

in der theoretisierenden Erfassung des Tonfilms allgegenwärtig; und sie verändert sich. Denn sollte man meinen, dass die Wiedervereinigung von Körper und Stimme, wie sie in den skizzierten Debatten zunächst schmerzlich vermisst und sodann als Novum akzeptiert wurde, von Seiten der Kritiker:innen und Theoretiker:innen der anbrechenden Tonfilmära euphorisch begrüßt wird, irrt man. So schlägt dem Tonfilm und seinen ›sprechenden Schatten‹ eher Skepsis entgegen. Im Folgenden soll daher nachgezeichnet werden, inwiefern das Anthropomorphe und Anthropophone in Form des *Audioviduum*s sich auch im Rahmen der Tonfilmdebatten als relevant erweist und inwiefern es fortgeführt zum Anlass für medientheoretisierende Ansätze der 1920er bis 1930er Jahre wird.

3.3 | Körper und Stimme – Zur Emergenz des *Audioviduum*s im Tonfilm

Die Frage nach der Verknüpfung von Filmbild und Ton wird – wie nachgezeichnet werden konnte – bereits und vor allem seit der Etablierung der sinn-separierenden Bewegtbild- und Tonmedien immer wieder diskutiert. Im Anschluss an die bisher fokussierten Debatten um die Zentralstellung von menschlichem Körper und Stimme in Stummfilm und Radio ist es dabei interessant zu sehen, wie die Tonfilmdebatte sich an diesen abgesteckten Feldern der vorangegangenen Medientheoriebildungen weiter arbeitet. Zum einen fällt dabei auf, dass die nun technisch mögliche Realisierung von *Audioviduen* auf zwiegespaltene Reaktionen stößt. Die Etablierung von Film und Radio als nicht nur Techniken, sondern Künste hatte es zuvor erforderlich werden lassen, jeweils aus der Not eine Tugend zu machen und somit die Stummheit des Stummfilms und die Immateriellität des Radios zur Grundlage ihrer Kunstfertigkeit und zum Alleinstellungsmerkmal (z. B. gegenüber dem Theater) zu erklären. Der dabei auf den Menschen gerichtete Fokus der Diskussionen half hier bei der Legitimation, um die Tragweite und Bedeutung dieser neuen Kunstformen für die Kultur, die ganze Menschheit und jeden einzelnen zu betonen; schließlich versprachen die ›neuen Medien‹ nun innovative, einzigartige und nie-dagewesene Formen der Selbsterkenntnis und des Kulturschaffens. Mit der technischen Etablierung des Tonfilms ab 1927 nun scheinen die Tugenden jedoch erneut gefährdet. Das zuvor diskursiv verhandelte und dann abgelegte Ideal des sprechenden Menschen zugunsten stimmfreier Körper und körperfreier Stimmen endet in Bezug auf den Tonfilm also in einem Dilemma, das jedoch erneut deutlich werden lässt, dass sich auch im Zuge der Ablehnung des neu geschaffenen *Audioviduum*s dennoch seine Relevanz als Denkfigur bestätigt – selbst wenn diese, wie zuvor bereits erprobt, darin besteht, es zu verneinen und zurückzulassen, zugunsten der Anerkennung der (sonstigen) Fähigkeiten des neuen Mediums. Wie diese Diskussionen sich entwickelten soll

im Folgenden erneut schlaglichtartig nachvollzogen werden, um die explorative Studie zur Genese des Audiovidiums damit abzuschließen.

3.3.1 | Körper und Stimme im Fokus der Tonfilmdebatten

Die Experimente zur Kopplung von Film- und Tonaufzeichnung sind so alt wie die Entwicklung der jeweiligen Techniken selbst (vgl. Müller 2003, 13; 75ff.)³⁰². Laut Müller (2003) ist es dabei allerdings nicht verwunderlich, dass die Durchsetzung des Tonfilms rund 30 Jahre auf sich warten ließ, denn ihrer These zufolge war die Etablierung des Films als Kunst dringend darauf angewiesen, dass dieser nicht im Kontext des rein maschinellen Dokumentarismus verblieb, sondern sich stattdessen zunächst als Medium des Fiktionalen etablierte, das nicht die Wirklichkeit schlicht abbildete, sondern einen Gestaltungsspielraum und damit Potenzial für künstlerische Interventionen ließ. Dennoch finden sich auch in der von Müller skizzierten »Tonfilmpause« (ebd., 13) bzw. »Tonfilmzäsur« (ebd., 82) zwischen Mitte der 1910er und Ende der 1920er Jahre Zeugnisse von Kontakten mit dem Tonfilm – so z. B. die bereits oben erwähnte Schilderung einer Vorführung des Kinetophons von Felix Salten, die er 1913 im *Berliner Tageblatt* veröffentlicht. Den ersten Eindruck der neuen Technik schildert er dabei folgendermaßen:

»[...]die weiße Leinwand schimmert auf, und wir kriegen endlich das Kinetophon zu sehen. Da ist eines jener entsetzlichen Zimmer, die man aus unzähligen Kin-

302 So arbeitete Edison parallel zu seinen reinen Filmsichtungsmaschinen ab den 1890er Jahren auch an der Entwicklung des Bild und Ton kombinierenden (wenn auch nicht direkt synchronisierenden) Phono-Kinetoskops (auch Kinetophon). Ludwig Stollwerck ließ diese Bild-Ton-Maschinen bereits ab 1895 auch in Deutschland neben den Edison'schen Kinetoskopen aufstellen (vgl. Müller 2003, 75). Ebenso waren bereits auf der Pariser Weltausstellung 1900 sogenannte ›Tonbilder‹ zu hören und zu sehen (vgl. ebd.) und für die etablierten und fest installierten deutschen Kinosäle produzierte Oskar Messter bereits zwischen 1906/07 und 1912 mit Tonwalzen kombinierte Filme (vgl. ebd., 76). Auch in den Texten der Etablierungszeit des Tonfilms selbst findet sich teilweise ein Bewusstsein dafür, dass die Entwicklungen vom Ende der 1920er Jahre gar nicht so neu waren – z. B. bei Kalbus (1935): »Der Ton in allen seinen Arten versuchte schon immer sich mit den Bildern auf der Leinwand zu vermählen, wenn auch zunächst mit wesensfremden und unkünstlerischen Mitteln. [...] Gleich nach der Erfindung des Kinematographen und Phonographen im Jahre 1894 beschäftigte es den Erfindergeist Thomas Alva Edison, diese beiden Erfindungen so zu kombinieren, daß gleichzeitig mit den bewegten Bildern der dargestellten Personen deren Sprache und Gesang (eventuell begleitende Musik oder Naturgeräusche) auf einem Phonographen mit aufgenommen und bei der Projektion wieder mit zu Gehör gebracht werden konnten. Diese Einrichtung, ›Kinetophon‹ nannte sie Edison, war der erste Vorläufer der modernen Nadelton-(Tonfilm-)Apparatur. Trotzdem können wir das Kinetophon heute nicht als die erste Tonfilmparatur bezeichnen, denn es fehlte ihr ein für den modernen Begriff ›Tonfilm‹ wesentliches Erfordernis: der Synchronismus« (ebd., 5).

toppvorstellungen kennt, ausgestattet mit jenem Schauder erregenden Luxus schwelender, persischer Sofas, präparierter Blattpflanzen, schlechter Gipsbüsten, erbärmlich kopierter Renaissancestühle. Ein junger, befrackter Herr kommt herein, tritt ganz vor, so weit, daß der Rahmen der Leinwand ihm die Füße abschneidet, und ... beginnt zu sprechen. Richtig und laut zu sprechen, wie der Herr Voredner, den wir soeben in *natura* gesehen und gehört haben. Zuerst denkt man natürlich an ein Grammophon, denkt bloß daran, daß eben ein Grammophon zu der Kinoaufnahme redet. Aber das dauert nur ein paar Sekunden. Dann wird Bild und Stimme vollkommen eins. Man hat nämlich dem Herrn da oben auf den Mund geschaut, und da merkt man, wie lebendig das ist, wie fabelhaft lebendig. [...] Die Stimme, die da zu uns redet, scheint aus diesem auf die Leinwand projektierten Mund eben zu entstehen, scheint aus der Brust dieses photographierten Menschen aufzusteigen. Deshalb vergißt man das Grammophon: vergißt überhaupt, daß der Mann dort ja gar nicht zu uns redet, sondern, weiß Gott wann im amerikanischen Atelier Edisons vor etwelchen Managern gesprochen hat. Dann zeigt der Mann noch etliche Proben. Er setzt ein Pfeifchen an den Mund und wir hören den Pfiff losschallen. Im selben Moment, in dem der Mann die Backen bläßt.« (Salten 1992 [1913]), 47f.)

In dieser kurzen Schilderung ist etwas von der Faszination zu spüren, die in ähnlichem Maße auch die ersten (stummen) Kinobilder erzeugt haben müssen, nur dass es dieses Mal ein neues, erweitertes Staunen gibt, das z. B. nicht mehr den abgeschnittenen Füßen des filmisch Vortragenden gilt, sondern eben der erstmaligen Kopplung von filmischem Menschenbild und Stimme in ihrer Ko- bzw. Kompräsenz. Das Verschmelzen von jeweils maschinell projiziertem Körper und maschinell reproduziertem Klang wird dabei als ein kurzer Moment der Irritation beschrieben, der aber dann der Illusion einer audiovisuellen Ganzheitlichkeit des Menschenmotivs weicht, die als doppelte Emergenz anzusehen ist: Mit der Manifestation eines dieser ersten Audioviduen im Raum der von Salten geschilderten Filmvorführung, ebenso wie in seiner Beschreibung dieses Vorgangs in der Zeitung, bildet sich das Audioviduum als Phänomen und Argument heraus.

Saltens Text ist dabei gleich in mehrfacher Hinsicht bemerkenswert, weil seine subjektiven, journalistischen Schilderungen dieses ersten Tonfilmerlebnisses von 1913 bereits vorausweisen auf die Aspekte, die auch in der erst rund 15 Jahre später beginnenden Kernzeit der Tonfilmdebatten eine Rolle spielen werden. Neben dem genannten Aspekt der Verschmelzung von Bild und Ton (die hier in Bezug auf das Menschenmotiv erläutert wird) ist es zudem einerseits der Status der menschlichen Stimme innerhalb der filmischen Geräuschwelt, die sich ankündigt, und andererseits die Frage nach den ›Gefahren‹ und dem Mehrwert, die die neue Technikkopplung mit sich bringt.

So verharrt Saltens Text nur zu Beginn seiner Schilderung bei der Zentralstellung der menschlichen Sprache, nennt dann aber – ergänzend zum »Pfeifchen« – als weitere Beispiele aus der Filmvorführung auch noch eine Glocke, eine Geigerin, eine Sängerin und bellende Hunde, die das neue Funktionsprinzip des Tonfilms veranschaulichen. »Das Publikum lacht und applaudiert. Die Geigerin und die Chansonette waren unnötig. Die beiden schönen Hunde in ihrer Natürlichkeit hätten genügt« (ebd., 48f.) – so wertet Salten das gesehene-gehörte Programm aus und lässt darin erkennen, dass der Klang der menschlichen Stimme (ob gesprochen oder gesungen) – so faszinierend er initial geschildert wird – am Ende auch in Konkurrenz zu anderen Geräuschen (Instrument, Tierstimme, Dingsound) steht.³⁰³

Doch trotz des einfallslosen, insgesamt stark musiklastigen Programms, das Salten kurz resümiert und dann kritisch demontiert, bleibt für ihn am Ende die Erkenntnis, dass es sich beim Tonfilm um eine bahnbrechende Entwicklung handelt: »Das alles ist so abgeschmackt, wie die Erfindung selbst, die uns solchen Schmarrn vermittelt, genial ist« (ebd., 49), so schreibt er und sieht auch schon einen Gewöhnungsvorgang initiiert:

»Nach den ersten zehn Minuten erscheint einem das Kinetophon als etwas ganz selbstverständliches. Die Bilder bewegen sich. Das ist man vom Kino her längst gewöhnt. Nun reden und singen sie auch. Nach zehn Minuten staunt man gar nicht mehr darüber. Als ob das so einfach und so natürlich sei.« (Ebd., 50)

Die anthropomorphisierende Übertragung von Gesang und Rede auf die Bilder selbst (nicht nur auf die in ihnen sichtbaren Gestalten) ist dabei möglicherweise ein Detail, das auf eine Veränderung der Perspektive vorverweist – nämlich die sich erst parallel in der Stummfilmkritik ausbreitende Erkenntnis, dass für den Film möglicherweise nicht die ›Menschen‹ allein relevant sind, die von der Leinwand sprechen, sondern ebenso die Dingwelt – und das gilt im erweiterten Sinne schließlich auch für die Leinwand bzw. die hinter ihr liegenden Lautsprecher und den filmischen Gesamtaufbau selbst, der sich akustisch zu äußern lernt. Dennoch macht Saltens Artikel deutlich, dass er vor allem in der sprachlichen bzw. stimmlichen Vertonung den größten Nutzen der neuen Technik sieht. Dies zeigt sich auch in der Frage nach Risiken und Wertigkeit der neuen Technologie, die er zum Abschluss seines Artikels anekdotenhaft verdeutlicht:

303 Zumal oftmals gerade die menschliche Stimme die noch junge Tontechnik vor Herausforderungen stellt, die es in Bezug auf Klanglichkeit und Verständlichkeit erst noch zu meistern gilt (vgl. z. B. Arnheim 2002 [1932], 191).

»Hansi Niese, die neben mir sitzt, sagt auf einmal: ›So ... jetzt können wir Schauspieler einpacken ... jetzt braucht ma uns gar nimmer ...‹ Ganz ängstlich und melancholisch ist sie geworden, während sie das sagt. Und am Ausgang, wie wir uns draußen in der Sonne wieder finden, ruft mir ein Freund lachend entgegen: ›Jetzt werden die Autoren ganz überflüssig! Aber ich glaube, es stimmt nicht. Weder die ängstliche Melancholie, noch der Scherz. Im Gegenteil. Das Leben wird ins Hundertfache gesteigert durch diese Erfindung, wird gleichsam lebendiger durch sie, und der Tod kann es jetzt nicht mehr so vernichten, wie einst. Diese Erfindung wird alles, was sterblich an uns gewesen ist, erhalten und aufbewahren bis in ferne Tage. Diese Erfindung wird große Persönlichkeiten in ihrer Echtheit erhöhen, und sie wird die großen ›Macher‹ unerbittlich entlarven. Deshalb werden uns die guten Schauspieler und die guten Autoren künftig beinahe noch nötiger sein als bisher. So bequem wie unsere Väter und Vorfäder dürfen wir nun nicht mehr leben. Denn wir werden eine strengere Nachwelt haben, und sie wird genauer prüfen können, was ihr überliefert wurde, als wir es vermochten. Eine Nachwelt, die hören kann und sehen, was die Toten geleistet haben ... die wird nichts auf Treu und Glauben hinnehmen brauchen. Wir haben das noch getan. Aber von jetzt an ist es damit zu Ende.« (Salten 1992 [1913], 50f.)

Salten schildert hier also einerseits die vermeintliche Angst der Filmschauspieler:innen und -autor:innen vor dem bereits wieder Obsoletwerden ihres gerade erst in den letzten 20 Jahren neu gelernten Ausdrucksrepertoires – im Angesicht eines Mediums, dessen neue Qualität, so wird hier wie selbstverständlich angenommen, nicht mehr in Schauspiel und Drehbuch, sondern in der Aufzeichnung und Zugänglichmachung der ›reellen‹ Welt bestehen muss. Salten betont dabei den historischen Wert der Erfindung, die ihm dazu in der Lage scheint, den Tod zu überwinden³⁰⁴ und als dokumentarisches Medium bisher ungekannte und überzeitliche ›Erhöhungen‹ und gleichzeitig ›Entlarvungen‹ des Menschen zu ermöglichen. In der befürchteten Abkehr von Darstellung und Dialog, Schauspieler:innen und Autor:innen, scheint sich hier eine weitere Problematik des Audioviduum anzukündigen.

Mit Saltens frühem Text sind dabei bereits exemplarisch einige der zentralen Achsen der sich ab Ende der 1920er Jahre intensivierenden Tonfilmdebatten vorgezeichnet: die Verschmelzung von Körper und Stimme bzw. das Verhältnis von Bild und Ton, die Frage nach dem künstlerischen Ausdruckspotenzial und/oder der (rein) dokumentarischen Funktion des Tonbildes sowie die Rolle des Men-

304 Ein ähnliches Argument war bereits in Bezug auf die ewige Jugend der Schauspieler:innen im Stummfilm bei Polgar (1992 [1911/12], 164) zu finden (siehe Kapitel 3.1.2) und die Kopplung von Tod und Film erinnert natürlich auch an Bazins spätere »Mumie der Veränderung« (2009, 39).

schen innerhalb des filmischen Soundgefüges bzw. die Rolle von Geräusch und Musik. Diese Aspekte sollen in den nächsten Abschnitten im Fokus stehen.

Zur Verschmelzung von Körper und Stimme

Die Tatsache, dass sich zum Zeitpunkt des aufkommenden Tonfilms der Film und teils auch das Radio als eigene Kunst- und Medienform bereits nahezu uneingeschränkt etabliert hatten, zeigt sich vor allem daran, dass die frühen Debatten um die Einführung und Durchsetzung des Tonfilms diesen in der Tendenz eher kritisch denn euphorisch begrüßen. Wie bereits angedeutet ist es vor allem der Verlust der Reduktion auf einen Kanal (Körper *oder* Stimme), der in den Schriften zum Tonfilm als bedauernswerter Mangel geschildert wird. Die Wiedervereinigung von Auditivem und Visuellem (vielfach beschrieben am Beispiel von Körper *und* Stimme) birgt die Gefahr, so wird geschildert, dass der gerade erst als Kunst angekommene Film seinen künstlerischen Gestus verliert zugunsten einer rein dokumentarischen Praxis.

