verbindlich festgelegten Aktivitäten und Verhaltensweisen, die einem biologischen Geschlecht zugesprochen werden. In der Regel sind die genderspezifischen Verhaltens- und Handlungsweisen in einer Gesellschaft für ihre Mitglieder so eng mit der biologischen Zuschreibung verknüpft, dass sie als biologisch fundiert gedacht werden (2008: 12).

Zudem weist Ayaß auf den arbiträren Charakter dieser Zuschreibungen hin. Sie sind „insofern beliebig, als sie ihrerseits nicht biologisch begründbar sind“ (ebd.: 13). In diesem Zusammenhang zeigt sie außerdem auf, dass in vielen Studien Gender als binäres System angelegt ist. Als Reaktion darauf hätten bereits Gildemeister und Wetterer (1992) kritisiert, dass diese Studien „die gesellschaftliche Klassifikation in zwei Geschlechter als unhinterfragten Ausgangspunkt wissenschaftlicher Argumentationen setzen“ (Ayaß 2008: 13) und somit „an der sozialen Konstruktion von Zweigeschlechtlichkeit beteiligt“ (Gildemeister/Wetterer 1992: 204, nach Ayaß ebd.) seien. Dementsprechend liegt dieser Untersuchung die Annahme diverser Gender zugrunde.

Bezogen auf das Konzept des Doing Gender zeigt Ayaß auf West/Zimmerman verweisend, dass Menschen nicht nur ihr eigenes Geschlecht, sondern auch „das Geschlecht der anderen fortwährend erzeugen“ (ebd.: 15). Darüber hinaus sei dieser „Herstellungsprozess“ (ebd.: 16) von Gender zum einen unvermeidbar,

78 Mit Stereotypen ist hier (sprachlich tradiertes) Wissen über „Merkmale, die den Mitgliedern sozialisierter Gruppen lediglich aufgrund ihrer Gruppenzugehörigkeit zugeschrieben werden“ (Jonas/Schmid Mast 2007: 69) gemeint, die sich auch auf Gendergruppen beziehen können: „This can be called Gender-Stereotypes; expecting someone to behave a certain way because of their gender“ (Bloom 2021: 5).

zum anderen sei er im Alltag als solcher überhaupt nicht wahrnehmbar. Dies ist insofern interessant, als Menschen andersherum die „Bedingungen und die Konsequenzen“ (ebd.) dieses Prozesses durchaus wahrnehmen. In Anlehnung an Garfinkels (1967) Formulierung verlaufe Doing Gender somit „seen but unnoticed“ (Ayaß 2008: 16). Dieser Aspekt ist auch hinsichtlich des Zusammenhangs von Kommunikation und Geschlecht relevant. Denn Doing Gender sei neben den oben beschriebenen Praktiken „in erster Linie ein kommunikatives Geschehen“ (ebd.: 19), da Menschen als ein Gender kommunizieren und sich somit als dieses in der Kommunikation hervorbringen. Dementsprechend würden auch Prozesse eines *undoing gender*,⁷⁹ die eine „Desexuierung von Geschlecht“ (ebd.) anstreben, vor allem in Kommunikation auftreten.

Rund zehn Jahre nach der Erscheinung des Einführungsbandes von Ayaß veröffentlichten Kotthoff und Nübling (2018) eine Einführung in Genderlinguistik,⁸⁰ in der sie auf ältere Studien zurückgreifen und diese um jüngere Erkenntnisse ergänzen und somit einen umfassenden Überblick über sprachwissenschaftliche Forschung mit und um das Thema Gender liefern. Neben einer Diskussion des Genderbegriffs, dem Genus und Sexus in der deutschen Sprache und Ausführungen zum Gendern, gehen sie auf viele soziolinguistische, gesprächsanalytische und in Teilen medienwissenschaftliche Forschungsfelder sowie auf kommunikative Variablen ein, und auf die Frage, inwiefern diese geschlechtsspezifisch gebraucht werden (darauf wird im methodischen Abschnitt näher Bezug genommen). Zunächst ist hinsichtlich der Titelgebung des Bandes relevant, dass sie Genderforschung, auch unter Abgrenzung zu einem binären Konzept von Gender, als auf diverse Gender ausgerichtete Disziplin verstehen: „Wir konzipieren die Genderlinguistik so, dass auch Fragen einer ‚queeren Linguistik‘ (Motschenbacher 2012) eingehen“ (ebd.: 19).

Hinsichtlich der konkreten Erforschung von Doing Gender weisen sie darauf hin, dass dies analytisch schwierig zu fassen wäre, zumal Gender realisierende Praktiken alltäglich „nur bei eklatanter Abweichung vom Erwartbaren ins Bewusstsein [treten]“ (Kotthoff/Nüblings 2018: 36). So würden im Zuge des Doing und Undoing Gender unterschiedliche Relevantsetzungen der Kategorie deutlich werden, die mithilfe eines Konzepts von „Indexikalisierung/Indizierung von kulturellem Geschlecht“ genauer gefasst werden könnten, da dieses „von vorn herein auf ein Erkennen von Typisierungsgraden und -merkmalen innerhalb von

79 Mit Hirschauer (1994) werden unter *undoing gender* „Prozesse der Irrelevantsetzung von Geschlecht“, also das „Nicht-Notiz-Nehmen von Aktualisierungen von Geschlecht“ verstanden (Ayaß 2008: 174).

80 Siehe dazu außerdem den Sammelband von Günthner, Hüpper und Spieß, der einige Jahre zuvor erschien (2012).

Handlungsgemeinschaften setzt, deren Wissen man aus der Forschungsperspektive rekonstruieren kann“ (Kotthoff/Nübling 2018: 37). Durch Anwenden dieses Ansatzes könnten somit auch „nicht-exklusive und indirekte Beziehungen zwischen stilistischen Merkmalen, Sprechaktivitäten und sozialen Kategorien, z.B. derjenigen von Gender“ (ebd.: 299) erfasst werden.