Bereits im Vorfeld der Etablierung des Tonfilms wird dieser Verlust an künstlerischer Wertigkeit thematisiert – z. B. bei Hugo Münsterberg, der in Bezug auf den (noch) stummen Film antizipiert:

»Even if the voices were heard with ideal perfection and exactly in time with the movements on the screen, the effect on an esthetically conscientious audience would have been disappointing. A photoplay cannot gain but lose if its visual purity is destroyed.« (Münsterberg 1916, 203)

Und auch wenn die bereits im vorherigen Kapitel geschilderten Positionen von Münsterberg und Bloem in Bezug auf den Stummfilm teils deutliche Unterschiede aufweisen, so sind sie sich in diesem Punkt einig, denn auch Bloem beschreibt bereits Anfang der 1920er Jahre die Gefahren des Tonfilms für die Kunst des Films:

»So bleibt noch die dritte Frage: welcher künstlerischen Wertung untersteht der sprechende Film? Einige Männer des Films, die ihrer Überzeugung schriftlich Ausdruck gaben, vertreten die Ansicht, daß der sprechende Film aus künstlerischen Gründen zu wünschen sei (Gad und Pordes). Doch mit dem Fortfall der Stummheit würde der letzte, der Hauptdamm niederbrechen, der das Lichtspiel vor der hemmungslosen Hingabe an die Darstellung der blanken Wirklichkeit schützt. Ein völlig zuchtloser Realismus würde den letzten Rest von Stilisierung, den jetzt noch das armseligste Filmmachwerk aufbringt, hinwegfegen; heute schon erkennen wir ein kaum gebändigtes Streben zu diesem Ziele hin, das namentlich durch verschwenderische Titelgabe und durch die Verwendung wirklicher Natur als Hintergrund künstlerischen Geschehens gefördert wird.« (Bloem 1922, 36)

Interessant an beiden Äußerungen – sowohl der Münsterbergs als auch Bloems – ist, dass der Tonfilm hier wie selbstverständlich mit Stimme und Sprechen gleichgesetzt wird. Über die Geräuschwelt wird an diesen Stellen weniger nachgedacht; zentral erscheint also die Antizipation einer dadurch ausgelösten, zu großen Nähe zur *›menschlichen Realität‹*. Bloem sieht durch dieses Mehr an Wirklichkeits-treue das Ausdruckspotenzial des Films gefährdet. Deswegen lehnt er auch zu Zeiten des Stummfilms bereits jegliche Sprachsubstitute im Kino kategorisch ab – seien dies Zwischentitel, Vertonungen durch Tonmacher:innen oder Sprecher:innen oder (im schlimmsten Fall) in Form von Lichtspielopern, in denen echte Sänger:innen ergänzend zur Leinwand operieren. Ihre Kombinatorik ist ihm zufolge eine zu vermeidende Kunstvermischung, da in der Verquickung das jeweilige Ausdrucksmedium seine Eigenständigkeit verliere. Die Individualität des Mediums geht hier also mit einer Idee der Einheit und Reinheit einher. Doch am Ende macht sich Bloem um die Kunst des Films keine Sorgen, wenn er resümierend prophezeit:

»Wehe dem Lichtspiel wenn es seine Stummheit verliert! Aber alle Befürchtungen sind überflüssig. Selbst wenn diese Erfindung einst in vollkommener Form vorliegen wird – durchdringen wird sie nicht.« (Ebd., 38)

Dass sich diese Annahme als falsch erweisen wird, zeigt sich natürlich erst Ende der 1920er und zu Beginn der 1930er Jahre mit der nachhaltigen Etablierung des Tonfilms. Während bei Münsterberg und Bloem allerdings die Ablehnung des Sprechfilms möglicherweise noch nachvollziehbarer erscheint, weil zu Zeiten einer noch nicht zufriedenstellend oder flächendeckend möglichen Realisation des Tonfilms eine Aufwertung des Stummfilms vonnöten ist, so wird mit dem Aufkommen erfolgreicherer Tonfilmtechnik die Kritik an dieser zunächst dennoch lauter. So beschreibt Willi Wolfradt 1929 den Ton als unnötige Einschränkung der Filmkunst, die doch gerade erst ihre Formen zur Reife entfaltet habe und nun ihre rein visuellen Gestaltungspotenziale durch die Verklanglichung wieder zu verlieren drohe, weil das Auditiv das Visuelle blockiere:

»Zu viel Gespräch gefährdet den Film, weil es ihn fesselt, – wie man denn an jedem Sprechfilm beobachten kann, daß er überhaupt nicht vorwärts kommt. Die Pantomime ist extreme Bühnenkunst, daher die bekannten Forderungen von Einheit der Handlung der Zeit, der Oertlichkeit hier ganz besondere Geltung haben, ferner auch die Tilgung des Individuellen durch Masken, durch kostümliche Irrealisierung usw. Das Schweigen bedeutet da eine weitere Uniformierung und echt bühnenhafte Limitation.« (Wolfradt 2002 [1929], 446)

Erneut ist es also das Gespräch und generell das Sprechen, das für den Tonfilm ein Problem darstellt, weil es eine Nähe zum Theater und seinen Formen suggeriert, die dem Film als eigenständige Kunst abträglich ist. Interessant für die Frage des Audiovidiums erscheint zudem die Beschreibung des Verlusts von Individualität durch den Verlust der Stimme, der im Sinne künstlerischer Reflexion und einer Art Purifizierung allerdings wünschenswert ist. Wolfradt befürwortet dies, so ist anzunehmen, im Hinblick auf eine überindividuelle Allgemeingültigkeit und Übertragbarkeit von Spielhandlungen. Die Stimme hingegen erscheint als individualisierende Barriere, weil sie das Audioviduum eher als ganzheitliches Gegenüber erscheinen lässt denn als abstraktes Motiv, das anschlussfähig bleibt für reflektierte Rezeptionen.

Mit Wolfradts Einschätzung, dass das Gespräch das Problem des Films sei, stimmt auch René Clair überein, der dazu aus ebenfalls tonfilmkritischer Perspektive dennoch einen – möglicherweise nicht ganz ernst gemeinten – Kompromissvorschlag unterbreitet:

»Wir Stummfilmgetreuen wollen uns der tönenden Invasion nicht länger verschließen. Machen wir gute Miene zum tönenden Spiel. Fassen wir den Stier bei den Hörnern. – Einen Ausweg sehe ich zum Beispiel im Tonfilm ohne Dialog. Vielleicht lässt sich durch ihn die Gefahr noch bannen. Man könnte sich doch vorstellen, daß die das Filmband begleitenden Geräusche und Klänge die Masse so sehr unterhielten, daß sie auf den Dialog verzichtete. So könnte man ihr eine der Bilderwelt weniger abträgliche Illusion der Wirklichkeit verschaffen. Das Publikum dürfte dieser Lösung jedoch kaum freiwillig zustimmen.« (Clair 1952 [1929], 95)³⁰⁵

Doch auch in Aussicht solcher Optionen bleibt für Clair die Kunst des Films gefährdet durch den Tonfilm, weil allein der Stummfilm das Potenzial habe als visuelle Kunst seine Wirkung zu entfalten. Vor allem die sprachlichen Äußerungsformen sind für ihn eher Störung denn Gewinn, weil sie die visuelle Sprache des Films übertönen:

305 An anderer Stelle schlägt er zudem die alleinige Ersetzung der Zwischentitel durch Sprache im Tonfilm vor: »Und doch wäre es möglich, sich des Tons zu bedienen, ohne auf die Vorteile des Stummfilms zu verzichten. Der gesprochene Text könnte an Stelle des Zwischentextes treten und nur als Hilfsmittel benutzt werden. Ich stelle mir einen bündigen, neutralen Text vor, der keinerlei visuelle Opfer erforderte. Mit etwas gutem Willen und Verstand müßte dieser Kompromiß zu erreichen sein. Ich bin nur nicht sicher, ob man sich lange daran halten würde, denn die Versuchung durch den künstlichen Dialog, durch die fertige Phrase, vor der wir uns einst aus dem Theater in das goldene Schweigen der Leinwand flüchteten, ist groß. Der Tonfilm läßt uns bei aller Jugendtorheit noch Hoffnung; vom Sprechfilm erwarten wir trotz seiner Anfangserfolge nicht viel« (Clair 1952 [1929], 103).

»Der Stummfilm hatte gerade erst eine neue, noch lallende, aber reine und weltweite Lyrik geschaffen, und nun tönen uns aus dieser Welt abgedroschene Phrasen entgegen, die Dichtung und Traum von der Leinwand treiben. Die imaginären Worte, die wir den Schatten in den Mund legten, waren unendlich viel reizvoller als die präzisen Sätze, die sie nun äußern. Damals sprachen sie die Phantasie an und hatten zum besten Bundesgenossen das Schweigen. Künftig werden sie Albernheiten von sich geben, und man kann nichts dagegen tun, kann sie nicht mehr zum Schweigen bringen.« (Ebd., 102f.)

Das Verschwinden des Stummfilms als eigene Kunstform ist es also, das von ihren bisherigen Verfechter:innen befürchtet wird, wobei für die Krise aber nicht einfach der Ton generell, sondern vor allem das sprechende Audioviduum als »sprechender Schatten« verantwortlich ist. An ihm entzünden sich somit die intensivsten Diskussionen.

Doch es gibt auch Autor:innen, die den Sprechfilm als Chance des Films begreifen, weil sie den Stummfilm und seine gestalterischen Mittel zum Ende des Jahrzehnts an seine Grenzen gekommen sehen. Als Beispiel für eine solche Position sei Herbert Leisegang erwähnt, der sich in einem Artikel mit dem Titel *Sprechfilm – Rundfunk – Bildfunk – Filmfunk – Zukunftstheater* (2002 [1929]) nicht nur sehr ausführlich mit den Potenzialen des Tonfilms, sondern auch den Möglichkeiten seiner Verbreitung (über Funk) beschäftigt.

»Die kommende Vereinigung von Rundfunk und Film bedeutet mehr als eine blosse Entwicklungsphase, sie scheint vielmehr ein bedeutsamer Schritt auf dem Wege einer bisher so oft gescheiterten Verwirklichung künstlerischer Massenwirkung werden zu wollen. Schauspielkunst, trotz mancher Bestrebungen doch immer noch mehr oder weniger eine Angelegenheit Einzelner, kann so *auf dem Umwege über die mechanischen Künste* eine Sache Aller, eine wahrhafte Volkskunst werden.« (361; Herv. i. O.)

Während die früheren Filmkritiker:innen also im Kino und »Kientopf« noch einen Apparat zur Verdummung des Volkes hatten kommen sehen, sieht Leisegang sehr klar einen Bildungsauftrag, der sich mit Hilfe der Techniken Film und Funk (gerade in ihrer Kombination) realisieren ließe. Mit der Etablierung des Tonfilms wähnt er sich darum »am Vorabend umwälzender technischer Ereignisse« und er kommt zu dem zeitgeistigen Schluss: »[...] die mechanische Kunst ist im eigentlichen *die Kunst unseres Jahrhunderts*« (ebd., 362; Herv. i. O.).

Da Bildfunk und Filmfunk (also Fernsehen) aber noch nicht vollends realisiert sind, sieht Leisegang die größten Konsequenzen der Audiovision zunächst für den Film gekommen, der seiner Meinung nach von den neusten Entwicklungen nur profitieren kann: »Der Film hatte diese Erfindung nötig, er brauchte diese Auffri-

schung!« (Ebd., Herv. i. O.). So sei mit »dieser technischen Umgestaltung die Steigerung der künstlerischen Qualität des Films« (ebd., 364) möglich, weil die banalen Zwischentitel durch das dichterisch verfasste »lebendig gesprochene Wort« (ebd.) austauschbar seien und so durch die Sprache als kunstvolle Form auch der Kunststatus des Films gesichert werde.³⁰⁶ Gleichzeitig sei er dennoch lediglich eine »mechanisierte« (ebd., 362) Form der ihm ursprünglich zugrundeliegenden Kunst – des Theaters, dem sich der Film durch den Tongewinn nun wieder anzunähern schiene.³⁰⁷ Den mechanischen »Künsten« käme es laut Leisegang also zu, dass sie erstmals die »Keimzelle« des Theaters – die Präsenz von sprechenden Körpern – in ihre Einzelteile zerlegt und dadurch die »uralte theatralische Totalitätskunst« in »Scheingesichte« und »Scheintöne« gespalten habe (ebd.). Diese erstmalige Trennung in »Augen- und Ohrenkunst« (ebd.) (Film und Radio) sei nun aber von den mechanischen Entwicklungen erneut eingeholt worden:

»Der Hauptvorwurf, den alle Kunstreunde gegen den Film immer und immer wieder erhoben, betraf die fehlende Wortgestaltung. Da man bis vor ungefähr 10 Jahren noch nicht daran glauben konnte, daß diese zweidimensionalen Schattenfiguren auf der Leinwand einmal anfangen könnten zu sprechen, bezeichnete man diesen Mangel irrigerweise als den Kernpunkt der künstlerischen Minderwertigkeit des Films. Dieser Gedanke muß heute infolge der überraschenden technischen Entwicklungen fallen gelassen werden.« (Ebd., 263)

306 Wobei in Leisegangs Ausführungen doch relativ klar wird, dass er den Film dabei weiterhin als die mindere Kunstform (z. B. im Vergleich zum Theater) ansieht. So urteilt er über den Film: »Mehr als ein formales Kunstwerk zu sein, wird er indessen mit seinen technischen Mitteln niemals zu erreichen imstande sein« (Leisegang 2002 [1929], 364).

307 Darin sieht er allerdings nicht nur für den Stummfilm und Rundfunk selbst (in ihren bisherigen künstlerischen Formen), sondern auch für das Theater eine Gefahr gekommen: »Wir hatten ausdrücklich hervorgehoben, daß unsere beiden modernen Spezialkünste, der Film und der Rundfunk, Sprößlinge der einen großen Urzelle Theater sind. Die furchtbarste Gefahr wird dem Theater dann drohen, wenn beide Künste sich auf ihren gemeinsamen Ursprung zurückbesinnen und sich, jede für sich voll ausgebildet, zu einer ungeheuren Weltmacht neu zusammenschließen. Die heutigen Formen beider sind als Durchgangsstadium zu diesem Endzustand anzusehen. Wie die visuelle Kinematografie ihrem Machtbezirk das Wort eingliedert, so bemächtigt sich von der anderen Seite der nur für das Ohr eingestellte Funk des Bildes (der Fernseher wird ein Bestandteil des Funks werden), der *Bildfunk* ist für uns heute schon ein geläufiger Begriff. Fernseher und Funk bilden die letzte Ergänzung: denn beide bedeuten die Überwindung des irdischen Raumbegriffs, der eine für das Auge, der andere für das Ohr. Jedem Nachdenkenden wird es einleuchten, daß es dann nur noch eine Frage der Zeit bedeutet, bis die starre Fotografie im eigenen Heim sich zur kinematografischen Beweglichkeit umwandelt. Das bedeutet natürlich auch das Ende der Lichtspieltheater in jetziger Form. In der Weise wie heute der Funk das Wort von einzelnen Sendern aus ins Haus liefert, wird Europa dann von einigen Zentralen mit dem Lautfilm versorgt« (Leisegang 2002 [1929], 366f.).

Leisegang beschreibt also die (gefundene)³⁰⁸ Audiovision als die neue, zeitgemäße weil mechanische Form der Kunstverbreitung, bleibt aber insofern dennoch konservativeren Haltungen treu, als er auf lange Sicht weiterhin das Theater als die eigentliche, größte Kunst des Menschen begreift. Das liege vor allem an der zwar schon gesteigerten, aber immer noch unvollkommenen Technik des Films:

»Trotzdem sehen wir uns dieses mechanische Gebilde näher an, wird ein letzter ungelöster, unkünstlerischer Rest bleiben. Die Zweidimensionalität filmischen Geschehen auf der Leinwand dissoniert mit der Dreidimensionalität der Musik oder Sprache. Schattenhafte Flimmergestalten bedienen sich der Ausdrucksmit- tel bluthafter Menschen.« (Leisegang 2002 [1929], 367)

Der Kontrast von akustischer Realitätsnähe und visueller Artifizialität erscheint Leisegang somit als ein Problem, dass sich potenziell erst mit der Einführung des stereoskopischen und farbigen Films lösen ließe. Dies würde aber wiederum eine zu große Nähe zum Theater bedeuten und daher eine solche »mechanische Form« einer »Guckkastenbühne« (ebd.) den Kunstwert des Films erneut gefährden. Und ein weiterer Punkt spricht gegen das rein mechanische Theater: das Fehlen des »soziologischen Moments« (ebd., 368), womit Leisegang erneut die (schon in Bezug auf den Stummfilm verhandelte) Anwesenheit der Darsteller:innen im Theaterraum im Gegensatz zu ihrer Absenz im Kino meint. Diese »menschliche warme Berührung vom Menschen auf der Bühne zum Menschen im Zuschauerraum« scheint dabei das letzte zu sein, »was ihm von der Technik nicht genommen werden kann« (ebd.). Leisegang entpuppt sich hier als doch nicht so fortschrittsliebend, wie es zunächst den Anschein hat. Implizit geht er davon aus, dass sich das Kino als Kunst durch seine voranschreitenden technischen Innovationen und Optimierungen auf lange Sicht selbst überflüssig mache, wenn man nicht genug aufpasse:

»Die Technik wird den angegebenen Weg zielsicher gehen, aber – man gestatte die subjektive Ansicht – sie wird im siegenden Vorwärtsstürmen ihr eigenes Grab schaufeln. Heute noch zieht man dem Theater mit seinen menschlichen Zufällen das präzisere Arbeiten der Maschine vor, es wird die Zeit kommen, wo man sich vor ihrer exakten Präzision angeödet wegwendet wird, und je vollendet die Technik die königlichste Kunst, die Schauspielkunst, wird reproduzieren können, desto

³⁰⁸ Das Fernsehen sieht er dabei auf lange Sicht als Ziel dieser Entwicklung – und erneut ist und bleibt das »Schauspiel« der zentralste Inhalt: »Damit sind wir beim idealen Endstadium ange- langt: Wir setzen uns häuslich in den Klubsessel, drücken auf den Knopf und sehen und hören Schauspielkunst wieder in erster Vereinigung. Das ist der Technik höchster Triumph« (Leisegang 2002 [1929], 367).

tiefer wird die Sehnsucht wieder nach der verlästerten Menschenbühne. Diese Sehnsucht jedoch erfordert eine andere Kultur, die mit heutiger Zivilisation und Technik in *geistigen* Fragen nichts mehr zu tun hat.« (368; Herv. i. O.)

Doch von der »exakten Präzision« ist der Tonfilm zu Beginn noch weit entfernt, weshalb auch die technischen Finessen in der Umsetzung der Audiovision thematisiert werden. Vor allem das Verhältnis von Technik und menschlichen Akteur:innen (vor Kamera und Mikrofon sowie auf der Leinwand) steht dabei im Mittelpunkt

Laut|Sprecher: Der Ton als technisches Problem kinematografischer Körper

Die neue Filmtechnik stellt sich für alle Filmschaffenden – von den Darsteller:innen bis zu den Techniker:innen – als eine Herausforderung dar. »Wir hören, daß die berühmten Filmstars von Hollywood verzweifelt nach Sprechlehrern rufen, die sie in die Anfangsgründe der Sprechtechnik einweihen können« – so schildert Leisegang (2002 [1929], 365) das Geschehen und auch Kurt Karl Eberlein (2002 [1928]) beschreibt die Veränderungen, die das neue Medium für die spielende Zunft bedeutet³⁰⁹:

»Zweifellos wird die neue Filmtechnik den Schaunaturalismus, Hörnaturalismus, Spielnaturalismus verstärken und im Guten wie im Bösen manches verändern. Dies trifft zunächst die Filmspieler, die Stars, die weder sprechen noch singen können. Der wie eine Uhr repetierende, sprachlose Schaumime der Jupiterlampen, der grünlich wie eine Wasserleiche in dem Geschrei und Gesumm wie ein Gliedermann schau-spielte [...] und bei Grammophongeheul, Lachen und Scherzen Tragisches andeutete, und der schöngebaute Gelbstern, halb Mannequin halb Schönheitskönigin, all das wird wohl verschwinden auch wenn man den Spielenden, Worte-kauenden fremde Stimmen mitgibt. Die Sprechschauspieler treten hervor, dem Theater entzogen oder entfremdet, dem Filmtheater gewonnen. Was sie sonst im Ablauf eines Abends in spielerischer Ordnung auf der Raumbühne mit organischer tönender Wortgeste gestaltet, das wird nun in Raten und Fetzen, in Tagen und Wochen dem atemlosen Perateuren – Weh dem, der niest! – vorgespielt, den Apparaten zugespielt. Dieser Schein eines Scheines, dies Theater eines Theaters, dies Bild im Bild wird zwischen Natur und Bühne schwankend ein peinliches Zwischen-
ding, ein Alsob und Wiewenn schaffen, das weder Film noch Theater im Zwischen-
reiche der Räume und Flächen eine eigene Gesetzlichkeit fordert.« (Ebd., 287)

Alfred Polgar hingegen beschreibt die Auswirkungen eher von Seiten des Publikums aus, bei dem die – im wahrsten Sinne des Wortes – Übereinstimmung von

309 Siehe hierzu auch Müller (2003, 174ff.).

gesehenem und gehörtem Menschen zu irritierenden Effekten zu führen droht, nämlich wenn man sich die Stimmen anders vorgestellt habe:

»Dahin nun die wundervollen Texte, die wir von lautlosem Munde ablasen, die leidenschaftlichen Schweigereden der Liebenden, die strahlende Witzigkeit, die den Komikern von stummen Lippen floß! Nicht mehr haben die schönen Frauen im Film die schönen, die edlen Männer die edlen Stimmen, die wir, sie nicht vernehmend, vernahmen, nicht mehr geht das Gesprochene (dank seiner Sublimierung durch die Unhörbarkeit der Sprechenden) geradezu ins Herz, allen höchsten Sinnes voll, den die Phantasie der Hörer ihm zu schenken vermochte [...]. Vorbei mit der vollkommenen Melodie, die wir den holen Frauen des Films auf die Lippen legten [...].« (Polgar 1929, 176f., zit. n. Müller 2003, 146)

Während Polgar hier dem Stummfilm nachtrauert, entpuppt sich Eberlein in seiner kritischen Sicht auf den Tonfilm wie Leisegang als eigentlicher Verfechter des Theaters³¹⁰, zu dem er das Kino in seiner neuen Form nicht wirklich als Konkurrenz sieht, weil der geeignete Schauspielstil für dieses neue Medium – wenn überhaupt – noch gefunden werden müsse. Sein Fazit fällt daher in Bezug auf das Konkurrenzverhältnis von Theater und Tonfilm und vor allem den Status des letzteren als Kunst recht eindeutig aus:

»Warum soll es nicht Zeiten ohne Kunst geben, Zeiten der Zwischenkunst? Vielleicht ist der Hiatus, diese schöpferische Pause mehr als nötig? Vielleicht wäre es gut, wenn das Wort einmal ruht, bis es wieder rein, stark und zeugungsfähig ist? Vielleicht brauchen die Maske, der Spielraum, der Spielkörper endlich etwas Ruhe, um wieder neu und eigen zu werden? Vielleicht wächst schon, während der Zirkus, das Stadion, das Kino in Schweiß und Nebel verblaßt, wie eine Osterblume das Theater von Morgen aus dem ungepflegten Boden, in dem so viele Tote ruhn. Vielleicht muß erst der kinetische Mensch ganz sterben, der dumme, bebende, flimmernde Flächenmensch der Maschinenlinse?« (Ebd., 297)

Die »Geburt des Tonfilms«, wie sie von einigen Autoren genannt wird (z. B. Kalbus 1935, 5; Arnheim 2002 [1932], 189), läuft somit bei Eberlein auf einen Tod des Menschen – bzw. einen Opfer- oder Stellvertreter:innen Tod des Audiovidiums –

³¹⁰ Und in gewisser Weise auch des Stummfilms wenn er schreibt: »Vielleicht kommt der Tag, da wir müde der Augen- und Ohrenhurei reumütig zum stummen, farblosen raumlosen Flächenfilm zurückkehren und seine tonige Naturferne ebenso erfrischend finden, wie man nach dem Genuß der Schab- und Farbstichblätter zu dem reinen Umrißstich und zu den schlüchten Holzschnitten zurückkehrte« (Eberlein 2002 [1928], 285).

hinaus,³¹¹ um daraus einen künstlerischen Neuanfang für das Theater möglich zu machen.

Doch nicht nur das mangelnde Vermögen der Schauspieler:innen und das seltsame Verhältnis vom Sprechfilm zum Theater – also von mechanischer Kunst zu ›tatsächlicher‹ Kunst – gefährden den tonfilmischen Genuss, auch die Technik scheint noch nicht in allen Details ausgereift. Dies äußert sich z. B. besonders schön in einer Kritik Rudolf Arnheims zum Film *DAS LAND OHNE FRAUEN* von 1929:

»Soweit wäre alles gut gegangen, wenn nicht der Tonfilm die Zeit für gekommen gehalten hätte, die Personen nun auch sprechen zu lassen. Es geschah also daß Conrad Veidts alte Mutter den Mund öffnete und mit der Trompetenstimme eines versoffenen Elefanten Trostworte flüsterte, die der Sohn nicht minder sonor, jedoch unter Zermalmung aller S-Laute, beantwortete. Nun war das Eis gebrochen, das Sprechen nahm überhand [...], aber keinem wollte die Sache so recht gelingen. Sie sprachen alle mit Pedal, und das S schien von der Zensur gestrichen: ›Verlaffe mich nicht!‹ flehte die Dame Nr. 68 und zerriß einen wichtigen Brief, was ein prasselndes Geräusch hervorrief.« (Arnheim 1929, zit. n. Müller 2003, 239)

Müller (ebd.) weist allerdings darauf hin, dass es für die zeitgenössische Kritik und vor allem für die Berliner Kritiker:innen nahezu zum guten Ton gehörte, den Tonfilm im Hinblick auf seine frühen Mängel zu verreißen. Auch das Publikum schien für diese Probleme sensibilisiert, wie eine Schilderung von Walter Jerven aus dem Filmkurier von 1930 nahelegt:

»Wenn die Hauptdarsteller sprechen, sieht man auf der Leinwand deren Partner, die nicht sprechen. Auf diese Weise will man die Klippe des Nichtübereinstimmens von Ton- und Mundbewegungen umsteuern! Das Publikum lacht.« (Jerven 1930, zit. n. Müller 2003, 306).

³¹¹ Dieser gewünschte Tod des Menschen widerspricht dabei nicht Eberleins eigentlichem anthropozentrischen Denken, das er im Hinblick auf die Kunst (seiner Zeit und generell) so zusammenfasst: »Das perpetuum mobile ist der Mensch, der Zeitgenosse, der Undankbare, der immer Hosiannah und Kruzifixe bereit hält und dabei immer der gleiche Komparse bleibt, während sich auf der Weltbühne das Schicksal unverständlich gestaltet. Ohne ihn gäbe es weder Film noch Filmproblem, weder Stoff- noch Formfragen, denn der Beschauer schafft unbewußt an dem Schauwerk mit, und es ist zumeist sein Spiegelbild, das wir im Bilde der Filmbühne wiederfinden, weil es für ihn bedacht, auf ihn bezogen, durch ihn verändert, scheinbar den Bildwillen Einzelner, tatsächlich aber die Bildwelt aller zeigt« (Eberlein 2002 [1928], 285). Auch hier also verbinden sich Anthropozentrismus und Anthropomorphozentrismus (des Menschen als Spiegelbild auf der Filmbühne) zu einer Medienreflexion.

Dass aber gerade in der »Nichtübereinstimmung« von Bild und Ton auch Potenziale für seine künstlerische Gestaltung liegen, das wird hier (noch) nicht bemerkt bzw. wird es an anderen Stellen sehr wohl verhandelt, wie noch zu zeigen sein wird. Dennoch – so lässt sich bis hierher feststellen – scheint gerade die Wiedergabe von Sprache für den Tonfilm zu seinem Beginn eine der größten Herausforderungen – sowohl was Frequenzen und Verständlichkeit als auch die Lautstärke betrifft.

Bei Hans Erdmann findet sich in Reaktion darauf eine hierarchische Liste von Klangsorten, die für das Kino Ende der 1920er Jahre am annehmbarsten erscheinen. Diese lautet: 1. Gesangsstimmen, 2. Geräusche, 3. Soloinstrumente (kleines Ensemble), 4. Großes Orchester, 5. Sprechstimme (vgl. Erdmann 2002 [1929], 344). Damit belegt das Audioviduum quasi sowohl den ersten als auch den letzten Platz zugleich. Erdmann begründet die Entscheidung, die Sprechstimme als Schlusslicht zu setzen folgendermaßen:

»An letzter Stelle steht die Sprechstimme, natürlich nicht etwa ihre Verständlichkeit, sondern ihr ästhetischer Wert. Es ist bisher noch kaum annähernd möglich gewesen, feinere Nuancierung des Sprechorgans, ohne die wir natürlich die Sprache als Kunstmittel nicht verwenden können, durch die Tonkamera hervorzubringen. Die Sprache wird gegenwärtig also nur kaschiert gebraucht werden können, wenn gewissermaßen die Unzulänglichkeit der Wiedergabe zum eigentlichen Wirkungsmittel wird, z. B. ein Lautsprecher in der Szene selbst, komische, groteske und grausige Figuren usw.« (Erdmann 2002 [1929], 344)

Die Klangqualität lässt also noch zu Wünschen übrig; wird dieses Problem der stimmlichen Inszenierung allerdings durch Musikalisierung getilgt, so ist der menschlichen Stimme (im Gesang) dennoch der erste Rang der kinematografischen Geräuschwelt zuzuordnen.

Diese Diagnose mangelnder Qualität scheint dabei allerdings ein vorübergehendes Problem, das der technische Fortschritt lösen kann. Folgt man erneut Leisegang (2002 [1929]) so ist es ein entgegengesetztes Qualitätsmerkmal, das in der Kombination von Bild und Ton den Unterschied macht – und zwar nicht die schlechte, sondern die im Vergleich zum Filmbild gerade zu gute Tonqualität. So findet sich bei Leisegang die Schilderung der Unvereinbarkeit von Dreidimensionalität des Raumklangs und Zweidimensionalität der graugestuften Leinwand, die nur durch 3D-Techniken zu kompensieren wäre. Das gleiche Problem schildert ebenso Eduard Rhein (2002 [1929]), der den »Gleichlauf« als das zentrale Kriterium des Tonfilms bezeichnet. Damit meint er gleichermaßen zeitliche und räumliche Dimensionen von Filmbild und -ton, die stets erst technisch miteinander gekoppelt werden müssten. »Der Künstler mag ihn anders sehen – dem Techniker aber bleibt der Tonfilm immer Bild plus Ton in genau festgelegter zeitlicher

und räumlicher Zuordnung zueinander« so beschreibt Rhein (ebd., 394) die Tücke bzw. Lücke des Films, die im Endeffekt besonders zwei Probleme mit sich bringt: zum einen den »Unterschied zwischen dem zweidimensionalen Bild und der dreidimensionalen räumlichen Wirklichkeit« des Tons (ebd., 396) und zum anderen die möglicherweise auftretenden räumlichen wie zeitlichen Fehlpassagen zwischen Bild und Ton, wobei das zeitliche Problem in der Regel die Synchronität betrifft, das räumliche die verhältnismäßige Positionierung der Tonquelle im Bild (z. B. ein sprechender Mensch) zur Tonquelle im Raum (i. d. R. ein Lautsprecher).

Die zeitliche Ebene ist die von vielen Kritiker:innen am ehesten adressierte, wenn es um eine Qualitätseinschätzung des Tonfilms geht. Auch wenn dabei stellweise anderes als der Mensch (bei Rhein z. B. ein Presslufthammer, ebd., 395) als Beispiel dient, um die Wichtigkeit des »Gleichlaufs« von Bild und Ton zu betonen, stellt der Mensch in dieser Hinsicht ein besonders illustratives Exempel dar, wie auch Pander feststellt:

»Wer gegenwärtig Tonfilme besucht, hat sich vom Reiz der Neuheit noch nicht völlig losgelöst; er hat täglich gelesen, was Synchronismus ist, wie das Frequenzband beschaffen sein muß usw. und kennt alle Fehler, auf die er stoßen kann, theoretisch – mit dem Erfolge, daß er sie ganz bestimmt bei der Vorführung bemerkt. Er sieht den Sprechern und Sängern auf den Mund, wie er es in der Wirklichkeit niemals tun würde, und verfolgt die Lippenbewegungen, um festzustellen, ob sie synchron sind, ob vielleicht zu einer Aufnahme eines amerikanischen Sprechers nachträglich ein deutscher Dialog hinzusynchronisiert ist – all diese Untugenden mögen die Zukunftshörer des Tonfilms vielleicht nicht haben, so daß sie eine stärkere Illusion haben könnten, soweit es sich um die bisher erwähnten Einzelheiten handelt.« (Pander 1930, 470, zit. n. Müller 2003, 305)

Die zeitliche und personelle »Synchronität« ist also das eine Problem, dem sich der Tonfilm ausgesetzt sieht. Aufbauend auf der Tatsache, dass Bild und Ton auch im Tonfilm eben voneinander getrennte Sphären bleiben, schildern Rhein u. a. darüber hinaus die räumlichen Aspekte, die es bei seiner Inszenierung im Kinosaal zu beachten gilt:

»Für den Tonfilm bedeutet diese Tatsache, daß es an sich nicht so sehr darauf ankommt, ob sich etwa beim sprechenden Menschen der Lautsprecher auch tatsächlich dort befindet, wo der Kopf des Schauspielers erscheint. Wir lokalisieren den Klang, der sogar vom anderen Ende der Leinwand herkommen könnte, auf den Kopf des sprechenden Menschen. Daraus aber schließen zu wollen, die Tonquelle brauche also nicht bewegt zu werden (mechanisch oder elektrisch), wäre zum mindesten etwas übereilt.« (Rhein 2002 [1929], 398)

Nicht die räumliche In-Eins-Setzung von visueller und räumlicher Tonquelle ist demnach ein Problem, sondern die Bewegung, die der Film seinen Gegenständen ermöglicht:

»Nicht weil die Tonquelle feststeht, läuft der Schauspieler von seiner Stimme fort, sondern weil der bildlichen Veränderung nicht eine akustische Veränderung entspricht. Überhaupt ergeben sich bei dieser Anordnung recht merkwürdige Erscheinungen. Es kann beispielsweise der Lautsprecher rechts stehen, wenn der zuerst links auf dem Bilde erscheinende Schauspieler, auf den wir den Ton dann lokalisiert haben, sich nach rechts bewegt, also in Wirklichkeit auf die Schallquelle zu, so bleibt doch der Eindruck bestehen, er entferne sich von seiner Stimme.« (Ebd., 399)

Was sich hier anhand des Menschenmotivs im Film offenbart ist somit die eigentliche Divalidität der Individualität des Audiovidiums und somit des Mediums Tonfilm. Auch wenn der immersive Eindruck einer klingenden und sichtbaren Welt im Einklang miteinander in ihm möglich scheint, ist diese Verschmelzung von Bild und Ton – hier verhandelt an Körper und Stimme – immer eine Illusion bzw. ein medial erst zu schaffendes Kunststück. Zur Lösung des Problems schlägt Rhein eine Art Stereo- bzw. Surround-System vor, wenn er schreibt:

»Der Fehler ließe sich schon dadurch beheben, daß hinter der Leinwand in mittlerer Höhe mehrere Lautsprecher aufgestellt würden, und zwar ein Lautsprecher links, einer in der Mitte und einer rechts, die vom Vorführer entsprechend den Bewegungen der sprechenden Person abwechselnd eingeschaltet würden.« (Ebd.)