Darüber hinaus weisen sie das Konzept des Doing Gender dahingehend als kritikwürdig aus, als es in seinem situativen Verständnis auch subjektivistisch orientiert ist und somit institutionelle und massenmediale Einflüsse sowie die „Geschichtlichkeit der normativen Konzeption“ (ebd.: 38) von Gender außen vor lässt. Der (aus der Semiotik stammende) Begriff des *indexing*, den die Autorinnen in Anlehnung zu Ochs (1992) vorschlagen, würde diesem Umstand Abhilfe schaffen. Ochs machte auf der Basis, dass hinsichtlich sprachlicher und kommunikativer Handlungen genderexklusive Beobachtungen selten sind, darauf aufmerksam, dass eine Betrachtung „nichtexklusiver Verfahren der Geschlechteranzeige“ (ebd.) aufschlussreicher sei. Diese Verfahren sind historisch etabliert und können somit von den Beteiligten interpretiert werden (vgl. ebd.). Dabei wird zwischen Indexikalität erster und zweiter Ordnung unterschieden: *Indexikalität erster Ordnung* ist klar verweisend (z.B. Namen oder Pronomen), während *Indexikalität zweiter Ordnung* sich über gesellschaftlich etabliertes assoziatives Wissen ergebe (vgl. ebd.: 39). Gender würde somit mit „bestimmten Praktiken und Verhaltensweisen in eine assoziative Verbindung“ (ebd.: 40) treten, die kulturell geprägt ist.

Der Begriff des Doings wird allerdings nicht völlig abgelehnt. Er wird vielmehr in einem engen Verständnis als „bemerkbare Relevantsetzung von Gender in der Interaktion“ (ebd.: 45) gefasst, wohingegen sprachliche Merkmale und kommunikative Praktiken einen „Genderindex ergeben [können], der an einer spezifischen sozialen Stilisierung teilhat“ (ebd.).⁸¹ Grundsätzlich schließt eine situative Verfertigung oder Referenz auf Gender im Sinne eines Doing zugrundeliegende kulturelle Zusammenhänge also nicht aus.⁸² So muss es bei der Analyse von Gender in Kommunikation – bzw. in Parainteraktion – darum gehen,

81 Zum Stilisierungsbegriff siehe Kapitel 3.3.2.

82 So liegt z.B. bei Richter in Anschluss an Maihofers (2006) Konzept von Geschlecht als *hegemoniale Existenzweise* ein übersituatives Verständnis von Performanz vor: „Der Begriff der Performances meint ausdrücklich nicht, dass Geschlecht nur dargestellt wird und beliebig aufführbar ist, sondern die Performances stehen mit der gelebten und gefühlten vergeschlechtlichten Existenz der Akteur_innen in Verbindung sowie deren Historizität und den Normen der Geschlechterordnung. Sie sind aber auch Resultat ihrer aktiven Selbstkonstituierung darin und werden als Verweise auf diese gelesen“ (2022: 79).

weniger im Sinne einer situationsbezogenen Gesprächsanalyse⁸³ Gender als gänzlich aus den jeweiligen einzelnen Situationen heraus beschreibbar zu verstehen, sondern vielmehr im Rahmen einer Rekonstruktion von Doing und Indexing Gender kulturelle Hintergründe mit einzubeziehen.

Insofern kommen bezüglich YouTube-Videos Indexe für Gender in den Blick, mit denen sich die Produzent*innen ihren Zuschauer*innen gegenüber mit einer Geschlechtsidentität präsentieren, vor allem auch in Hinblick auf die parainteraktive Ansprache dieser. Somit geht es nicht nur darum, welche Hinweise YouTuber*innen für ihr eigenes Gender liefern,⁸⁴ sondern inwiefern sie auch in der (multimodalen) Ansprache der Zuschauenden (deren) Gender relevant setzen. Bereits Richter wies im Kontext von Beauty-Videos darauf hin, dass es sinnvoll ist, die „YouTuber_innen als gesellschaftlich eingebundene, aber auch handelnde Akteur_innen zu verstehen, die mit dem Geschlechterdiskurs ihrer Umgebung konfrontiert sind, aber auch eine aktive Rolle in seiner Konstituierung haben“ (2002: 78). Besonders vor dem Hintergrund der YouTuber*innen als Identifikationsfiguren, an denen sich die Follower*innen orientieren können, erhalten nahegelegte Vorstellungen von Gender Relevanz. Gleiches gilt für die Ansprache des Publikums und inwiefern dabei (implizit) Annahmen über ihr Gender und damit verbundene Erwartungen deutlich werden. So weisen Hartmann, Schramm und Klimmt (2004b: 27) mithilfe des (sozialpsychologischen) Begriffs des Selbstkonzepts darauf hin, dass sich Rezipierende, je nach dem eigenen Selbstkonzept, zu den von der Medienperson adressierten sozialen Gruppen zugehörig fühlen können. Dies gilt auch für die Gruppen bestimmter Genderzugehörigkeiten. Darüber hinaus wird auch in den Fokusgruppen im Fokus stehen, inwiefern die Teilnehmenden Gender als Kategorie relevant setzten und damit verbundene normative Vorstellungen und Anforderungen. Durch welche Merkmale dies spezifisch analytisch erfasst werden kann, wird im Folgenden genauer erläutert.

83 Vgl. Schmitt (2001, 2005).

84 Dabei geht es nicht darum, (para-)interaktive geschlechtsspezifische Besonderheiten kausal auf das Gender der jeweiligen YouTuber*innen zurückzuführen, auch, da ihre Genderidentität meist gar nicht genau bekannt ist und sich die Analyse somit auf von außen rekonstruierbare Hinweise belaufen muss.

3.3.2. ANALYSEKATEGORIEN VON GENDER IN (PARA-)INTERAKTION

Da bei der Untersuchung genderspezifischer Fragestellungen gewisse Problematiken auftreten können (s. Kapitel 2.5), die bei der Analyse sozialer Kategorien (u.a. auch *age*, *race* oder *class*) beachtet werden müssen – wie einerseits das grundsätzliche, sozial antrainierte Wissen darüber, wie man Menschen unbewusst Geschlechtern zuordnet, und andererseits die Gefahr, Zuschreibungen der Formate nach bestimmten Zielgruppen erst herzustellen respektive etablierte Geschlechterunterschiede zu reaktivieren – ist ein methodisches Vorgehen notwendig, das die Analyse geschlechtsspezifischer Merkmale auf kommunikativer Ebene ermöglicht. Dementsprechend muss die Analyse, der Prämisse der Konversationsanalyse folgend, aus den Daten erfolgen. Bezogen auf das Untersuchungsmaterial bedeutet dies konkret aus den Daten der YouTube-Videos sowie aus den Ton- und Videoaufnahmen der Fokusgruppengesprächen. Im Folgenden sollen deshalb verschiedene linguistische Ansätze und Studien zur Analyse von Gender herangezogen werden, um zu prüfen, welche für die beiden Korpora anschlussfähig sind. Dabei geht es zunächst einmal um das Videokorpus.