Allerdings ist für andere eben nicht nur das Problem der Bewegung, sondern tatsächlich auch das der Lautsprecher-Position hinter der Leinwand virulent. So z. B. in der Einschätzung von Alexandre Arnoux, der ebenfalls weniger das zeitliche als das räumliche Synchronitätsproblem des Tonfilms sieht:

»Der erste Eindruck ist niederdrückend. Da der Lautsprecher hinter der Leinwand angebracht ist, kommen die Stimmen immer vom gleichen Punkt her. Die Synchronisation ist zwar gelungen, wirkt aber auf den Zuhörer irritierend. Analysiert man seine Irritation, so findet man heraus, daß die Übereinstimmung von Lippenbewegung und ausgesprochenen Silben strengere Anforderungen an die Natürlichkeit stellt und optische Anhaltspunkte im Raum erfordert. So wohnt man einem skurrilen Lustspiel bei, dessen Personen mit dem Mund Worte formen, denen ein geheimnisvoller, bauchredender Chorführer tief aus dem Innern der Leinwand Ton gibt. Zur Vertuschung dieses Gebrechens und wohl auch aus technischen Gründen vermied der Regisseur während der Wortwechsel Veränderungen der Kamera-Einstellung, was eine gewisse filmische Monotonie, eine schwerfällige Gangart,

ein Stagnieren zur Folge hat und unmittelbar ans Theater mit seinem an der Rampe gesprochenen Text erinnert. Sobald die Schauspieler zu sprechen aufhören, wird es wieder lebendiger, so daß man sich schließlich fragt, ob die Stimme dem Ausdruck nicht mehr schadet als nützt.« (Arnoux 1928, zit. n. Clair 1952, 90f.)

Was Arnoux beschreibt ist also genau das, was die Verfechter:innen des Stummfilms in Bezug auf den Tonfilm fürchteten, nämlich, dass sein eigenes visuelles Bildrepertoire, das er sich über die Jahre erarbeitet hatte, durch den Ton nicht nur akustisch überspielt wird, sondern dass es durch die Anforderungen des Tons zu tatsächlichen Einwirkungen auf die Visualität des Bewegtbildes kommt – und zwar, indem dieses seine Bewegung verliert, weil etwa statische Einstellungen gewählt werden oder auch die Schnittfrequenz sinkt, um den Menschen im Bild nicht – im wahrsten Sinne – das Wort abzuschneiden (oder zumindest nicht dessen visuelles Aussprechen gegenüber dem Erklingen). Arnoux gibt dabei zu, dass er »ein alter Filmliebhaber« sei und dass ihn das »Spiel in Schwarz-Weiß mit seinem Bildertakt und der Verbannung der leidigen Vokabel« (ebd., 91), das der Stummfilm bietet, mehr zusagt. Und erneut wird dabei die Frage nach dem Tonfilm eine Frage nach Leben und Tod:

»Und nun macht eine ungebettene Erfindung alles zunichte; muß einen das nicht verbittern? So viel Mühe und Hoffnung, um nun doch wieder beim Theater zu landen, dieser zu Tode gehetzten Methode, um das Wort wieder auszugraben und gar auszuposaunen. Ich weiß nicht, ob das Geburt oder Tod ist, was wir erleben. Eins aber steht fest: am Film- und Tonhimmel gehen gewaltige Wandlungen vor; man muß Augen und Ohren offenhalten. [...] Ist es eine Wiedergeburt oder Tod? Man kann's nicht wissen.« (Ebd.)

Auch bei Arnoux wird somit das mediale Problem des Tonfilms anthropozentriert über das Menschenmotiv verhandelt, worin allerdings deutlich wird, dass die große Kunst des Films eigentlich in seiner Ablösung von eben diesem Menschenmotiv bestehen müsste.

Und eine letztes Beispiel sei beigefügt für die Probleme, die sich vor allem an das Menschenmotiv des Films koppeln lassen, wenn es um die technische Verschmelzung seiner körperlichen und stimmlichen Verfasstheit geht. Denn auch Hans Pander beschäftigt sich mit diesen Fragen, wenn er den Tonfilm auf seine Möglichkeiten der Illusionserzeugung hin befragt.

»Der Zuschauer wünscht den Eindruck zu haben, daß der Sänger oder Redner die Töne mit dem Munde, nicht mit einem anderen Körperteile hervorbringt; wohin seine Wünsche bei mehrstimmigem Gesang, bei Orchesterstücken, bei der Wiedergabe des Stimmengewirrs bei Menschenmassen, bei dem ohrenbetäubenden

Geräusch eines Maschinensaals gehen, weiß man nicht, und in gewissen Fällen bringt der Tonfilm sogar Töne, denen im Bilde überhaupt keine Schallquelle entsprechen soll. Das ist der Fall, wenn etwa ein Kulturfilm von einem unsichtbaren Redner durch einen Vortrag begleitet wird, ebenso, wenn der Tonfilm nur die bisher übliche Orchestertermusik ersetzen soll. In diesem Falle wird sich aber niemand daran stoßen, wenn die Lautsprecher unterhalb der Bildwand dort ihren Platz haben, wo sich sonst das Orchester befindet.« (Pander 2002 [1929], 391)

Der Tonfilm scheint also auch bei Pander einerseits Einschränkungen unterworfen im Hinblick darauf, ob die sichtbaren und hörbaren Klangquellen (z. B. Mund/Lautsprecher) miteinander korrelierbar sind, andererseits aber scheint er auch andersartige Freiheiten mitzubringen, die die varierende Kombination von Bildern und Tönen erlaubt. »Die wirkliche und die scheinbare Tonquelle können bei einer Tonfilmvorführung übereinstimmen« (ebd., 389), so beschreibt es Pander zunächst, aber wie sich in seinen weiteren Ausführungen zeigt, müssen sie es nicht unbedingt.

Diese Freiheit von Bild- und Tonebene generiert aber in Bezug auf die räumliche Dimension noch zwei weitere Probleme, nämlich im Hinblick auf Proportionen und Lautstärken. In Bezug auf die Lautstärke (z. B. von Menschenstimmen) könne im Kino laut Pander »[v]on Natürlichkeit oder Unnatürlichkeit [...] überhaupt nicht mehr die Rede sein« (ebd., 391) und auch durch die varierende Einstellungswahl des Filmbildes sei eine einheitliche Übereinstimmung von Klangvolumen und sichtbarem Setting nicht in der gleichen Varianz möglich. So sei »[g]ewöhnlich [...] das projizierte Bild, soweit es sich um die Darstellung von Menschen handelt, über groß im Vergleich mit den Originalen, und bei Groß- und Teilaufnahmen finden sich ganz ungewöhnlich starke Vergrößerungen«, im Gegensatz zu Landschaftsaufnahmen z. B. – so Pander (ebd.). Der Abstand des Publikums zur Leinwand könne diese Ton-Bild-Diskrepanzen zwar hin und wieder kompensieren, aber generell seien die visuellen Varianzen problematisch für einen naturalistischen Ton, der sich über alle Einstellungen hinweg als solcher erhalten sollte. Die zuvor für den Stummfilm festgestellten perspektivischen Verzerrungen des Menschen, die z. B. Salten (1992b [1913], 364) über die Standardisierung der Leinwand und deren feste Verpflichtung auf das Menschenmaß als Maßstab kompensieren wollte,³¹² wären im Angesicht des Tonfilms der einzige Weg diese scheinbar illusionsgefährdenden Sprünge zwischen Bild und Ton zu vermeiden. Dass es aber genau diese Sprünge und Inkonsistenzen zwischen Bild und Ton sein könnten, die den Kunstwert des Films aus Stummfilm und Rundfunk ins »neue« Medium herüberretten, auch diese Idee entwickelt sich relativ früh.

312 Siehe Kapitel 3.1.2.

3.3.2 | Synchronität, Asynchronität, Kontrapunktik

Anhand der Reflexion der geschilderten technischen Problemstellungen und Herausforderungen des Tonfilms durch seine Kritiker:innen und Macher:innen, schärft sich somit ein Bewusstsein dafür, dass mit der *Entkopplung* von Bild und Ton in ihrer Aufzeichnung sowie ihrer Rekopplung im fertigen Film Freiheiten gegeben sind, die es eben nicht nur im Hinblick auf weitere Fortentwicklungen des Films zu überwinden gilt, sondern die sich – ganz im Gegenteil – gerade künstlerisch nutzen lassen. Es findet daher eine intensive Auseinandersetzung mit den Grenzen und Potenzialen von Synchronität, Asynchronität und Kontrapunktik statt. So schildert z. B. der zunächst tonfilmkritische Clair auch Möglichkeiten, die die ins Audiovisuelle verdoppelten Mittel des Tonfilms produktiv ausnutzen, etwa in Bezug auf die Dialoggestaltung. Ähnlich wie Jerven beschreibt er dabei eine Gesprächssituation, in der auf der Bildebene von den Sprechenden »weggeschnitten« wird, allerdings nicht zum Zweck der Kaschierung von technischen Unzulänglichkeiten, sondern, dem entgegengesetzt, als Stilmittel von Eigenwert.

»Jedenfalls zeigen uns die Sprechfilmdialoge, daß es meist interessanter ist, das Gesicht des Zuhörenden als das des Sprechenden zu sehen [...]. Das ist schon ein beachtlicher Schritt vorwärts und ein Zeichen, daß der Tonfilm aus den Kinderschuhen heraus ist, aus der Zeit, als der Mund des Schauspielers sich im gleichen Augenblick zu öffnen hatte, in dem der Ton erklang, um dem Beschauer das gute Funktionieren des Mechanismus zu zeigen. Die besten Ton und Sprecheffekte werden durch den abwechselnden Gebrauch von Bild und Ton erzielt, nicht durch deren Gleichzeitigkeit. Vielleicht ist diese erste sich aus dem Chaos lösende Regel eines der technischen Axiome von morgen?« (Clair 1952 [1929], 99)

Der ästhetische Mehrwert des Tonfilms ergibt sich hier also eben nicht aus seiner technischen Realisierbarkeit, die das Audioviduum ermöglicht, sondern gerade aus der Abkehr von dieser Gestalt(ungs)option. Das wechselnd unvollständige bzw. zu vervollständigende Audioviduum dient so in produktiver Weise der medienreflexiven Demonstration der Dividualität des Audiovisuellen.

Dass Clair mit der Antizipation dieser Axiomatik nicht ganz unrecht hat, zeigt ein Blick auf das bereits im Jahr zuvor veröffentlichte *Manifest zum Tonfilm* von Sergej M. Eisenstein, Wsewolod I. Pudowkin und Grigorij W. Alexandrow (2001 [1928]), das schließlich zu einem der vielbeachtetsten Texte in Bezug auf die Tonfilm-Etablierung wird.³¹³

313 Wie Müller (2003) zum Tonfilmmanifest ausführt: »Die Bedeutung dieser Theorie wird von der Filmgeschichtsforschung und Filmtheorie unterschätzt, weil man ihren Niederschlag zu sehr allein in den Filmen der russischen Regisseure selber suchte. Das *Tonfilm-Manifest* beeinflusste

Audiovisuelle Kontrapunktik: Das Tonfilmmanifest

Im sogenannten Tonfilmmanifest setzen sich Eisenstein, Pudowkin und Alexandrow (2001 [1928]), ausgehend von ihren Konzeptionen einer formalistischen Montagetheorie, mit den Auswirkungen auseinander, die der Ton im Hinblick auf die visuellen Gestaltungsweisen des Films entfaltet. Während Farbe und Stereoskopik, die potenziell als nächste Schritte in der Transformation der Filmkunst anstünden, die Montage wenig beeinflussten (vgl. ebd., 55), sei der Ton ein entscheidenderer Faktor im Gefüge des Films. Die Autoren starten daher mit der Feststellung:

»Es ist bekannt, daß das elementare (und einzige) Mittel, das dem Film eine solche mächtige Kraft gegeben hat, die MONTAGE ist. Die Anerkennung der Montage als vorrangigem Gestaltungsmittel ist zum unbestreitbaren Axiom geworden, auf dem eine weltweite Filmkultur errichtet wurde.« (Eisenstein/Pudowkin/Alexandrow 2001 [1928], 54; Herv. i. O.)

Das heißt in der Konsequenz, dass der Ton diese Vorherrschaft der Montage nicht stören darf bzw. dieser Logik zu unterwerfen ist. Eisenstein, Pudowkin und Alexandrow befürchten allerdings, dass bei der Tonfilmproduktion »die Klangaufzeichnung naturalistisch durchgeführt wird« und somit »in einer Weise, die genau mit der Bewegung auf der Leinwand korrespondiert und eine gewisse *Illusion* sprechender Menschen oder hörbarer Objekte etc. vermittelt« (ebd., 55; Herv. i. O.). In dieser Synchronität als »naturalistischem« Verfahren sehen die Autoren jedoch einen entscheidenden Widerspruch zu den Logiken der Montage, weshalb sie sich gegen diesen Modus des Originaltons wenden:

»Den Ton in diesem Sinne zu verwenden, würde aber die Zerstörung der Montage-Kultur bedeuten, denn jegliche ÜBEREINSTIMMUNG zwischen dem Ton und einem visuellen Montage-Bestandteil schadet dem Montagestück, indem es dieses von seiner Bedeutung löst. Dies wird sich zweifellos als nachteilig für die Montage erweisen, da es sich in erster Linie nicht auf die Montage-Teile auswirkt, sondern ihre ÜBERLAGERUNG. NUR EINE KONTRAPUNKTISCHE VERWENDUNG des Tons in Beziehung zum visuellen Montage-Bestandteil wird neue Möglichkeiten der Montage-Entwicklung und Montage-Perfektion erlauben. [...] DIE ERSTE EXPERIMENTELLE ARBEIT MIT DEM TON MUSS AUF SEINE DEUTLICHE ASYN-

jedoch zumindest in Deutschland die Erwartungen an die Tonfilmästhetik auf breiter Ebene und bewirkte, dass man sich vom Tonfilm eine ebenso radikal neue Kunstform versprach, wie es der Stummfilm in der Kultur gewesen war. Das *»Tonfilm-Manifest«* hatte in Deutschland jedoch nicht nur auf die allgemeinen Erwartungen an die Tonfilmästhetik Auswirkungen, sondern auch auf die frühe Tonfilmproduktion und machte sich sogar bis in konzeptionell ausgewiesene *»Publikumsfilme«* hinein bemerkbar« (ebd., 18; Herv. i. O.).

CHRONISATION MIT DEN VISUELLEN BILDERN AUSGERICHTET WERDEN. Nur eine solche Operation kann die notwendige Konkretheit herbeiführen, die später zur Schaffung eines ORCHESTRALEN KONTRAPUNKTES visueller und akustischer Bilder führen wird.« (Ebd., 55f.)

Pudowkin schreibt diese Idee einer Kontrapunktik schließlich in seinem Buch *Über die Filmtechnik* (1961 [1928]) weiter fort. Seiner Meinung nach dürften Bild und Ton eben »nicht durch naturalistische Nachahmung aneinander gekettet sein« (ebd., 217), sondern müssten sich unabhängig voneinander entfalten, was dann allerdings auch Konsequenzen für die Art des Tons hätte.

»Freilich verstehe ich unter dem Ausdruck ›Tonfilm‹ keinen Dialogfilm, in welchem die Sprache UND verschiedene Geräuscheffekte genau mit dem entsprechenden Bild auf der Leinwand synchronisiert sind. Solche Filme sind nichts anderes als eine Art photographierte Theaterstücke. Die wirkliche Zukunft gehört einem Tonfilm anderer Art.« (Ebd., 202)

Und er konkretisiert dies in einer Forderung:

»Niemals sollte man jedoch auf der Leinwand einen Menschen zeigen, dessen Rede genau mit der Bewegung seiner Lippen synchronisiert ist. Das ist billiges Nachäffen, ein bloßer Trick, der keinem etwas nützt.« (Ebd., 203)³¹⁴

Sprache und menschliche Stimme sind also nicht kategorisch aus Pudowkins filmischen Gestaltungsmitteln ausgeschlossen, aber wenn sie auftauchen, dann nicht in Dialogen oder anderen Formen lippensynchroner Rede. Stattdessen solle man sich den Umgang mit Ton, Sprache und Bild wie die verschiedenen Stimmen und Melodien eines Orchesters vorstellen, so Pudowkin (vgl. ebd.). In dieser Konzeption eines orchesterlichen Tonfilms, in dem sich Bild und Ton kontrapunktisch ergänzen und dadurch erst künstlerische Potenziale freisetzen, erweist sich Pudowkin als klarer Verfechter des Tonfilms, den er auch als »das größte Ausdrucksmittel von heute und von morgen« (ebd., 238) bezeichnet. In der Gestaltung dieser neuen Form sieht er die »Zukunft« (ebd.) eines produktiv erweiterten Kinos der Audiovision:

³¹⁴ Diese Äußerung ist dabei auch in einer gewollten und expliziten Abgrenzung zum Theater begründet: »Ich glaube, daß nur eine Entwicklung in diese Richtung den Film von der Bühnen-nachahmung befreit und ihn über die Grenzen, die dem Theater gesetzt sind, hinauswachsen läßt, Grenzen der Vorherrschaft des gesprochenen Wortes, der Unveränderlichkeit in der Lage des Bühnenbildes und der Requisiten, der Abhängigkeit der Handlung wie auch der Aufmerksamkeit des Publikums vom Schauspieler, des Zusammenschrumpfens der weiten Welt zu einem einzigen Zimmer ohne vierte Wand« (Pudowkin 1961,[1928], 231f.).

»[D]er Tonreichtum wird überwältigend sein. Alle Geräusche der ganzen Welt, anfangen beim Flüstern eines Menschen, dem Weinen eines Kindes bis zum Donnerschlag einer Explosion! Der Expressionismus des Films kann unerhörte Tiefen erreichen. Er kann den Zornesausbruch eines Menschen mit dem Brüllen eines Löwen in Verbindung bringen. Die Sprache des Films wird die Ausdruckskraft der Sprache der Literatur erhalten.« (Ebd., 202f.)

Was sich in dieser Schilderung Pudowkins aber, ebenso wie schon im Tonfilmmanifest, ankündigt, ist dass in diesen Beschreibungen des neuen Films der Mensch mehr und mehr in den Kontext der weiteren Geräusche eingegliedert wird, was auch daran festzumachen ist, dass in vielen der (schon zuvor zitierten) Texte der Mensch oftmals nur »ein Beispiel unter anderen« ist (wie schon bei Saltons Hunden und Pfeifchen). Auch in den Restriktionen, die Pudowkin und das Tonfilmmanifest für die Verwendung von Klang im Film festsetzen, ist dabei zu bemerken, dass es stets auch darum geht der Sprache keinen vorrangigen Status vor anderen Formen des Geräuschs einzuräumen. Der Anthropozentrismus, der in den früheren Ansätzen zum Film und auch zum Rundfunk noch zu spüren ist, weicht also einer stärkeren Egalisierung seiner Motive und dadurch einer heterogenen Auseinandersetzung mit dem Medium selbst. Gleichzeitig ist in diesen Ausschließungsmechanismen und Argumenten gegen »naturalistische« Originaltöne eine klare Absage an das Audioviduum implizit, die für die hier vorliegende Arbeit natürlich höchst interessant ist: Es scheint, dass die zuvor ausgefochtenen Kämpfe um Film und Funk als eigenständige Kunst- und Kommunikationsformen hier ihren Nachhall finden. Während diese sich zunächst am Audioviduum (als utopischem Maßstab) insofern abarbeiten, als die jeweils fehlende Stimme und der fehlende Körper des medialisierten Menschen zunächst von einem *Mangel* zur einzigartigen *Qualität* erhoben werden müssen, wird nun – nachdem das Audioviduum möglich ist – erneut sein Ausschluss zum zentralen Kriterium auch einer Theorie des Tonfilms. Dem Audioviduum wird seine Unteilbarkeit (Individualität) versagt, sodass es weiterhin (maximal) entweder als nur-sichtbares oder nur-hörbares Verwendung finden soll. Diesen Ideen wird noch weiter zu folgen sein.

Doch es gibt wenigstens einen Bereich, in dem der Mensch auch in Zeiten des Tonfilms im Zentrum bleibt, nämlich der Bereich der Synchronisation für den internationalen Markt. Auch auf dieses Problem sieht Pudowkin den kontrapunktischen Film als Antwort, wenn er über seine Konzeption des Films als Bild-Ton-getrenntes Machwerk schreibt:

»So stelle ich mir einen Film vor, der tönt, und ich möchte darauf hinweisen, daß ein solcher Film im eigentlichen Sinne international wäre. Worte und Geräusche,

die man hört, aber auf der Leinwand nicht entstehen sieht, könnten in jede Sprache übertragen und je nach dem Land für den Film ersetzt werden.« (Ebd.)³¹⁵

Da aber – entgegen Pudowkins Idee – der massentaugliche Tonfilm der Zeit in der Regel (auch) Sprech- und Dialogfilm bleibt, ist auch die Frage der audioviduellen Verschmelzung über Körper-, Sprach- und Landesgrenzen hinweg ein zentrales Thema. Weil diese Debatten folglich exklusiv um das Audioviduum kreisen, sei auch auf sie ein kurzer, exemplifizierender Blick geworfen.

Mit fremden Zungen über falsche Lippen: Synchronfassungen und problematische Internationalität

Das Problem, dass sich dem Tonfilm im Hinblick auf die menschliche Stimme stellt, wird deutlich, wenn man zunächst in einen vor seiner Etablierung angesiedelten Text von Oskar Kalbus hineinliest, der 1920 über den Stummfilm schreibt:

»Der Film steht abseits vom Babelgewirr der Sprachen, er ist das einzige wahrhafte Verständigungsmittel der Nationen, das Esperanto der Seelen. Der Film kann, aus dem Hirn eines wahren Künstlers geboren, Weltkriegswunden heilen und Menschheitsgedanken vermitteln. Das ist der unverkennbare Wert der Stummheit des Films.« (Kalbus 2002 [1920], 137)

Genau dieses visuelle »Esperanto der Seelen« ist aber natürlich darauf angewiesen eine möglichst weltweite Verbreitung zu finden, was wiederum im Hinblick auf den Tonfilm und seine visuelle Offenheit bei sprachräumlicher Geschlossenheit zum Problem wird. Auch Leisegang (2002 [1929]) verweist auf diese Problematik, wenn er bezüglich der Umstellung vom Stummfilm auf den Sprechfilm schreibt:

»Es kann hier nicht unsere Aufgabe sein, den wirtschaftlichen Bedenken dieser Entwicklung im einzelnen nachzugeben. Nur das schwierigste und bisher noch

315 Auch im Tonfilmanifest wird auf diese Idee einer vereinfachten Internationalisierung der Filmkunst selbst unter Vorzeichen des Tons hingewiesen: »Der Ton wird, wenn er als neues Montage-Element verstanden wird (als ein vom visuellen Bild getrennter Faktor), zwangsläufig gewaltige Möglichkeiten zum Ausdruck und zur Lösung der kompliziertesten Aufgaben mit sich bringen, die wir jetzt mit den Mitteln einer mangelhaften Filmmethode, nämlich der ausschließlichen Arbeit mit visuellen Bildern, nicht lösen können. Die KONTRAPUNKTISCHE METHODE für die Konstruktion des Tonfilms wird das INTERNATIONALE KINO nicht schwächen, sondern statt dessen seine Bedeutung zu unvergleichlicher Kraft und kultureller Höhe steigern. Eine derartige Methode würde sich nicht, wie das Abfilmen von Dramen, auf einen nationalen Markt beschränken müssen, sondern wird wie niemals zuvor die Möglichkeit bieten, eine filmisch gestaltete Idee weltweit zu verbreiten« (Eisenstein/Pudowkin/Alexandrow 2001 [1928], 55f.; Herv. i. O.).

gänzlich ungelöste Problem der *Internationalität des Sprechfilms* sei herausgegriffen. Augenkunst hat vor Ohrenkunst den großen Vorteil, nicht an sprachliche Grenzen gebunden, also international zu sein; und gerade die schwer ringende deutsche Filmindustrie, die die Weltmarktsproduktion an zweiter Stelle beherrscht (Amerika versorgt heute den Markt mit 85-90%, nach Amerika kommt Deutschland vor Frankreich mit 4%), muß, um sich behaupten zu können, auf diese Internationalität höchstes Gewicht legen. Sie ist aber außerordentlich gefährdet, sobald jeder Film die Sprache seines Herstellungslandes spricht, die ihn selbstverständlich in den meisten Fällen für anderssprachige Nationen untauglich macht.« (Ebd., 366; Herv. i. O.)

Über die Frage der Sprache halten somit nationale, kulturelle, wirtschaftliche und politische Gesichtspunkte Einzug in die Tonfilmdebatten. Leisegang verbindet diese Ebenen aber auch mit ästhetischen Fragen, indem er die Unterschiede zwischen europäischem und amerikanischem Filmgut in Bezug auf den Einsatz des Dialogs differenziert: Während Erstere den Tonfilm stärker als »Lautfilm« interpretierten in dem eine »Vereinigung von Sprech- und Augenkunst« stattfinde (ebd., 364), konzipierten ›die Amerikaner‹ den Tonfilm eher als musiklastige Filmoperette,

»d.h. sie halten das gesprochene Wort im Kino für eine ›Kinderkrankheit, die nach spätestens einem Jahr überwunden sein wird. Anstelle des ernsteren Wortes glauben sie an die leichtere Musik. Stellt sich Europa also zum Lautfilm, so bekennt sich Amerika zum Ton- oder Musikfilm.« (Ebd., 364f.)

Ähnlich erläutert auch Oskar Kalbus (1935) rückblickend die national spezifischen Unterschiede zwischen einem »Amerika im Tonfilmtaumel« (ebd., 9), in dem die Umstellung auf die neue Technik sich ab 1927 sehr rasant vollzieht, und dem deutschen Verhältnis zum Tonfilm. Bei ihm ist es, im Sinne einer nationalstaatlichen und damit einhergehend kulturellen Differenzierung, dabei erneut das US-amerikanische Kino, das eher Musik und Gesang favorisiert, während er dem deutschen Film eine Affinität zur »Wortkunst« zuschreibt:

»Hollywooder Künstler wollten vom vertonfilmten Theater, vom Bühnenersatz nichts wissen, und so galt damals der reine, hundertprozentige Sprechfilm so gut wie überwunden. Die menschliche Sprache sollte aus dem tönenden Film zwar nicht völlig ausscheiden, sie sollte aber nur als Sprech- oder Dialogpartie dort eingestreut werden, wo sie einen besonderen Sensationswert oder für eine bestimmte Situation einen dramatischen Höhepunkt erreichen kann. So wird man verstehen, daß der Amerikaner lange Zeit die Zukunft des Tonfilms erblickte in der ›Musical Comedy‹, einem Gemisch von Operette, Revue und Spielfilm. Die

Spielhandlung ist hier immer eine süßliche Liebesgeschichte mit sensationellem Einschlag und humoristischen Kontrastwirkungen, und dazu immer und überall Schlager, Tanznummern, Girlparade. [...] Diese ›Musical Comedies‹ bildeten lange Zeit das ganze Arbeitsgebiet von Hollywood und seiner Tonfilmproduktion, auf jeden Fall so lange, bis die Deutschen in Temephol und Neubabelsberg deutsche Filmkunst und deutsche Tonkunst zu einem Ganzen zusammenschweißten und hieraus eine neue Tonfilmdramaturgie schufen, vor der die Amerikaner die Augen groß und ängstlich aufrissen.« (Ebd., 9)

Kalbus proklamiert hier also eine durchaus politisch interpretierbare Konträrstellung zwischen Amerika und Deutschland, indem er nationalkulturelle Unterschiede an den Menschenmotiven ihrer jeweiligen Kinos festmacht (amerikanische Girlparade vs. deutsche Film- und Tonkunst). Doch das ist eben nicht das einzige Problem.

Das nächste ist die Schwierigkeit, diese unterschiedlichen Kinos dennoch zueinander durchlässig werden zu lassen, indem man die in ihnen genutzten Sprachen über Synchronisationsverfahren³¹⁶ übersetzbarmacht. Müller (2003) und Wahl (2005) geben in ihren Auseinandersetzungen mit dem Tonfilm und seiner Etablierung sehr ausführliche Überblicke zu den Debatten, die sich um dieses Problem entwickeln, und schildern jeweils die vehemente, differenzsetzende Abneigung, die der Neu-Kombination von Körpern und Stimmen in der Nachsynchronisation entgegenschlägt. So werde durch diese »jener Sinn verfälscht, mit dem man lebende Ganzheiten sieht und erfaßt und unterscheidet von Mißgeburten, Zwittergebilden, mechanischen Krüppel-Produkten« – so ein Zitat aus der *Vossischen Zeitung* von 1933 (zit. n. Wahl 2005, 116), und auch laut einem Artikel im *Kinematograph* vom 29. Oktober 1930 sei die Synchronisation »nicht glücklich, weil die Stimme kein Zufall ist, sondern zu einer ganz bestimmten Persönlichkeit gehört, jener Persönlichkeit nämlich, die man auf der Leinwand sieht und welche die Zuschauer gleichzeitig zu hören wünschen« (o. A. 1930, zit. n. Müller 2003, 302). In diesen Beschreibungen wird also die Zulässigkeit des Audiovidiums an die normierende Vorstellung eines Menschen geknüpft, der körperlich wie persönlich ›ganz‹, unversehrt und nur dadurch als individuell zu denken ist, während abweichende Varianten als monströse Abnormitäten oder persönlichkeitsgefähr-

³¹⁶ Müller z. B. schildert mehrere Verfahren, die zur Internationalisierung der Filme genutzt wurden: »die Aufführung im Original mit Zwischentiteln in landessprachlicher Übersetzung / die stumme Aufführung mit landessprachlichen Zwischentiteln / die Aufführung im unbearbeiteten Original / die Produktion polyglotter Filme / die Produktion von Mehrsprachenversionen / die Aufführung mit Untertiteln / die Aufführung landessprachlich synchronisierter Filme« (Müller 2003, 292)

dende Eingriffe diskreditiert werden.³¹⁷ Die Individualität – also die Ganzheit und Unteilbarkeit – des Audiovidiums wird somit nicht nur von der Synchronität von Bild und Ton abhängig gemacht, sondern ebenso von ihrer Übereinstimmung mit einer vorausgesetzten außerfilmischen Individualität der aufgezeichneten Körper und Stimmen – wie es z. B. auch in der *Lichtbildbühne* vom 30. Dezember 1930 eingefordert wird:

»Selbst wenn es gelingt, durch geschickten Schnitt und Dialog eine technisch vollkommene, nämlich 100prozentige synchrone Arbeit zu erreichen, selbst dann ist das ganze System widersinnig. Denn die Stimme ist ein Teil der Persönlichkeit des Menschen, und man kann nicht auf Dauer dem Publikum Mary Pickford mit der synchronisierten Stimme von Fräulein Meyer zeigen oder Emil Jannings mit der spanischen Stimme irgendeines Señor.« (H. F. 1930, zit. n. Müller 2003, 302)

Auch Wilhelm Karl Gerst fasst in einem Text zur »künstliche[n] Berechtigung des synchronisierten Films« (1931) in *Die Filmtechnik* diese zeitgenössischen Diskurse zusammen. So erscheine es dem Gros des Publikums und der Kritik als nicht zulässig eine

»im Filmbild dargestellte Rolle durch einen anderen Künstler sprechen zu lassen, als durch den, dessen Bild der Film zeigt. [...] eine Rolle in einem Film könne doch nur von einer künstlerischen Persönlichkeit erfaßt werden [...]. Ein Kritiker faßte seine ablehnende Meinung kürzlich in den Satz: ›Sprache ist Leib. Damit wollte er sagen, daß die Stimme ein Teil der körperlichen Wesenheit des Schauspielers sei und daß dessen Stimme nicht durch die eines anderen Menschen ersetzt werden könnte.« (Gerst 1931, zit. n. Müller 2003, 303)³¹⁸

317 Michael Wedel (1998) spricht im Zusammenhang mit der Synchronisation z. B. auch von einer »schizophrene[n] Technik« und Kurt London schildert sie zeitgenössisch als »gespenstische Technik« (London 1932, zit. n. Müller 2003, 305).

318 Auch Müller selbst nimmt eine Zusammenfassung des Anti-Synchron-Diskurses vor, in dem sie vor allem die Schaffung eines »künstlichen Menschen« als Problem identifiziert: »Denn der Tonfilm erzeugte durch die Synchronisation erstmals einen synthetischen, ›virtuellen‹ Menschen, dessen Körperlichkeit keine Einheit mehr ist. Selbst wenn es sich nur um ein medial erzeugtes ›Mensch-Produkt‹ handelte, so schuf die Synchronisation durch die Verschmelzung der optischen Erscheinung des Körpers eines Schauspielers mit der Wiedergabe der Stimme eines anderen doch einen ›Menschen‹, der nicht existierte. Bei diesem Aspekt überschneiden sich die Bereiche Tonfilmtechnik, Ästhetik und Kultur des Fiktionalen, weil man sich zum ersten Mal damit konfrontiert sah, dass die gültigen Konventionen fiktionaler Darstellungen durch eine Technik ästhetisch gebrochen wurden. Die Abwehr der Zeitgenossen gegen die Synchronisation war entsprechend auch nicht ethisch, sondern ästhetisch motiviert« (Müller 2003, 20). Dies ist natürlich insofern interessant, als Vertov, Tynjanov, Pudowkin und andere im »künst-

Das menschliche Individuum scheint sich also insofern auf das mediale Audioviduum auszuwirken, als die audiovisuelle Spaltung überwunden werden muss zugunsten der Repräsentation »ganzheitlicher Menschen« – verstanden als solche mit Stimmen und (genormten) Körpern. Doch Gerst selbst ist anderer Meinung. Ähnlich wie Clair und die russischen Formalisten³¹⁹, sieht er in der Synchronisation ein eigentlich produktiv nutzbares, künstlerisches Ausdrucksmittel, in einem Medium (Film), das ohnehin den Menschen fragmentiere und in wechselnden Größen und Anschnitten zeige. Es sei daher »unwesentlich, ob die Sprache, die wir im Film hören, von demselben Menschen stammt, dessen Körper wir im Film sehen, oder ob es der Technik möglich geworden ist, zu diesem Bilde die Stimme eines anderen Menschen einzufügen«, so Gerst (ebd.). Dennoch ist dies laut Müller (2003, 304) eine Ansicht, die sich (zumindest zunächst) nicht wirklich durchzusetzen vermag. Den Filmschaffenden bleibt also nur übrig, Modalitäten zu finden, mit denen diesen Problemen körperlicher und stimmlicher Ganzheit zu begegnen ist:

»Die dem Sprechdoppel gestellte Aufgabe ist, sich in die Rolle einzuleben, die ein anderer vorgespielt hat und sich trotzdem eine gewisse Individualität zu wahren; denn ein Kopieren schlechthin kann nie zu einer befriedigenden Darstellung führen.« (Fch. 1932, zit. n. Müller 2003, 308f.)