Zuerst ist darauf hinzuweisen, dass sich viele der angewandten linguistischen Studien auf Kommunikation beziehungsweise Interaktion in Face-to-face-Situationen beziehen und somit oft von einer vor dem Hintergrund von Wahrnehmungswahrnehmung entstehenden wechselseitigen Relevanzsetzung sozialer Kategorien ausgehen (vgl. Hausendorf 2000; Garfinkel 1967), was für (meist) monologische YouTube-Videos nicht zutrifft. Mit Ausnahme der wenigen Videos, in denen mehrere YouTuber*innen miteinander interagieren, ist somit für den Großteil der Daten ein Analyserahmen notwendig, der nicht von wechselseitiger Hervorbringung ausgeht, sondern vielmehr die sequenzielle Verfertigung von Gender relevant setzt. So weist beispielsweise Hausendorf darauf hin, dass sein Modell für die interaktionsemantische Untersuchung kommunikativ ausgedrückter Zugehörigkeit zu sozialen Gruppen – in seinem Fall die der regionalen Zugehörigkeit – nicht nur auf Face-to-face-Gespräche, sondern auch auf „schriftsprachlich bzw. massenmedial konstituiert[e]“ Kommunikation, die „anwesenheitsübergreifend auf Druck und Funk beruht“ (2000: 89), also etwa aus textlinguistischer Perspektive auf geschriebene und auditiv wahrgenommene Sprache angewandt werden kann. Zugehörigkeit meint in diesem Kontext, dass „Personen Mitglieder verschiedenartigster Gruppen sind“ und dies in Kommunikation einfließen lassen respektive diese Zugehörigkeit erst hervorbringen, was also „in und mit Kommunikation erfolgt auch und gerade kraft Sprache“ (ebd.: 3). Diesen Prozess fasst er in Anlehnung an Sacks (1992) als „soziale Kategorisierung“

(ebd.: 4).⁸⁵ Dabei geht es nicht darum, welchen Gruppen man Menschen von außen objektiv zuordnen würde, oder welchen gesellschaftlichen oder rechtlichen Gruppen sie tatsächlich angehören, sondern darum, welche Mitgliedschaften ein Individuum für sich wählt und – und dies ist der springende Punkt – in Kommunikation als bedeutend markiert und hervorbringt (vgl. ebd.: 9).

Da es bei dem Ausdruck beziehungsweise der Relevantsetzung von Gender auch für YouTube, besonders vor dem Hintergrund, dass ein bestimmtes Publikum angesprochen wird, auch um Zugehörigkeit zu einem bestimmten Geschlecht gehen kann, ist Hausendorfs Ansatz für diese Untersuchung in zweifacher Weise anschlussfähig: Zum einen spielt Zugehörigkeit auf Produktseite, vor allem in Bezug auf Parainteraktion eine Rolle. Darüber hinaus ist Zugehörigkeit im Kontext von YouTube generell relevant, wenn es beispielsweise um die Bildung oder Teilnahme an einer Community geht. Aus Rezeptionsperspektive geht es außerdem darum, ob man sich einer bestimmten Zielgruppe zugehörig fühlt. In Bezug auf Geschlecht ist dies hier insofern relevant, da die Jugendlichen in den Fokusgruppen sich als (geschlechtsspezifischen) Zielgruppen zugehörig einschätzen können oder Zuschreibungen über Zielgruppen bei gleichzeitiger Selbstabgrenzung vornehmen können, die wiederum auf geschlechtsbezogenen Unterschieden basiert. Dementsprechend beschreibt Hausendorf für die Hervorbringung von Zugehörigkeit die kommunikativen Aufgaben *Zuordnen*, *Zuschreiben* und *Bewerten* als zentral (vgl. ebd.: 5). Er definiert diese Aufgaben wie folgt: „Zuordnen meint die Darstellung von Gruppenzugehörigkeit, Zuschreiben die Darstellung gruppenspezifischer Eigenschaften und Bewerten die Darstellung einer auf Gruppenzugehörigkeit und gruppenspezifische Eigenschaften und Verhaltensweisen bezogenen Einstellungen“ (ebd.: 23f).

Sein Modell für die Beschreibung sozialer Kategorisierungen geht dabei von drei Ebenen aus, die jeweils ineinander übergehen. Demnach werden Aufgaben auf der ersten Ebene durch eine Vielzahl an möglichen Mitteln, also Verfahren (Ebene zwei) gelöst, die wiederum kommunikativ durch viele verschiedene Formen (dritte Ebene) realisiert werden. Rekonstruktion von Mitteln erfolgt „im Sinne typischer, wiederkehrender *Interaktionsmuster*“ (ebd. 126), also im Rahmen einer Sequenzanalyse und kann darüber hinaus anzeigen, ob die jeweilig erledigte Aufgabe eher im Vorder- oder im Hintergrund des Geschehens stattfindet (vgl. ebd. 127). Mithilfe welcher „sinnlich wahrnehmbare[n] Gestalt“ (ebd. 140), also mit welchen grammatischen und lexikalischen Formen die jeweiligen

85 Er verweist damit auf das Konzept der *membership categorizations*, „das als einer der Eckpfeiler für die konversationsanalytische Beschäftigung mit der kommunikativen Hervorbringung von Zugehörigkeit angesehen werden kann“ (Hausendorf 2000: 10).

Mittel realisiert werden, wird wiederum auf Form-Ebene deutlich. Die drei Ebenen bauen dementsprechend aufeinander auf. Dennoch ist methodisch je nach Einzelfall zu entscheiden, an welcher Ebene sich sinnvoll orientiert werden sollte (vgl. ebd. 144). Für die Aufgabe des Zuordnens unterscheidet Hausendorf drei Mittel: Die „*Klärung von Zugehörigkeit*“ und die „*Hervorhebung von Zugehörigkeit*“, die jeweils als Hauptaufgabe im Vordergrund der Interaktion stehen, sowie das „*Anzeigen von Zugehörigkeit*“, was unauffälliger im Hintergrund der Interaktion stattfindet (vgl. ebd. 179, Herv. i. O.). Beim Anzeigen von Zugehörigkeit wird diese eher als „ein gegebener, bereits eingeführter Sachverhalt“ gefasst, der „lediglich (wieder)aufgerufen und von Fall zu Fall aktualisiert wird“ (ebd. 180). Zuschreibungen können vorgenommen werden, indem gruppenspezifische Eigenschaften und/oder Verhaltensweisen entweder *ausdrücklich festgestellt, exemplifiziert veranschaulicht* oder lediglich *inferentiell nahegelegt* werden (vgl. ebd. 312). Ausschlaggebend bei der Rekonstruktion von Zuschreibungen ist, dass nur das eine Rolle spielt, was kommunikativ geleistet wird. Hausendorf macht unter Bezug auf sozialpsychologische und linguistische Stereotypenforschung deutlich, dass es nicht darum geht, ob stereotype Einstellungen vorliegen, „sondern ob ein Stereotyp oder Vorurteil kommunikativ aktualisiert wird und wie diese Aktualisierung empirisch nachweisbar ist“ (ebd.: 313).