– so beschreibt ein gleichnamiger Artikel in *Der Kinematograph* »Das Versionen-Problem« (1932). Bei der Wahl der jeweiligen Stimmenbesetzung müsse deshalb diese sonst so stabile Verbindung des Menschen mit seiner Stimme Beachtung finden:

»Die Auswahl der Versionensprechers [...] wird dadurch noch erschwert, dass der Zuschauer, wenn er das Bild eines Menschen sieht, sich eine gewisse Vorstellung macht, wie dieser Mensch voraussichtlich sprechen wird; es besteht eine Art von Kongruenz zwischen stimmlicher und bildhafter Persönlichkeit, sind wir hierin getäuscht, so wirkt das Ganze unwahr und unbefriedigend.« (Ebd., zit. n. Müller 2003, 309)

Es bleibt also festzuhalten, dass – trotz der parallel z. B. von den russischen Formalisten unternommenen Bemühungen um Kontrapunktik und Asynchronität – das Problem der Sprache im Tonfilm im Hinblick auf seine Nachsynchronisation

lichen Menschen« – parallel zu den Tonfilmdebatten am Übergang zu den 1930er Jahren – den eigentlichen Sinn und Zweck des Films sehen. Siehe auch Kapitel 3.1.4.

319 Wobei diese natürlich Lippensynchronität generell ablehnen.

und Internationalität bestehen bleibt.³²⁰ Das Audioviduum steht also auch hier im Zentrum der Verhandlung des Tonfilms als ‚neuem‘ Medium. Was sich in diese Debatten um eine Art unumgängliche und einmalige Verbindung vom Menschen und seiner Stimme zudem einschreibt, ist ein weiterer Aspekt, der im Hinblick auf den Tonfilm relevant wird, nämlich seine dokumentarischen Qualitäten bzw. sein ‚Naturalismus‘.

Der Naturalismus des sprechenden Menschen

Wie schon in Bezug auf Pander und auch Leisegang angedeutet, spielt die Frage der Realitätsnähe des Films innerhalb der Tonfilmdebatte eine ebenfalls zentrale Rolle.³²¹ So hatte Leisegang bereits die unterschiedliche Realitätsnähe von 3D-Ton und 2D-Bild als problematisch angemerkt, die quasi nur über Stereoskopik und Farbe zu überbrücken sei (vgl. Leisegang 2002 [1929], 367). Und auch Pander (2002 [1929]), der sich vor allem mit der Illusionsfähigkeit des neuen Mediums befasst (etwa in Bezug auf die Platzierung von Lautsprechern; s. o.), sieht mit dem Tonfilm einen weiteren Schritt zur vollständig realistischen Kunst getan:

320 So schreibt René Clair etwa noch 1952 zur Frage der Synchronfassung: »Dies Problem harrt noch immer seiner Lösung, und der Turmbau zu Babel ist nach wie vor das Symbol des Tonfilms. Die Beschriftung ausländischer Filme, die den Besucher zum Lesen zwingt, statt daß er sieht und hört, ist ein Unding, und von der Synchronisation sagt Jean Renoir mit Recht, ihr Erfinder wäre im Mittelalter auf den Scheiterhaufen gekommen. Einem Körper eine fremde Stimme zu geben sei so schlimm wie Hexerei. Es nimmt Wunder, daß sich die Schauspieler, die doch sonst, was ihren Ruhm und das Prestige ihres Standes betrifft, so empfindlich sind, diese degradierende und standeswidrige Praktik gefallen lassen« (Clair 1952, 96f.).

321 Erneut sei hier auf Müller (2003) verwiesen, deren Buch *Vom Stummfilm zum Tonfilm* ja die zentrale These aufstellt, dass sich die frühen Tonfilmtechniken deshalb zunächst nicht etablieren, weil sie als Tatsachenmedium wahrgenommen wurden und sie dadurch für die Frage der Kunst nicht relevant bzw. eher hinderlich erschienen. Erst mit der Durchsetzung der Idee, dass der Film ein Medium des Fiktionalen sei (und eben nicht des Dokumentarismus), habe sich der Film als Kunst etablieren können. Die Verfremdung durch den fehlenden Ton sei dabei ein wichtiger Grund für diese mögliche Abgrenzung (vgl. ebd., 14). Müller fasst dies so zusammen: »Diese Ausrichtung auf einen materialtechnischen ‚Realismus‘ filmischer Darstellungen war unvereinbar mit dem Anspruch, der Kultur des Fiktionalen anzugehören, die verlangt, dass die Differenz fiktionaler Äußerungen und Darstellungen von der empirischen Wirklichkeit und auf diese bezogene Äußerungen und Darstellungen deutlich erkennbar signalisiert werden muss. Sein Verstummen, so die hier zu diskutierende These, war die Antwort des Films auf diese Prämisse, um sich in der kommerziell ertragreichen Kultur des Fiktionalen, in Kunst und Unterhaltung durchzusetzen«(ebd., 15). Mit dem Tonfilm aber gerät genau diese Argumentation in eine Krise: »Wenn aber ein Medium, das sich prononciert als ‚nicht-realistisches‘ Medium ausgewiesen und etabliert hat, plötzlich ‚realistisch‘ wird, dann ist das weniger eine ‚natürliche Entwicklung‘ der filmischen Gestaltungsmittel, als vielmehr ein Problem« (ebd., 16).

»Bekanntlich schreitet die Technik unaufhaltsam fort; besonders rasch die des Films; zum Sichtbaren kommt jetzt, illusionssteigernd, der Schall, das Hörbare hinzu, es fehlen nur noch Farbe und Plastik, dann wird der vollkommene Film nicht mehr von der Wirklichkeit unterschieden werden können. [...] Es wird also Zeit, daß man sich mit der Frage der Illusion beschäftigt. Für den Tonfilm hat sie zwei Unterabteilungen: die nach der Illusion beim Sichtbaren und beim Hörbaren.« (Ebd., 388)

Wie bereits in Bezug auf das Tonfilmmanifest festgestellt,³²² wurde diese Illusionskraft dabei vorrangig als in der parallelen Aufzeichnung und dann auch Wiedergabe von Originalton und tönendem Objekt begründet angesehen; einen Bild-Ton-Modus, den z. B. Eisenstein, Pudowkin und Alexandrow auch als »naturalistisch« bezeichnen (und ablehnen). Nach Pudowkin (1961 [1928]) jedenfalls gehe es nicht um eine gesteigerte Naturnähe des Films, nicht darum die »Lebensähnlichkeit des Bildes zu verbessern« (ebd.), sondern in der Kontrapunktik müssten Bild und Ton gerade (ihren realweltlichen Entstehungskontexten gegenüber) separiert Verwendung finden.

»Im Stummfilm wurde der Sprechende von Titeln begleitet, heute hören wir seine Stimme. Die Rolle, die der Ton im Film zu spielen hat, ist von weit größerer Tragweite, besteht in viel mehr als in sklavisch naturalistischer Nachahmung.« (Ebd., 216)

Die Wiedergabe von ganzheitlich aufgezeichneten und als solche (re)produzierten Menschen hält Pudowkin folglich für den banalsten Effekt des Tonfilms, der daher im Sinne künstlerischen Schaffens gerade abzulehnen sei.

Julius Bab (2002 [1925]) sieht dies im Grunde ähnlich, betont aber auch den Nutzen, den der Film als Reproduktionstechnik von Kunst haben könnte (er schreibt 1925 zu einer Zeit, in der diese Technik sich erst noch etablieren wird). So fragt er sich zunächst, warum noch nicht längst ein Filmarchiv für Tanzveranstaltungen angelegt sei, um diese erstmals für die Nachwelt festhalten zu können (ebd., 163). Und auch als Archivierungswerkzeug für das Theater sei speziell der Tonfilm nun (bald) geeignet, wenn er auch darin natürlich keine eigene Kunst werden könne (ebd.). Vor allem am menschlichen Körper sieht Bab dabei diese theatralische Filmkunst in ihren Realitätseffekten scheitern:

»Und der *sprechende* Film wird, wie gesagt, ein sehr wichtiges neues Mittel sein, schauspielerische Reproduktionen zu schaffen; aber, wo er sich selbstständig gebärdet, wird er nicht weniger den Charakter einer Minuskunst haben. Denn er wird bei Klang und Bildeindruck doch noch ein Wesentlichstes vermissen lassen, auch

322 Siehe Abschnitt 3.3.2.

wenn wir uns das Bild farbig, die Klangwiedergabe höchst vollkommen denken: der Eindruck eines lebendigen Körpers geht nicht nur durch Aug' und Ohr. Schon der Blick des Auges, der für viele Schauspieler eine so wichtige Wirkung ist, bedeutet mehr als eine optisch faßbare Erscheinung. Der *lebendige* Körper wirkt über die Rampe hinweg mit vielerlei unsichtbaren und unhörbaren Zeichen, die sich jeder technischen Wiedergabe entziehen werden. Schon vom Schauspieler aus gesehen ist also die Nachahmung des dramatischen Theaters durch den Film hoffnungslos.« (Ebd., 166; Herv. i. O.)

Bab geht also davon aus, dass die eigene Kunst des Films in allen seinen Formen daher nicht in der Reproduktion anderer Künste (des Theaters im Tonfilm, der Pantomime im Stummfilm etc.) bestehe, sondern stattdessen in der durch die Bewegung im lebendigen Bild möglichen, besonderen, künstlerischen Behandlung von Raum und Zeit. Diese Möglichkeiten sieht er daher vor allem im Trickfilm realisierbar (vgl. ebd., 169f.), der ja gerade die vermeintliche Realitätsnähe des Films konterkariert. Alles in allem ist somit auch für Bab die Reproduktionsfähigkeit und dokumentarische Qualität, die der Tonfilm ermöglicht, seinem Status als Kunst eher abträglich.

»Sobald er seine Kräfte auf das wirklich schöpferische Zentrum seiner Sphäre versammelt hat, – d.h. auf den Apparat, für den Naturaufnahmen aller Art, einschließlich der schauspielerischen Pantomime, nur *Rohstoff* sind, mit denen die Herrichtung des eigentlichen Filmwunders *beginnt!* – sobald der Film diese Einstellung erreicht hat, wird er ein neues und echtes Kunstwerk schaffen.« (Ebd., 175; Herv. i. O.)

Ähnlich wie bei Vertov, Pudowkin und anderen wird also auch hier die Kunst des Films nicht mehr vom Menschen direkt abhängig gemacht, sondern vom Apparat und der Möglichkeit alles – ob Mensch, Natur oder Ding – zum Rohstoff seiner Kunst zu deklarieren.

Die Kritik an der vermeintlichen Realitätsnähe des Films äußert sich aber auch noch in anderen Hinsichten. Eberlein (2002 [1929]) etwa sieht in dem (auflängere Sicht sogar mit Farbe und Dreidimensionalität) ausgestatteten Tonfilm kein poetisches »Gesamtkunstwerk [...] von Form, Farbe und Ton« nahen (bzw. »Raumfilm, Farbfilm, Tonfilm«), sondern ganz im Gegenteil eine »Unkunst« deren »Illusion und Schein, durch Natura Nähe und Naturwirkung nur noch ärmer« würde (ebd., 286). »[E]in großer Aufwand wird umsonst vertan« – so wertet Eberlein die technischen Bemühungen; der Naturalismus sei lediglich eine »technische Scheinform« (ebd., 287). Ähnlich wie Eberlein und auch Bab sieht es Clair (1952 [1929]), der im Tonzuwachs des Films (vor allem im Vergleich zum Stummfilm) keinen künstlerischen Mehrwert für das Medium feststellen kann, weil es sich – und das passt

wiederum zu Babs Erwähnung des Trickfilms als genuin künstlerischem Modus des Films – mit der Wirklichkeit zu sehr von seiner eigentlichen Domäne, der des Traumes, verabschiedet habe. Zwar sei der Film auf dem Weg, sich vom schlicht aufgezeichneten Theater zu lösen, aber dennoch sei die »Wirklichkeit« nicht der ideale Modus wenn man Kunst sein und schaffen wolle.

»Die Stummfilmparteigänger werden immer etwas unsicher, wenn sie sich zum Ton- und Sprechfilm äußern sollen, denn dieser ist in seiner besten Form doch schon weit mehr als fotografiertes Theater. Er ist etwas Selbständiges geworden und scheint durch die klangliche Diversität und das Stimmorchester zunächst reicher als der Stummfilm. Aber ist diese Fülle und Pracht nicht Talmi? Ist der Tonfilm nicht ein Koloß mit tönernen Füßen? Hat die Leinwand durch ihn nicht mehr verloren als gewonnen? Zwar hat sie sich das Reich der Stimmen erobert, aber dafür das des Traumes eingebüßt. Ich beobachtete mehrmals eine aus dem Tonfilm kommende Menschenmasse. Sie wirkte, als käme sie aus einer Music-Hall. Kein Zeichen jener wohltätigen Verzauberung. Die Leute lachten, unterhielten sich oder trällerten den letzten Refrain. Sie waren in der Wirklichkeit geblieben.« (Ebd., 101)

Clair attestiert dem Tonfilm also durchaus Potenziale, die ihn von einer reinen Theaterreproduktion abheben könnten. Dennoch scheinen die Stimmen – als Realitätsanker – dieträumerischen und illusionären Freiheiten des Kinos zu tilgen, was sich nicht zuletzt im Publikum niederschlägt. Und auch in einer weiteren, gesellschaftlichen Hinsicht machen sich am vermeintlichen Naturalismus des Tonfilms (im Vergleich zu Stummfilm oder Theater) Zweifel breit; etwa bei Joseph Roth (1956 [1934]), der die unterstellte Wirklichkeitsnähe des Films, vor allem im Hinblick auf dessen Menschenmotive, als Trugschluss entlarvt – was er in diesem Falle weniger der filmischen Technik selbst, als den menschlichen »Vorlagen« zuschreibt, deren Realität bzw. Irrealität im Film zum Vorschein komme:

»Und wenn wir es zustande gebracht haben, daß sich Schatten auf der Leinwand der Kinotheater wie lebendige Menschen bewegen und sogar sprechen und singen, so sind doch keineswegs ihre Bewegungen, ihre Worte und ihr Gesang echt und ehrlich; vielmehr bedeuten diese Wunder der Leinwand das eine: daß die Wirklichkeit, die sie so täuschend nachahmen, deshalb gar nicht schwierig nachzuahmen war, weil sie keine ist. Ja, die wirklichen Menschen, die lebendigen, waren bereits so schattenhaft geworden, daß die Schatten der Leinwand wirklich erscheinen mußten.« (Roth 1956 [1934], 701f., zit. n. Wahl 2005, 34)

Roth integriert also insofern die Illusion wieder in den Film, als die Welt selbst sich in ihm als unwirklich offenbart. So erscheinen die Audioviduen der Leinwand als nur noch schattenhafte Spiegelbilder ebenso schattenhafter Menschen, wodurch

ein Naturalismus auf beiden Seiten des Spiegels nicht mehr behauptet werden kann und sich der Zweck des ganzen in nichts auflöst. Ob nun also die Wirklichkeit oder der Film nicht (mehr) taugen, um filmische Realitäten herzustellen – es bleibt dabei, dass auch hier das Audioviduum nicht als Lösung eines Problems, sondern als (Teil des) Problem(s) selbst erscheint. Was die Frage aufwirft, ob der Tonfilm abseits seiner Menschenmotive weniger problematisch wäre.

3.3.3 | Jenseits von Anthropomorpho- und Anthropophonozentrismus

Damit kommen wir im letzten Teil dieses Kapitels zu zwei finalen Aspekten, die innerhalb der bisher skizzierten Tonfilmdebatten schon angeklungen sind, die sich aber erneut in den Schriften zweier Autoren auf engerem diskursiven Raum verdichten. Dabei handelt es sich um Rudolf Arnheims Schriften zum Tonfilm, vorrangig *Film als Kunst* (2002 [1932] und *Neuer Laokoon* (2004 [1938])), sowie Bruno Rehlingers Abhandlung zum »Begriff Filmisch« (1938). Beide Ansätze verbindet dabei der Versuch der dominanten Konzentration auf das Audioviduum, die die Debatten und Bestimmungen des Tonfilms kennzeichnen, zu entgehen. Die Medientheorie des Tonfilms bewegt sich daher in einen Bereich jenseits von Anthropomorpho- und Anthropophonozentrismus. Das Audioviduum wird somit als diskursives und materielles Motiv noch einmal intensiv verhandelt, um dann vor ihm zu kapitulieren (Arnheim) oder es hinter sich zu lassen (Rehlinger).

Sprechfilm vs. Tonfilm

Zunächst soll also noch einmal Rudolf Arnheim mit seinen Schriften zum Tonfilm zu Wort kommen, der sich ab Ende der 1920er Jahre ebenfalls intensiv und kritisch mit dieser »neuen« Art von Film und seinen Gestaltungsmöglichkeiten auseinandersetzt. Dies geschieht immer wieder in kleineren Artikeln, aber auch in seinem Buch *Film als Kunst* (Arnheim 2002 [1932]), deren Argumentationslinien im Hinblick auf den Tonfilm hier nachgezeichnet werden sollen – bis zu seinem Text *Neuer Laokoon* (Arnheim 2004 [1938]), in dem Arnheim seine Ideen zum Tonfilm Ende der 1930er Jahre erneut reflektiert.

So schildert Arnheim seine ersten Eindrücke zum Tonfilm zunächst in einem kleinen Text zum »tönende[n] Film« (2004 [1928]) in der *Weltbühne*, in dem er trotz kritischem Blick die Faszination beschreibt, die vor allem das Kunststück des nun sprechenden Menschen auf der Leinwand in ihm auslöst:

»Der Tonfilm ist aber da, wenn der Schauspieler Paul Grätz riesengroß auf der Leinwand erscheint, ein bißchen verlegen, denn er ist wohl nicht gewöhnt sein Gesicht mit nichts als etwas Zivilmimik bekleidet so groß vor den Leuten zu zeigen, und wenn er nun zu uns redet. Das S klingt noch, als habe Paule sein Gebiß einzuhalten vergessen, und noch schütten sich verwischende Echos über die Stimme, so als

ständen die vier Wände einer leeren Kathedrale drumherum, aber sonst ist schon alles sehr schön: wenn der photographierte Paule mit einer Selbstverständlichkeit, die einen bei so einem zweidimensionalen Gebilde vorläufig noch wundert, die Lippen zum Sprechen öffnet, gibt es keine Blamage, sondern die Stimme sitzt ihm am rechten Fleck und gehorcht dem leisen Zucken seiner Mundwinkel. Da steht er nun und redet, und vor diesem Wunder wird man naiv wie ein Kind: der Geiger Weißgerber tritt, nachdem er vor Auge und Ohr ein Virtuosenstück aufgeführt hat, mit einer Verbeugung zurück, und da klatscht man ihm spontan Beifall und meint zu sehen, wie er sich darüber freut, und schiebt dann ängstlich und beschämmt zum Nachbarn, aber der hat auch geklatscht. Der Eindruck, daß da nicht ein Abbild, sondern ein leibhaftiges Wesen agiert, ist völlig zwingend. Aber in dem Augenblick wo dies geschieht, tritt die Filmkunst ihren mühselig eroberten Platz wieder an den guten, alten Guckkasten ab« (ebd., 67f.).

Vor allem das Audioviduum ist es also, in dem einerseits die Qualität des ›neuen‹ Mediums eindrücklichste Gestalt gewinnt, in dem sich andererseits aber auch diese Errungenschaft des neuen Mediums – dass es sprechende Menschen zeigen kann – als zentrales Problem darstellt, weil dadurch eine Nähe zum Theater wiederhergestellt wird, die der Stummfilm gerade erst glücklich hinter sich gelassen hatte. Der ganz eigene Kunstwert des Tonfilms bleibt für Arnheim daher zunächst fraglich: »Der Tonfilm als verbessertes Opernglas und als Konservenbüchse – herrlich! Der Tonfilm als eine eigene Kunstform ...?« (Ebd., 70) – so endet sein Text und verweist damit auf die zentrale Frage, die auch seine weiteren Schriften zum Tonfilm prägt. Deutlich wird aber ebenso, dass Arnheim als eigentlicher Verfechter des Stummfilms schreibt, der den Ton mehr als unerwünschtes Anhängsel denn als Zugewinn für den Film ansieht. Dies zeigt sich schließlich auch in seiner Monografie *Film als Kunst*, in der er zunächst einen systematischen Überblick über die grundsätzlichen Gestaltungsmittel des Films auf Bildebene liefert, bis er dann (erst) im vorletzten Kapitel ausführlicher auf den Tonfilm zu sprechen kommt, dessen Versäumnisse und weiterhin bestehenden Unzulänglichkeiten er zum Einstieg mit kritischem Blick bewertet. Die im Kontext dieser Arbeit zuvor skizzierten Tonfilmdebatten (etwa in Bezug auf mangelnde technische und ästhetische Qualität, Banalität und Probleme der Synchronisation etc.) erscheinen darin noch einmal zusammengefasst:

»Denn wenn man erschreckend grunzende Frühgeburten auf das Publikum losließ, Schnellprodukte einer von der Konkurrenz gehetzten Industrie, wenn ein Kuß wie ein Donnerwetter klang und eine Frauenstimme wie eine Dampfsirene, wenn minderwertig aufgenommene Filme auf minderwertigen Projektoren vorgeführt wurden, wenn die Dialoge der Schauspieler das letzte Schmierenstück an Albernheit übertrafen, wenn der Ton mit dem Bild nicht zusammenstimmte,

weil der Synchronismus versagte, wenn man stumme Filme mit eilends dazu aufgenommener Begleitmusik als »Tonfilm« propagierte, wenn man den Schauspielern ausländischer Filme die Worte, die sie gesprochen hatten, wegschnitt und ihnen dafür eine ganz fremde Stimme und Worte einer ganz andern Sprache in den Mund legte oder aber die unverständlichen ausländischen Reden beibehielt und dafür gespenstische Spruchbänder in der eignen Muttersprache über die Bilder huschen ließ, wenn man die Originalfassung und Bearbeitung szenenweise durcheinander schnitt oder alle Dialogstellen einfach herausnahm und sie durch Zwischentitel ersetzte, und wenn man sich mit noch wilderer Entschlossenheit, als das beim stummen Film schon der Fall gewesen war, auf die erprobtesten und abgedroschensten Operetten- und Rührstückthemen warf – ja, dann konnte man sich nicht wundern, daß die Besucher protestierten und fernblieben und über den Tonfilm zu dem Urteil kamen: So jung und schon so verdorben!« (Arnheim 2002 [1932], 191)

Auch in dieser Bestandsaufnahme wird also deutlich, dass Arnheim dem Tonfilm grundsätzlich kritisch-prüfend gegenübersteht und dass er vorrangig in der Inszenierung der in ihm auftretenden Menschenmotive das größte Problem für eine tönende Filmkunst sieht. Doch das scheint die Tatsachen nicht zu ändern: »Der Tonfilm ist da, und es hat wenig Zweck, darüber zu streiten, ob das zu begrüßen oder zu bedauern sei« (ebd., 192) – so fasst Arnheim die Entwicklung zusammen, auch wenn deutlich wird, dass bei ihm das Bedauern überwiegt. So sieht er im Tonfilm eine Gefahr für die Kunst des Stummfilms, der seiner Meinung nach eben »nicht aufgebraucht« (ebd., 192) gewesen sei und es daher verdiene, dass man ihn weiter pflege.

Um aber dennoch dem Tonfilm als unvermeidbarer Tatsache analytisch zu begegnen, macht er zunächst, wie es innerhalb des Diskurses gängig ist, Vergleiche zu anderen mit Ton und Sprache ausgestatteten Künsten auf – namentlich: zum Theater, zur Literatur und zum Hörspiel. Für seine Vorstellung von Kunst ist es dabei entscheidend, dass es ihm nicht um Rand- und Überschneidungsbereiche dieser Einzelkünste geht, sondern ganz konkret um ihren jeweiligen zentralen Kern, um den das künstlerische Vermögen kreist. »*Nein, nicht auf die Grenzen eines Kunstbezirks kommt es an, sondern auf sein Zentrum*« (ebd., 196; Herv. i. O.) – so lautet Arnheims Credo, das er zunächst vor allem in Bezug auf das Theater entfaltet.

Entlang des Vergleichs mit der Bühne kommt Arnheim – ähnlich wie Bab (2002 [1925]) – zu dem Schluss, dass höchstens in der filmischen Aufzeichnung von Theateraufführungen ein Mehrwert des Tonfilms für das Theater liegen könnte; umgekehrt sei der Tonfilm als Kopie des Theaters jedoch für sich wertlos (vgl. Arnheim 2002 [1932], 194ff.). Und auch im Vergleich mit Sprache als genereller, literarischer Kunstform offenbaren sich Probleme: Hier sieht Arnheim die Gefahr, dass bei einer gleichzeitigen, künstlerischen Verwendung von Bild und Wort im

Tonfilm beide »einander nicht ergänzen und vereinigen, sondern nur unerträglich stören« (ebd., 199). Jedoch sei die Sprache eben »außer einem *Kunst-Mittel* auch zugleich *ein Stück Natur*« (ebd.), weshalb der Film gerade durch seine Wirklichkeitsnähe die Möglichkeit habe, die Stimmen (im Gegensatz zu Bühne und Literatur) stärker als Teil der natürlichen Welt zu inszenieren – z. B. durch die Verwendung von Alltagssprache oder auch partielle Unverständlichkeit. Für den Tonfilm sei die Stimme somit »als bloßes Geräusch unter Geräuschen« aufzufassen und folglich hafte dem Film der »Reiz einer unbühnenhaften, intimen Sprechkunst« an (ebd., 200). Auch bei Arnheim erfolgt somit eine Eingliederung des Menschen in die Welt der Dinge/Geräusche, die sich schon in den zuvor skizzierten Debatten zum (Ton)Film – vor allem bei Pudowkin und den russischen Formalisten – andeutete und die sich auch im Folgenden (bei Arnheim, Balázs und auch bei Rehlinger) weiter entfalten wird.

In einem nächsten Schritt führt Arnheim die Idee der menschlichen Stimme in ihrer Materialität als Geräusch schließlich noch im Hinblick auf das Hörspiel weiter aus. Auch in ihm sei es möglich, den Menschen gleichwertig einzugliedern in die Welt der Klänge; ebenso wie der Stummfilm sich »der Gesamtheit des optisch Wahrnehmbaren« bediene, könne auch das Hörspiel (und ebenso der Tonfilm) »die gesamte Geräuschwelt der Wirklichkeit« einbeziehen (ebd., 202). Zudem sei in beiden Formen die menschliche Stimme in ihrer Lautstärke erstmals ihrem Realklang am nächsten:

»Zum erstenmal (vom Grammophon abgesehen) ist es hier möglich die Stimme des Schauspielers in natürlicher Lautstärke zu verwenden während er im Theater bisher immer vor der zwiespältigen Aufgabe stand, brüllen zu müssen, auch wenn er flüsterte.« (Ebd., 208)

Dennoch sieht Arnheim in der Sprachverwendung nicht den zentralen Kern des Tonfilms. Es scheint also, als habe die Stimme zwar einen Gewinn von ihrer Einbettung in Hörfunk und Tonfilm als neuartigen Sprechkünsten, gleichzeitig aber – so wurde zuvor deutlich – sind diese selbst, als Kunstformen, eben nicht (mehr) zentral auf die Stimme und damit den Menschen angewiesen, um Kunst hervorbringen zu können.

Aufbauend auf diesen Überlegungen zu Theater, Sprache und Hörspiel kommt Arnheim schließlich zur Frage des Tonfilms als »Sprechfilm« (ebd., 213) und somit zur Rolle, die das Wort in ihm spielen sollte. Wieder geht es Arnheim dabei um eine Verhältnisbestimmung – einerseits der Art und Weise wie sich die Sprache (als eigenes Kunstmittel, aber auch als natürlicher Teil der Welt) in das übergeordnete Kunstwerk des Films einfügen lässt und andererseits, wie sich dann die

Kunst des Visuellen dazu verhält. Das Ergebnis ist eine zentrale Unterscheidung in Sprechfilm, Tonfilm und Stummfilm³²³ als graduelle Abstufung:

»Obwohl das Wort imstande ist, ganz allein geschlossene Kunstwerke mit vollkommenem Weltbild zu schaffen, ist es dennoch möglich, das Wort mit Sehbildern und Geräuschdarstellungen zusammen in ein einheitliches Kunstwerk einzufügen, weil die Sprache nicht nur ein Kunstmittel sondern ein Stück Welt ist und als solches in die übrige Welt eingeordnet werden kann. Welcher Art nun kann und soll die Funktion des Wortes im Tonfilm sein? Die eine Grenze verkörpert der extreme Sprechfilm, die andere der stumme Film: das Wort übernimmt entweder die gesamte Gestaltung, für die Bild und Geräusch nur, nach Art des Theaters, eine untermalende Begleitung abgibt, oder aber Wort und Geräusch verschwinden ganz und überlassen dem Bild die Gesamtgestaltung. Zwischen beidem liegt der eigentliche Tonfilm, in dem Bild, Wort, Geräusch gemeinsam und gleichberechtigt formend wirken. Und wie schon beim Vergleich zwischen Tonfilm und Theater gesagt worden ist, wieder kommt es nicht darauf an, diese Gebiete scharf gegeneinander abzugrenzen – denn diese Grenzen sind fließend – sondern die Zentren, die Schwerpunkte dieser drei Gebiete zu bezeichnen, zu denen hin sich der Künstler orientieren muß.« (Ebd., 213)

Im Tonfilm erscheint das Wort (und damit der Mensch) somit auf eine Ebene gestellt mit den Geräuschen und Dingen, während der Sprechfilm das quasi an-

323 Eine solche Unterscheidung von Sprechfilm und Ton- bzw. Lautfilm hatte er bereits in einem Aufsatz von 1929 unternommen, wobei es darin noch stärker um die Hierarchisierung von Bild und Ton im Filmgefüge ging – also ob Bild- oder Ton-Ebene die ästhetisch bestimmende ist. Zu dieser Zeit scheint er allerdings den Tonfilm an sich abzulehnen (aus Angst vor gestalterischen Einbußen in Bezug auf Bildgestaltung und Montage): »Man muß bei der Betrachtung der Lage sorgfältig den Sprechfilm vom eigentlichen Tonfilm unterscheiden. Diese Bezeichnungen sind so ungeschickt wie möglich gewählt. Tonfilm – das ist zunächst die mechanische Wiedergabe der Musikbegleitung, die bisher von den Kinoorchestern ausgeführt wurde und die eine nützliche und notwendige Einrichtung ist. Es steht ganz außer Zweifel, daß in allerkürzester Zeit die Kinoorchester verschwinden und jeder Film mit akustischer Begleitung geliefert werden wird. Soweit gibt es keine Diskussion. Sowie aber die Begleitung zur Nachahmung oder Reproduktion von Geräuschen, von Klingeln, Schüssen, Instrumentenspiel übergeht, haben wir die Grenze zum Sprechfilm, der darin besteht, daß alle akustischen Begleiterscheinungen der optisch dargestellten Szene mitgeliefert werden, und gegen den [...] die stärksten künstlerischen Bedenken bestehen. Der ›illustrierte Film‹, wie wir ihn im Gegensatz zum ›Lautfilm‹ nennen wollen, wird seine Untermalungstechnik vervollkommen, alle Raffinements der Geräuschmusik werden eingespannt werden; aber niemals wird das Akustische hier etwas anderes als Begleitung des Optischen sein. Während beim ›Lautfilm‹, wenn er überhaupt einen Sinn haben soll, grundsätzlich Gehörtes und Gesehenes koordiniert sein müssen« (Arnheim 2004 [1929], 73).

thropozentrische Extrem dieser Verhältnisse im Gestaltungsspektrum des Films bildet.

Den Modus des Sprechfilms sieht Arnheim dabei aus verschiedenen Gründen nicht als die erstrebenswerte Variante des Tonfilms an, denn erstens sei dieser technisch (noch) nicht darauf ausgelegt, dass man alles immer gut verstehe und daher sei der reine Sprechfilm in der Rezeption sehr anstrengend (vgl. ebd., 214ff.); zudem zwinge die bei der Aufzeichnung notwendige Nähe des Mikrofons auch die visuelle Ebene des Films in einen Modus monotoner Großaufnahmen, sodass der Variantenreichtum des Bildes unter den Einschränkungen der Tonproduktion leide (vgl. ebd.),³²⁴ natürlich auch, weil sich das Visuelle entgegen dem Sprachlich-Akustischen nicht in den Vordergrund drängen dürfe:

»Die Mimik des Schauspielergesichts wirkt als eine bloße zwanglose Untermalung des Dialogs, fängt man aber an mit auffälligen perspektivischen Einstellungen oder gar mit Montage zu arbeiten, so ist das ein offensichtlich neues, eignes Gestaltungsprinzip, das sich mit der Dominanz des Wortes nicht verträgt. So etwas ist im reinen Sprechfilm ebenso stillos und unmöglich, wie wenn ein ernsthaftes Orchester seine Musik durch Lichteffekte unterstützt [...]« (Ebd., 216)

324 Auch diesen Gedanken, dass die Beschaffenheit der Sprache der Freiheit der Montage Grenzen setze, hatte Arnheim in einem vorherigen »Beitrag zur Krise der Montage« (2004 [1930]) bereits ausgeführt. Dort schildert er als Beispiel eine Filmszene bei der ein Rundfunk sprecher auf der Galerie eines Ballsaals steht und in ein Mikrofon spricht: »Sieht man den Rundfunk sprecher vor sich, so ist es sinnvoll, seine laute Stimme zu hören. Springt aber das Bild 50 m fort ins Parkett des Festsaals und hört man dennoch die Stimme des Sprechers mit unveränderter Lautstärke fortklingen, so kann man die Stimme plötzlich nicht mehr lokalisieren, sie taumelt im Raum herum und stört. Man fühlt sich versucht unter den Gästen im Thronsaal, die man jetzt im Bilde sieht, einen herauszusuchen, aus dessen Munde die Stimme kommt – und nur wenn eine Szenenkombination gelänge, bei der die fortönende Stimme in Szene II ebenso sinnvoll lokalisiert wäre wie in Szene I, wäre der Überblendungstrick möglich und genügend elegant. (So, wenn etwa in Szene II ein Lautsprecher zu sehen wäre, aus dem die Stimme tönte!) Nun wird aber eine solche Möglichkeit nur ganz außerordentlich selten gegeben sein, und es scheint sehr fraglich, ob für dieses im stummen Film so brauchbare Montageprinzip im Tonfilm eine große Zukunft besteht« (Arnheim 2004 [1930], 79). Deutlich wird hier, dass Arnheim als überzeugter Anhänger des Stummfilms im Tonfilm eine Gefahr für die Kunst der Montage sieht, auch wenn das von ihm geschilderte Beispiel stark konstruiert erscheint. Interessant ist, dass hier die Formation des Audiovidiums (als körperlich-stimmliche Einheit) erneut das Problem wird: Weil die Stimme, die über den Schnitt plötzlich und abrupt vom Körper ihres Sprechers als Quelle gelöst wird, sich schnell einen neuen, sichtbaren ›Wirk‹ suchen muss, kommt es zur Verwirrung. Der Drang des Audiovidiums zur Unteilbarkeit wird somit zum Problem einer in der Montage immer schon dividuierenden Filmkunst. Dass aber natürlich nicht nur auf visueller, sondern auch auf akustischer Ebene mit dieser realitäts- und stabilitäts-konträren Fragmentierung der Welt gespielt werden könnte, diesen Gedanken lässt Arnheim zu dieser Zeit, vermutlich aus Sorge um den Stummfilm, nicht zu.