Die Kategorie Geschlecht wird von Hausendorf außerdem explizit genannt und u.a. als Verweiskategorie verwendet, so auch wenn es darum geht, inwiefern man sich seiner eigenen Zugehörigkeit bewusst ist. So würde man sich bei manchen sozialen Kategorien direkt und selbstverständlich als Mitglied ansehen, während man sich der Zugehörigkeit zu anderen Kategorien eventuell gar nicht bewusst sei. Hausendorf beschreibt wie folgt:

Über Zugehörigkeit kann Ungewißheit, aber auch Gewißheit herrschen: So muß man z.B. häufig nicht lange nachdenken, welches Kästchen man bei der Option Geschlecht ankreuzt, und oft wird man nicht einmal wissen, daß man in diesem Fall überhaupt etwas ‚weiß‘ – vielleicht die stärkste Form der Gewißheit, die man mit Bezug auf Aspekte der eigenen Zugehörigkeit haben kann (ebd.: 6).

Unabhängig davon, dass die eindeutige, sichere Zuordnung zu (k)einem Geschlecht nicht die Lebensrealität aller beschreibt, weist diese Aussage darauf hin, dass Gender oft als eindeutig verstanden und unbewusst oder nebenbei behandelt wird. Auch in Bezug auf andere sei die Zuordnung von Geschlecht offensichtlich:

Soziale Kategorisierungen können aufwendig und thematisierungsbedürftig, aber auch unscheinbar ausfallen und weit unter der Bewußtseins- und Thematisierungsschwelle verbleiben: Man hört oder sieht, daß die Person, die einem gegenüber sitzt, ein *Kind* oder ein *Erwachsener*, eine *Frau* oder ein *Mann* ist (ebd.: 6).

Diese Einschätzung deckt sich mit der Beobachtung, nach der Geschlecht als Kategorie erst dann kommunikativ relevant wird, wenn sie auffällt respektive die Zuordnung zu einem Geschlecht nicht mehr offensichtlich, also ‚einfach‘ möglich ist (vgl. Garfinkel 1967; Ayaß 2008). Dementsprechend ist besonders für die Videos relevant, inwiefern YouTuber*innen es ‚für nötig halten‘ Geschlecht (z.B. in Bezug auf das angesprochene Publikum) zu thematisieren oder inwiefern sie selbst (unbewusst) Zugehörigkeit zu einem Gender ausdrücken. Besonders die Aufgabe der Zuordnung, aber auch die der Zuschreibungen und Bewertungen sowie die dafür verwendeten Mittel und kommunikativen Formen werden daher, wenn sie Gender betreffen, bei der Videoanalyse näher in den Blick genommen.

Neben dem Ausdruck von Gender-Zugehörigkeit auf diesen drei Ebenen werden weitere Kategorien für die Videoanalyse zurate gezogen, die nicht von bestimmten kommunikativen Aufgaben ausgehen, sondern von geschlechtsspezifischen Markern im kommunikativen Verhalten. So gehen Selting und Hinnenkamp (1989) mit ihrem Ansatz einer interpretativ-soziolinguistischen Stilanalyse davon aus, dass bestimmte Stile und Stilisierungen im Sprachgebrauch vorkommen, die nicht nur typisch für verschiedene kommunikative Gattungen oder für soziale Milieus sein können oder sich eher auf situative Charakteristika (wie höflich oder unpersönlich) oder die Verwendung sprachlicher Varietäten beziehen, sondern sich zudem Formen von „männlichem“ versus ‚weiblichen Stil“ (Selting/Hinnenkamp 1989: 7) finden lassen. Dabei ist zunächst bemerkenswert, wie ‚männlicher‘ und ‚weiblicher‘ Stil als binäres Paar auftritt und darüber hinaus deren Gegensätzlichkeit herausgestellt wird, was Unvereinbarkeit beider Stile suggeriert. Zudem ist entscheidend, dass Stile meist nicht explizit gemacht werden, sondern Wissen über Stile „auch mit metakommunikativen Ausdrücken benannt werden“ und Interagierende „aufgrund ihrer kommunikativen Erfahrungen und Geschichten über relativ Zeit-konstante Typisierungen verfügen, die sie in Interaktion als lokal relevante Interpretations- und Interaktionsressourcen nutzen und ggf. auch strategisch ausnutzen können“ (ebd.: 7f). Interagierende verfügen folglich über bestimmtes (gemeinsames) Wissen und evozieren dementsprechend prozesshaft gewisse für die Situation relevante Stil-Typen. Sie schaffen damit nicht nur einen Orientierungsrahmen für ihr Gegenüber, sondern gleichzeitig „schaffen und reproduzieren sie soziale Realität“ (ebd.: 8).

Von diesem (holistischen) Verständnis von Stil ausgehend definieren die Autor*innen den Begriff der Stilisierung, der wirksam wird, sobald ein in Kommunikation ausgedrückter Stil über die einzelne Situation hinausgeht und sich auf allgemeine „soziale Typen“ bezieht: „„Stilisierung“ meint die Repräsentation, Induzierung, Inszenierung etc. sozial typisierter und interpretierter Sinnfiguren in der Interaktion. Hierbei sollte zwischen Fremd- und Selbststilisierung differenziert werden [Herv. i. O.]“ (ebd.: 9). Wie schon bei Hausendorf werden auch hier u.a. Interagierenden Mitgliedschaften sozialer Gruppen zugeschrieben. Bei Fremdstilisierungen wird man entweder in gewisser Weise stilisiert oder man stilisiert die andere Person (vgl. ebd.). Die Zuschreibung zu einer bestimmten Gruppe, einer sozialen Kategorie oder eines bestimmten Merkmals könne dabei im Sinne von *Individualtypen* auf die Ganzheit einer Person bezogen werden, meistens würden Merkmale jedoch mit anderen Merkmalen gemeinsam wahrgenommen bzw. stilisierend angenommen. Bei Selbststilisierungen hingegen verallgemeinert man den eigenen sozialen Typ bzw. hebt einen Aspekt oder eine Handlung hervor, während man sie gleichzeitig einem bestimmten sozialen Typus zuordnet, als der man sich inszenieren will (vgl. ebd.). Neben kommunikativen Ressourcen, also sprachlichen Stilen, körperlichen Ressourcen wie Mimik und Gestik – die im Rahmen der Analyse in Hinblick auf (genderspezifische) Parainteraktion einer der relevanten Beschreibungskategorien bildet –, können auch äußerliche Merkmale („Haartracht, Kleidung“ (ebd.: 9)) oder übersituative Kategorien („Betätigungen, Lebensweise“ (ebd.)) für Selbst- und Fremdstilisierung genutzt oder zurate gezogen werden. Darüber hinaus geht dieser Ansatz davon aus, dass Interagierende nicht nur Stilisierungen vornehmen, sondern auch ihr Verhalten im weiteren Verlauf der Kommunikationssituation dementsprechend anpassen (vgl. ebd.: 10). Demzufolge werden bestimmten auf (Alltags-)Wissen beruhenden Stilisierungen und Zuschreibungen mit weiteren kommunikativen und verinnerlichten Registern geantwortet, die den jeweiligen sozialen Typen angemessen erscheinen. Bezogen auf die Kategorie Geschlecht kommen hier Studien in den Blick, die feststellen, inwiefern sich Personen – auch bereits gegenüber Kleinkindern – abhängig vom (angenommenen) Geschlecht des Gegenübers (unterschiedlich) verhalten (vgl. Seavey/Katz/Zalk 1975; Will/Self/Datan 1976; Mascaro et al. 2017). Inwiefern den Interagierenden Zuschreibungen bewusst sind und inwiefern die Kategorie Gender relevant gesetzt wird, kann dabei variieren. Kotthoff und Nübling betonen in diesem Kontext jedoch, dass „explizit[e] Referenzen auf Geschlechternormen“ eher selten vorkommen „im Vergleich zu Stilisierungen, die quasi immer mitlaufen und von den Mitgliedern einer Gesellschaft als Normalität angenommen worden sind (wenngleich sie kulturell hergestellt sind)“ (Kotthoff/Nübling 2018: 32). Die

Autorinnen beziehen sich hier u.a. auf Schegloff (1997), der an erkennbaren Verweisen auf die Kategorie Gender in Interaktion interessiert ist, woraufhin sie zu bedenken geben, dass die „Ausrichtung auf Genderdifferenzen hin [...] für die Interagierenden selbst aber oft versteckt [bleibt] hinter speziellen Stilphänomenen, Handlungsrollen oder Themen“ (ebd.: 34).⁸⁶

Bezogen auf die Stilanalyse respektive die Rekonstruktion stilisierender Strategien weisen Selting und Hinnenkamp zudem darauf hin, dass Interagierende und Analysierende in analysierten Situationen nicht notwendigerweise die jeweiligen Konstruktionen von Stil gleich interpretieren, sodass eine sequenzielle Detailanalyse notwendig ist (vgl. ebd.: 10). Stilisierungen sind außerdem nicht mit Vorurteilen oder Stereotypen zu verwechseln (vgl. Auer 1989). Gleichzeitig gilt: „Geschlechtliche Stile dürfen nicht mit dem eigentlichen Sprachverhalten gleichgesetzt werden. Kein Mensch praktiziert unentwegt ein und denselben Stil“ (Motschenbacher 2006: 373).

In dem gleichen Band von Selting und Hinnenkamp fasst Auer in seinen Ausführungen zu sprachlicher Natürlichkeit als Stil den Stilbegriff etwas enger und grenzt ihn von anderen linguistischen Konzepten zum (individuellen) Sprachgebrauch ab. In Anlehnung an Hymes (1974) beschreibt für ihn Stil eine bestimmte Weise zu sprechen, die zunächst unbewertet bleibt (vgl. Auer 1989: 28). Auer weist jedoch darauf hin, dass dieses Stil-Verständnis nicht mit einer scharfen semantischen Trennung von Form und Inhalt zu verwechseln ist. Stattdessen grenze sich dieser Stilbegriff insofern von anderen linguistischen Kategorien wie Kontextualisierungshinweisen (nach Cook-Grumperz/Grumperz 1976) oder auch Variation und Varietät ab, als „Stil als Menge konkurrierender, sozial interpretierter Merkmale aufgefaßt wird“ (Auer 1989: 29). Sprachlicher Stil steht demnach nicht für sich allein, also innerhalb einer Äußerung oder eines Kommunikats, sondern muss nach dieser Definition als solcher interpretierbar sein bzw. interpretiert werden, und zwar „von den Mitgliedern einer Kultur bzw. Sprechgemeinschaft“ (ebd.).

Während bei diesem Ansatz zwar von bestimmten Stilen und Stilisierungen ausgegangen wird, die sich auf Geschlecht beziehen können beziehungsweise je nach Geschlecht unterschiedlich ausfallen, gehen andere Forschungen im Rahmen linguistischer Traditionen von grundsätzlich unterschiedlichen ‚weiblichen‘ und ‚männlichen‘ Stilen aus (siehe bspw. Lakoff 1973 & 1975, die jedoch auch soziale Strukturen und Stereotypisierungen im Blick hatte).⁸⁷ Ayaß (2008)

86 Dies wird neben einer Explizit-Machung von Genderrelevanz auch im Rahmen der Fokusgruppengespräche zu beachten sein.

87 Zur kritischen Überprüfung von Lakoffs Ergebnissen siehe Günthner/Kotthoff (1991).

zeichnet diese Entwicklungen in ihrer Einführung unter anderem unter Rückbezug auf den damaligen Stand der „Genus-Debatte“ (Ayaß 2008: 27) nach und macht deutlich, dass reine geschlechtsspezifische ‚Sprachen‘ für das Deutsche nicht existieren, sich jedoch durchaus Unterschiede im Sprachgebrauch ausfindig machen lassen (vgl. ebd.: 47). Dabei greift Ayaß auf die Unterscheidung von Bodine (1975) zwischen *geschlechtsexklusivem* und *geschlechtspräferenziellem* Sprachgebrauch zurück, durch die „sich Unterschiede im Sprachgebrauch von Männern und Frauen präziser erfassen“ (Ayaß 2008: 47) lassen.⁸⁸ Im Zuge dessen verweist sie darauf, dass es keine hinreichenden Hinweise für einen geschlechtsexklusiven Sprachgebrauch gibt, sondern eher von verschiedenen Ausprägungen geschlechterpräferierter Tendenzen auszugehen ist (vgl. ebd.: 49). Außerdem macht sie stark, dass im Gegensatz zu früheren Studien, die einzelne isolierte Variablen in Hinblick auf Geschlechtsspezifika untersuchten, eine kontext- und situationsbezogene Herangehensweise der Sprachwissenschaft und Soziologie ergiebiger ist, die „nach dem Zusammenspiel kommunikativer Vorgänge im *interaktiven* Verlauf [Herv. i. O.]“ (ebd.: 65) fragt. Sie bezieht sich dabei auf eine Vielzahl von Studien, die sich mit der Untersuchung ‚weiblicher‘ und ‚männlicher‘ Gesprächsstilen befassen und für Frauen und Männer deutlich unterschiedliche Stile attestierten. Es wird erläutert, inwiefern diese Studien kritisch zu bewerten und bereits teilweise diskutiert oder widerlegt worden sind. Obwohl die Ergebnisse einzelner Studien somit oft als zu pauschaliert eingestuft werden, wird trotzdem deren Relevanz deutlich, nicht nur im Hinblick auf das Anstoßen eines wichtigen Forschungsgebiets, sondern auch aufgrund der Tatsache, dass sie immer wieder und oftmals unreflektiert rezipiert und ihre durchaus streitbaren Ergebnisse als Ausgangspunkt weiterer Argumentationen herangezogen werden (vgl. ebd.: 66). Dabei fasst die Autorin die Ergebnisse diverser Studien wie folgt zusammen: „Männer verwenden einen dominanten, hierarchischen Gesprächsstil, sind an der Sache orientiert, setzen ihre Themen durch, unterbrechen häufig. Der kommunikative Stil der Frauen hingegen ist unterstützend, emotional, persönlich, kooperativ und egalitär“ (ebd.: 65). Sie bezieht sich damit u.a. auf Studien von Aries (1987), Tannen (2007 [1990], 2011 [1986]), Wodak (1981) und im besonderen Maße auf Fishman (1978), die argumentierte, dass Frauen in Gesprächen die „Drekarbeit“ übernehmen würden, sowie auf West und Zimmerman (1983; West 1979), die das Unterbrechen als typisch männliche kommunikative Strategie kennzeichneten. Ayaß verdeutlicht anhand dieser Beispiele, indem sie gesondert auf das jeweilige Datenmaterial eingeht, inwiefern

88 Ayaß bezieht sich hier ebenfalls auf binäre Kategorien, verweist aber an derer Stelle auf Diversitäten in sozialem und biologischen Geschlecht, siehe Kapitel 3.3.1.

die Ergebnisse nicht ohne Weiteres haltbar sind und inwiefern unpassende Definitionen (wie die der Unterbrechung bei West und Zimmerman) oder einseitige Bedeutungszuschreibungen von kommunikativen Mitteln (wie für Fragen und *tag questions*) sowie eine „voreingenommen[e] Interpretationshaltung“ (ebd.: 71) dazu führen können, Ergebnisse in Daten zu lesen, obwohl aus diesen nichts eindeutiges oder gar Gegenteiliges hervorgehen würde (wie sie es für Fishman beschreibt). Neben diesem für die (Geschlechter-)Forschung grundsätzlich zentralen Postulat macht sie bezogen auf die Analyse von Gesprächsstilen deutlich, inwiefern diese nicht fruchtbar auf eine (soziale) Kategorie fokussiert stattfinden kann, sondern alle für die Bildung eines Gesprächsstils möglichen relevanten Kategorien und situationsbedingten Kontexte berücksichtigt werden sollten (vgl. ebd.: 83). Für eine geschlechtsfokussierte Kommunikationsanalyse bedeutet dies einerseits eine „Isolation einzelner kommunikativer Variablen von ihrem Kontext“ und andererseits eine „direkte Korrelierung derselben mit dem Geschlecht der Interagierenden“ zu vermeiden (ebd.: 84). Folglich bedeutet dies, nicht von einem grundsätzlich gegenderten Sprachgebrauch auszugehen, oder zu versuchen, Nachweise für diesen zu finden, sondern stattdessen kontextbezogen individuell nach Hinweisen für kommunikative Ausdrücke von Geschlecht zu suchen.

Den gleichen Ansatz vertreten auch Kotthoff und Nübling (2018). Die Autorinnen stellen einschlägige Studien vor, in denen versucht wird, sprachliche und kommunikative Handlungen in Hinblick auf mögliche Geschlechtspräferenzen zu analysieren. Dabei machen auch sie (in Anschluss an Cameron 1997) deutlich, dass grundsätzlich nicht von immer geltenden Zusammenhängen zwischen Geschlecht und kommunikativen Stilen ausgegangen werden kann, sondern „Stile (und meist auch Gattungen und andere Sprechaktivitäten) in soziale Praktiken eingebettet sind, in denen auch Gender (neben anderen sozialen Parametern) relevant gesetzt werden kann, aber nicht muss“ (Kotthoff/Nübling 2018: 39).

Konkret diskutieren sie u.a. Variablen, bei denen sich keine deutlichen geschlechtsspezifischen Unterschiede finden lassen wie bei der Phonologie, bei Fragen oder der Verteilung von Redezeiten, und welche, bei denen Tendenzen deutlich wurden wie bei der Prosodie, beim Lachen, bei Humor oder für Aufforderungen. So halten sie etwa für Prosodie fest, dass die Stimme meist für biologisch festgelegt gehalten wird, also auf Gender bezogen die Annahme herrscht, dass es deutlich unterschiedliche Stimmregister bei Frauen und Männern gebe. Dieser Unterschied sei jedoch – wie auch das soziale Geschlecht – konstruiert, sodass Stimmen sozialen Prägungen unterliegen und sich beispielsweise für verschiedene Kulturräume Stimmverläufe modulierter oder monotoner gestalten (vgl. ebd.: 53). Dementsprechend wären auch geschlechtsspezifische Unter-

schiede auf Gender zurückzuführen, also innerhalb einer bestimmten Gesellschaft erlernt, sodass sich folgende Beobachtung ergibt: „Frauen sprechen sehr häufig höher als natürlicherweise und Männer tiefer“ (ebd.: 54). Diese Differenz sei dabei nicht nur stark sozial geprägt, sondern unterlaufe auch ständigen Veränderungen. So stellen sie aus einer diachronen Perspektive mit Bezug auf die Studie von Fuchs (vgl. Berg/Fuchs/Wirkner et al. 2017) an der Universität Leipzig heraus, dass die Stimmfrequenz für Frauen heutzutage niedriger ist als noch vor 20 Jahren, was Fuchs mit einem sich ändernden Frauenbild in Verbindung bringt (vgl. Kotthoff/Nübling 2018: 55).⁸⁹ Außerdem würden Frauen mit tieferer/abgesenkter Stimme ernster genommen werden (vgl. ebd.).

Über die Tonhöhe hinaus sind außerdem genderspezifische Unterschiede für Modulationen zu erkennen. So würden deutschsprachige Frauen mehr modulieren, also größere Stimmgrundfrequenzschwankungen und Tonhöhenbewegungen vornehmen sowie von einer „Knarrstimme“ Gebrauch machen, was in eine Art „Singsang“ münden und eher kindlich wirken würde und „Emotionalität und Expressivität, aber nicht Autorität vermittelt“ (Kotthoff/Nübling 2018: 57). Außerdem würden sie Endungen von Äußerungen öfter mit steigendem Tonverlauf realisieren, was insgesamt unsicherer wirke (vgl. ebd.). Bei deutschsprachigen Männern hingegen seien Stimmverläufe grundsätzlich monotoner (vgl. ebd.: 58). Eine Knarrstimme würde bei Männern eher am Ende einer Äußerung vorkommen und dort auch mit Männlichkeit assoziiert werden (vgl. ebd.). Die Autorinnen verweisen außerdem darauf, dass Lautstärke, Sprechtempo und der Wechsel des Stimmverlaufs ebenfalls Gender anzeigen können und dass jeweils „die emotionalere, unsicherer bzw. kindlicher wirkende Variante mit Weiblichkeit assoziiert“ (ebd.) wird, ohne jedoch näher zu bestimmen, welche Variante das jeweils wäre. Um an dieser Stelle keine vorschnellen oder stereotypbehafteten Schlüsse zu ziehen, sei hier lediglich erwähnt, dass auch diese Variablen nicht losgelöst von anderen Faktoren als eindeutig genderspezifisch zu beschreiben sind. Ohne, dass im Rahmen dieser Analyse von YouTube-Videos dezidiert auf Hz-Angaben geachtet wird, lassen sich dennoch Anknüpfungspunkte finden, insbesondere, wenn es um Modulationen oder Auffälligkeiten hinsichtlich sich ändernder Tonhöhen geht.

Außerdem lassen sich für Scherzkommunikation und Humor geschlechtsbezogene Unterschiede in Gebrauch sowie in der Bewertung ausmachen. Dabei gebe es nicht nur Witze, deren Inhalte sich konkret auf Geschlechter beziehen

89 Demnach würden tiefere Stimmen bei Frauen, im Gegensatz zu höheren, mädchenhaften, besser zum Bild heutiger Frauen passen, die selbstsicher sind und „nicht mehr beschützt werden“ (vgl. Ahlers 2017) müssen.

und somit „geschlechtsrelevant“ sind wie etwa Witze, die auf (geschlechtlichen) Stereotypen basieren, sondern Humor und scherzhafte Kommunikation würden darüber hinaus generell andere Themen behandeln und sprachlich anders realisiert werden (vgl. ebd.: 294f). Männer und Jungen würden beispielsweise auf einem höheren „Angriffsniveau“ (ebd.: 294) und aggressiver scherzen (vgl. Apte 1985, nach ebd.). Deppermann und Schmidt (2001) haben außerdem für Jungen-Cliquen gezeigt, inwiefern auch hierarchische Strukturen in Scherzkommunikation einfließen können und thematisch u.a. von Schadenfreude und Verspotten Gebrauch gemacht wird (vgl. ebd.: 295). Generell waren bei Humor auch soziale Faktoren und das Alter ausschlaggebend.

Daran anschließend liefern Kommunikationsstudien auch Hinweise darauf, inwiefern Lachen geschlechtspräferentiell genutzt wird, sodass Frauen und Männer in unterschiedlichen Situationen unterschiedlich viel lachen (vgl. ebd.: 298). Besonders bezogen auf das „soziale Lachen“ (ebd.), also ein kommunikativ genutztes Handeln, das verschiedene Funktionen erfüllen kann und keine (körperliche) Reaktion auf z.B. etwas Lustiges ist, scheinen Frauen anteilig häufiger zu lachen als Männer (vgl. ebd.). Kotthoff und Nübling stellen allerdings deutlich heraus, dass Scherzpraktiken und Lachen nicht exklusiv gegendert sind, sondern auch hier soziale Stilisierungen genutzt werden, um sich als bestimmter sozialer Typus zu inszenieren (vgl. ebd.: 299). Dementsprechend werden bei der Analyse Humor und Lachen nicht im Fokus stehen, sondern als möglicherweise relevante Kategorien dann hinzugezogen, wenn sie mit anderen Variablen zusammen auftreten.

Neben sprachlichen Kategorien weisen die Autorinnen auf die Relevanz der Gestaltung des Äußerlichen hin. Obwohl es sich bei ihren Ausführungen um sprachwissenschaftliche Untersuchungen handelt, sind äußerliche Merkmale wie Kleidung, Frisuren etc. als semiotische Ressourcen zu verstehen und unter linguistischer Perspektive zu beschreiben. In diesem Kontext wurde beispielsweise gezeigt, inwiefern Jugendliche ihre Kleidung für Identitätsarbeit nutzen (vgl. Staubach 2018; 2020). Auch Kotthoff und Nübling weisen darauf hin, dass Jugendliche den Übergang zum Teenager aktiv durch die selbstbestimmte Änderung ihres Kleidungsstils kennzeichnen (vgl. Kotthoff/Nübling 2018: 26). Gerade bei Kleidung spielt Gender offenkundig eine Rolle, da bestimmte Kleidungsstücke und -schnitte wie in Kapitel 2.5 bereits angerissen eindeutig geschlechtsspezifisch zugeordnet werden. Dementsprechend ist Mode bzw. die eigene Auswahl der Kleidung ein Ausdruck von Geschlecht, der derartig gesellschaftlich etabliert ist, dass er als „habitualisiert“ (ebd.: 20) angesehen werden kann. Bereits bei Garfinkels „Agnes-Studie“ (1967) waren Kleidung, aber auch Frisur und Make-up Hauptkriterien für das Doing Gender, womit auch „heute noch die

auffälligsten Genderstilisierungen“ (Kotthoff/Nübling 2018: 26) realisiert werden. Dementsprechend ist besonders für die Videos des Beauty-Korpus im Hinterkopf zu behalten, dass alle durch Make-up gesellschaftlich etablierte Zuschreibungen aufgerufen werden, die Make-up und Schminken als Praxis ‚weiblich‘ konnotieren. Das Gleiche gilt für Videos, die sich mit Kleidung mit ‚weiblichen‘ Schnitten oder aus der Frauenabteilung befassen. Auch diese Kategorien, die die Videos auf den ersten Blick als deutlich gegendert auszeichnen, sind für die Analyse zu berücksichtigen, allerdings im Zusammenhang mit weiteren Hinweisen für die Anzeige von Geschlecht. Schließlich können semiotische Ressourcen auf anderer Ebene, etwa über (kommunikatives) Verhalten im Sinne einer „Neutralisierungsarbeit“ (Kotthoff/Nübling 2018: 35), ausgeglichen werden. Für die FIFA-Videos kann die Kategorie Mode auch in Bezug auf die Kleidung innerhalb des Spiels, also auf die Trikots der Spieler*innen, aufschlussreich sein (vor allem vor dem Hintergrund, dass im Bereich (Männer-)Fußball Mode und Sponsoren nicht nur für Trikots, sondern auch für beispielsweise Schuhe ein zunehmend medial auftretendes Thema sind).

Alle hier genannten Kategorien und Gesprächsvarianten sind also als sozial konstruiert anzusehen, mithilfe derer Gesprächsteilnehmende Hinweise auf Geschlecht erhalten und sie bewusst liefern können. Für die Videoanalyse wird dementsprechend untersucht, inwiefern die YouTuber*innen, auch wenn sie nicht in einer eigentlichen Interaktionssituation sind, sondern die Zuschauenden ihr Gegenüber bilden, Gender realisieren. Dies kann durch explizite Äußerungen geschehen, die die Kategorie Gender entweder aufrufen oder sogar zum Thema innerhalb des Videos machen. Dementsprechend kommen die parainteraktiv genutzten Modalitäten und Ausdrucksressourcen auch unter einer genderspezifischen Perspektive in den Blick, was sich neben den bisher erwähnten Kategorien auch auf mediale Komponenten und die Inszenierung des Raums bezieht. Zudem ist bezogen auf die Anrede der Zuschauenden sowie grundsätzlich in den Videos interessant, ob und inwiefern die YouTuber*innen gendern, da dies sowohl auf eine inklusivere ‚Einladung‘ einer diverseren Zielgruppe als auch auf ein grundlegendes Bewusstsein für geschlechtspolitische Themen schließen lassen würde. Genauso aufschlussreich sind jedoch implizite Hinweise auf Gender, die nicht notwendigerweise bewusst relevant gesetzt werden. Dies kann über den Ausdruck von Zugehörigkeit (über die Aufgaben des Zuordnens, Zuschreibens und Bewertens) geschehen, es kann sich in Stilisierungen ausdrücken oder/und mit kommunikativen Handlungen einhergehen, für die bereits an anderer Stelle Geschlechtspräferenzen nachgewiesen worden sind. Der Analyse liegt dabei immer ein situatives, kontextbasiertes Verständnis zugrunde.

Für die Fokusgruppengespräche gelten grundsätzlich die gleichen Voraussetzungen dafür, wie Gender innerhalb Interaktion hervorgebracht werden kann. Allerdings geht es hier nicht darum, wie die Teilnehmenden in Interaktion ihre Geschlechtszugehörigkeit über einen bestimmten Sprachgebrauch ausdrücken.⁹⁰ Stattdessen kommt bezogen auf die Fokusgruppengespräche Hausendorfs Zugehörigkeits-Begriff noch einmal dahingehend in den Blick, als es sich hier nun um Face-to-face-Interaktion handelt. Werden die Fokusgruppengespräche bezogen auf die anderen Forschungsfragen wie etwa bezüglich des Nutzungsverhaltens der Jugendlichen oder ihrer Meinung zu YouTuber*innen generell aus inhaltsanalytischer Perspektive beleuchtet, so wird es im Kontext von Pseudokommunikation und Gender darum gehen, auch aus gesprächsanalytischer Sicht auf die Daten zu blicken. Dementsprechend ist der Ausdruck von Zugehörigkeit dann von Interesse, wenn es um die Kategorie Geschlecht in Bezug auf bestimmte Videoinhalte geht, also Jugendliche sich beispielsweise selbst einer gegenderten Zielgruppe zuordnen oder abgrenzend Zuschreibungen von Zielgruppen einzelner Formate vornehmen. Vor diesem Hintergrund werden auch noch einmal die Aufgaben des Zuordnens, Zuschreibens und Bewertens relevant.

Grundsätzlich ist für die Gruppengespräche im Hinterkopf zu behalten, dass Aussagen der Jugendlichen bezüglich ihrer Mediennutzung vor dem Hintergrund sozialer sowie situativer Faktoren stattfinden und nicht mit der ‚Wirklichkeit‘ übereinstimmen müssen. Auch ist nicht unwahrscheinlich, dass sich die Teilnehmenden an dem orientieren, was als sozial erwünscht gilt (vgl. Ayaß 2008: 141). So hält Ayaß für das Fernsehen bezogen auf Geschlecht (unter Bezug auf Cornelißen 1998) fest, dass „Selbstauskünfte zu Programmreferenzen [...] ein Ort von *Selbstdarstellung* und mithin auch eine Situation [sind], an dem *Geschlechtsstilisierungen* („gender display“ im Sinne Goffmans) erfolgen [Herv. i. O.]“ (ebd.: 141)

90 Branner forschte z.B. zu genderspezifischer Scherzkommunikation in Mädchengruppen. In diesem Zusammenhang liefert sie einen Überblick über Untersuchungen zu geschlechtsspezifischen Interaktionsstrukturen in Mädchen- und Jungengruppen, wobei sie teilweise voreilige Zuschreibungen sowie grundsätzlich dichotome Forschungsparadigmen kritisiert (vgl. 2003: 2ff).