Im Sprechfilm bleiben also Bild- und Tonkunst getrennte Sphären, in denen das Wort in der Hierarchie klar oben steht. Die Montage ist eingeschränkt, weil auch die Logik des Wortes bestimmte Rhythmen und Strukturen vorgibt, die nicht von der Kamera durchbrochen werden dürfen.

»Montage ist aber im reinen Sprechfilm schon deshalb schwer möglich, weil man das dramatische Wort nicht, wie das Bild, in Sekundenszenen verwenden kann, sondern weil es Zeit braucht. Ein Dialog kann, darf und soll Viertelstunden lang dauern – und während dieser Zeit müssen die sprechenden Personen ruhig vor unseren Augen verharren. Es geht nicht an, daß sie zerstückelt, in einen Bilderstrauß zerlegt werden oder daß die Bildkamera inzwischen ihrer Wege geht und allerhand Zimmerwinkel, Zuhörer oder den Himmel mit Wolken einstreut.« (Ebd., 2017)

Auch bei Arnheim wird also – über diese Verhandlung des Sprechfilms – deutlich, dass es gerade das Audioviduum ist, das den Film in seiner Eigenständigkeit als Gesamtkunst (d. h. als Bild-*und*-Ton-Kunst) gefährdet. Der Anthropozentrismus bisheriger Kunstformen (wie der des Theaters etwa) ist im Angesicht der Technik scheinbar nicht der angemessene Rahmen für ein künstlerisches Schaffen, das eben dieser Technik und ihren Möglichkeiten angemessen Rechnung tragen will. Das Audioviduum bremst diese Potenziale aus – etwa weil es als audiovisuelles *Individuum* (hörbar im Ton und sichtbar im Bild) nicht durch die Technik (Montage) *geteilt* werden darf. Steht der Mensch (hier in Form von Sprache) im Vordergrund, so gerät das Medium in den Hintergrund, was es aber im Sinne seiner Bestimmung zu verhindern gilt. Arnheims Fazit fällt dementsprechend sehr eindeutig aus: »Der reine Sprechfilm hat mit Filmkunst nichts zu tun« (218; Herv. i. O.), so resümiert er, denn in dieser Form unterscheidet er sich »ästhetisch in nichts vom Theater« (ebd.).

Der Sprechfilm, und damit das Audioviduum, dient ihm für seine Argumentation somit als Kontrastfolie, die eine »Sichtbarwerdung« des Mediums (als Kunst) im Gegensatz zum Menschen möglich macht. Dort wo der Mensch im Zentrum steht, tut es das Medium nicht; dort wo man ihn als Zentrum relativiert, tritt das Medium hervor. In dieser Hinsicht ist auch seine Konstruktion des Tonfilms zu lesen, der sich vom Sprechfilm vorrangig darin unterscheidet, dass in ihm das Bild-Ton-Gefüge eben nicht mit alleiniger Rücksicht auf den Menschen gestaltet wird:

»Beim eigentlichen Tonfilm jedoch liegt für die Bildkamera kein Grund vor, in einen solchen Sympathiestreik einzutreten. Sie darf den sprechenden Schauspieler in beliebig schiefer Einstellung zeigen, ohne daß dadurch ein psychologisches oder ästhetisches Gesetz verletzt würde.« (Ebd., 217)

Nicht allein der Ton ist somit ausschlaggebend für Arnheims Definition des Tonfilms, sondern welche Konsequenzen sein Einsatz für die Bilder hat – und vor allem: welchen Stellenwert, er dabei dem Menschen als Motiv zukommen lässt.³²⁵ Sprech- und Tonfilm werden hier folglich gleichermaßen als anthropomediale Relationen beschrieben, aus denen sie hervorgehen; diese dienen als Ausgangspunkt, um medientheoretische Definitionsarbeit zu leisten.

Es ist hier also – auch in Weiterführung der bisher gezeichneten Theorieentwicklungen im Bereich des Stummfilms und Rundfunks – in der Gesamtschau eine interessante argumentative Bewegung festzustellen: Arnheims definitori-sche Ausführungen sorgen dafür, dass das Lob des alleinigen Menschenkörpers im stummen Film und der alleinigen Stimme im Radio Bestand haben kann, auch dann wenn der Tonfilm ihre (Wieder)Vereinigung ermöglicht. Die Domäne des Menschlichen (im Sinne seines künstlerischen Ausdrucks) ist dadurch für den Tonfilm wiederum eine Art Tabu, weil ansonsten Stummfilm und Rundfunk als Kunst gefährdet würden. Deshalb gewinnen die Geräusche bei Arnheim als Gegengewicht eine so große Bedeutung; um also den Ton im Film nutzen zu können, muss der Mensch (als geistig-sprachbegabtes Wesen) in die Geräuschwelt ebenso wie in die Dingwelt eingebettet und ihr gegenüber egalisiert werden, damit das Gefüge der Künste und ihrer verschiedenen Zentren keine Störung erfährt. Ähnlich wie es in den resümierenden Ausführungen zum Stummfilm geschildert wurde³²⁶, wird somit auch beim Tonfilm der Mensch als Zentrum in Frage gestellt, zugunsten eines Hervortretens des Audiovisuellen als Medium – und das gerade abseits seines menschlichen Inhalts oder zumindest: diesen relativierend. Da der Tonfilm sich zu einer Zeit etabliert, in der diese anti-anthropozentrischen Debat-ten (dinghafter Film, Geräuschradio etc.) bereits für Film und Funk angestoßen sind, bleibt für den Tonfilm argumentativ lediglich eine ebensolche Egalisierung des Menschen als legitimierender Gestus übrig.

Arnheim führt diese Ideen auch in den folgenden Abschnitten seines Buches weiter aus, z. B. indem er sich mit der »Räumlichkeit des Tonfilmbildes« (ebd.,

325 Diese Idee einer Egalisierung des Menschen im Hinblick auf seine Umwelt findet sich auch in weiteren Texten Arnheims, z. B. in einer Auseinandersetzung mit der *Bewegung* im Film, die er 1934 veröffentlicht: »Aber Bewegung gibt es im Film keineswegs nur bei den Schauspielern. Im Film ist der Mensch immer unauslöslicher Teil seiner Umwelt. Diese Umwelt spielt mit, und sie produziert Bewegungsmotive, die oft eindringlicher sind als jede menschliche Bewegung. [...] das Ziehen der Wolken, das Wiegen des Korns, der Sturz eines Wasserfalles, das Tanzen eines Bootes, das Pendeln eines Uhrpendelkels, das Auf und Ab eines Maschinenkolbens hat in schon so manche Filmszene stärkeren Ausdruck gebracht als alle Schauspielergärden. Und man versteht, warum: diese Naturbewegungen besitzen die großzügige, eindringliche, musikalische Einfachheit des Verlaufes die der pantomimischen Bewegung des Menschen mehr oder weniger abgeht« (Arnheim 2004 [1934a], 163f.).

326 Siehe Kapitel 3.1.4.

213ff.) beschäftigt. In gewisser Hinsicht anknüpfend an Leisegang, befasst er sich hierin mit der Tatsache, dass Bild und Klang unterschiedlichen Dimensionen angehören und ihre Verbindung daher immer schon gefährdet sei. So unterstelle man dem Film:

»Das auf die Leinwand geworfene Bild ist *Schein*; Scheinmenschen öffnen da den Mund – die Stimme aber, die aus diesem Munde dringen soll und die der Lautsprecher produziert, ist nicht Schein sondern eine *wirkliche* Stimme. Denn von Tönen kann es keine Scheinbilder geben! Schein und Wirklichkeit aber können sich niemals zu einem geschlossenen Kunstwerk zusammenfinden. Nun, das ist eine Frage des praktischen Experiments! Die Erfahrung hat gelehrt, daß die Vereinigung von Wort und Bild *zwei Typen von Reaktionen* bei den Zuschauern auslöste. Die einen vermohten Bild und Ton nicht im geringsten zu einem einheitlichen Eindruck zu verbinden. Die Menschen auf der Leinwand wirkten plötzlich erschreckend unkörperlich, flächig, schattenhaft und tot, während von irgendwo außerhalb des Bildes her für sich und ohne Beziehung zum Bild die Stimmen erklangen. Oder aber: die Stimme drang aus dem Munde der Leinwandfiguren, und diese, wie vom Ton belebt, erschienen plötzlich lebendiger, plastischer, leibhaftiger als beim stummen Film.« (Ebd., 220f.; Herv. i. O.)

Das Audioviduum wird hier also als ein ambivalentes Phänomen beschrieben, das zwischen glaubwürdiger Verschmelzung und dem Auseinanderfallen in seine Einzelteile schwankt. Der Ton wird dabei – wie selbstverständlich – als etwas angenommen, das weniger ›medialisiert‹, sondern stattdessen in seiner technischen Reproduktion ›echter‹ oder zumindest naturnäher erscheint. Doch selbst wenn Bild und Ton im Film auf technischer Ebene getrennt blieben, glaubt Arnheim doch daran, dass der Wille zur Verschmelzung beider Teile in der Rezeption deutlich höher ist. Dass das Audioviduum sich manches Mal nicht so einfach herstellen ließe, dürfte

»[...] vor allem darauf zurückzuführen sein, daß der Ton die Illusion eines tatsächlichen *Raumes* erweckt, während das Filmbild kaum räumliche Tiefe und Körperllichkeit hat. Dieser Widerspruch in den räumlichen Verhältnissen steht der Vereinigung der optischen und akustischen Reize entgegen. (Und natürlich auch die objektiv falsche Lokalisation der Sprechtöne, die nicht aus den Mündern sondern unten aus dem fixierten Lautsprecher kommen). Nun passen aber, hiervon abgesehen, die optischen und akustischen Phänomene aufs Beste zueinander – sind sie doch Reproduktionen ein und desselben wirklichen Vorgangs (die Mundbewegungen passen zu den Worten etc.) – Und so haben die beiden Reizgruppen andererseits die kräftigste Tendenz, miteinander zu verschmelzen. Je nachdem welche der beiden Tendenzen siegt, die auseinander- oder die zusammentreibend-

de, entsteht das eine oder das andere der oben angegebenen Erlebnisse.« (Ebd., 221; Herv. i. O.)

Arnheim beschreibt den Tonfilm damit – in explizitem Bezug auf seine Menschenmotive – als einen Prozess der Individuation, in dem eine innere Dynamik diesen individuierenden Verschmelzungswillen anstößt, dieser aber stets eingebunden bleibt in äußere, wiederum deindividuierende Bedingungen, aus denen das Audioviduum erst hervortreten muss. Doch welche Relevanz haben nun diese Beschreibungen auf inhaltlicher Ebene für die Frage der Medialität des Tonfilms auf theoretischer Ebene?

Folgt man Arnheims Ausführungen weiter, so setzt er sich in einem nächsten Schritt noch einmal intensiver mit Unterschieden zwischen Stummfilm und Tonfilm auseinander (vor allem im Hinblick auf die Rolle der Mimik und Gestik), um dann zur Frage von Parallelismus, Kontrapunktik und Asynchronismus zu gelangen (vgl. ebd., 235ff, 242ff.). In expliziter Abgrenzung von Pudowkin macht er dabei allerdings klar, dass er den Vorschlag einer rigiden Kontrapunktik nicht teilt:

»Bei dieser Gelegenheit sei noch ausdrücklich betont, daß nicht etwa jedes Bild eines sprechenden Menschen ein unerwünschter Parallelismus in unserm Sinne ist! Denn ehr häufig wird das Mimische eine Bereicherung dessen bringen, was der Sinn der Worte und der Klang der Stimme ausdrückt. Und durchaus nicht immer wird der Inhalt der Worte nur eine sprachliche Formulierung dessen sein, was man ohnehin mit den Augen sieht. Immerhin liegt hier die Gefahr sehr nahe, daß das Miteinander von Ton und Bild auf eine bloße abschwächende Verdopplung herausläuft. An der Kunst, den sprechenden Schauspieler trotzdem in wirklichen Tonfilmbildern zu zeigen, wird man den guten Tonregisseur besonders deutlich erkennen können.« (Ebd., 239)

Für Arnheim ist die Frage der natürlichen Synchronität, die Pudowkin und auch die weiteren Autoren des Tonfilmmanifests ablehnen, also nicht das grundsätzliche Problem des Tonfilms; lediglich das *Verhältnis* zwischen Bild und Ton müsse entsprechend stimmen, sodass weder der Ton das Bild dominiere noch umgekehrt. Zwar stehe »rein technisch [...] nichts im Wege, auch solche Töne und Bilder miteinander zu kombinieren, die durchaus nicht zugleich und am gleichen Ort aufgenommen worden sind« (ebd., 243) so Arnheim weiter. Dies sei aber eher im grotesken oder animierten Film der Fall und als gängige Praxis nicht unbedingt sinnvoll. Auch die Nachsynchronisation lehnt er ab, weil hier eben sichtbare Körper und hörbare Stimme nur unter schwierigsten Bedingungen die geeignete audiovisuelle (bzw. audioviduelle) Symbiose eingehen könnten. Die Synchronität ist somit nicht das Problem des Tonfilms, sondern stellt lediglich die entscheidende Herausforderung für die Filmschaffenden dar, die lernen müssen diese

beiden verschmelzungswilligen Ebenen – Bild und Ton bzw. in diesem Falle einmal mehr: Körper und Stimme – in einen ausgewogenen und damit künstlerisch wertvollen Einklang zu bringen.

Arnheim akzeptiert insofern den individuierenden Charakter des Films, den er anhand des Menschenmotivs erklärt, der aber – im Sinne der Egalisierung von Mensch und Dingwelt – auch für die Geräuschebene gilt. Welche Konsequenzen das für die medientheoretische Definition des Tonfilms hat, wird noch etwas deutlicher, wenn man einen weiteren Text Arnheims hinzuzieht, den dieser unter dem vielsagenden Titel »Tonfilm mit Gewalt« (2004 [1931]) schon zuvor als explizite Kritik an Pudowkins Kontrapunktik verfasst hatte.³²⁷

»Es ist recht zweifelhaft, wie weit das Prinzip des Asynchronismus überhaupt angewendet werden soll. Ob es gut ist, etwa die Stimme eines unsichtbaren Sprechers mit Bildern zu illustrieren. Wie weit sich Begleitmusik und Synchronegeräusche vertragen. Die Einwände dagegen sind vor allem ästhetischer Natur. Man verkennt den elementaren Wert der Anschaulichkeit im Film, wenn man Dinge übereinanderklebt, die nur dem Sinn nach zusammenpassen, aber keine sinnliche Einheit ergeben. Es mag nicht verboten sein, so etwas zu tun, aber zweifellos ist das künstlerische Niveau solcher Einfälle im Prinzip erheblich geringer als das jenes anderen Typus [...]. Von der Theorie her klingt die Methode des Asynchronismus vornehmer, exklusiver, ästhetischer. In Wirklichkeit, ist siedürftig, weil sie im Begriff kleben bleibt, statt die Verleiblichung des Gedankens in der unzerschnittenen, unmißhandelten Wirklichkeit selbst aufzusuchen.« (Ebd., 88f.)

Diese und auch die vorherigen Ausführungen sind für die Idee einer Individualität der Medien natürlich äußerst interessant. So wird die Eigenständigkeit der Tonfilmkunst bei Pudowkin (u. a.) von seinem Vermögen abhängig gemacht, bestehende, ›natürliche‹ Einheiten zu zerstören. Arnheim hingegen hält diesen Ansatz scheinbar für zu prätentiös; dem gewollten Zerstören von Einheiten hält er die stets notwendige Generierung neuer Einheiten entgegen, die dabei sehr wohl der »unzerschnittenen [...] Wirklichkeit« entnommen sein könnten, um sie dann nur anders zu kontextualisieren.

»Gute Tonfilmkunst wird nicht dort entstehen, wo man Mikrophon und Bildkamera zum Hahnenkampf aufeinanderhetzt. Man respektiere die natürliche Einheit des Augen-Ohren-Erlebnisses. Denn in der Kunst regieren die Sinne.« (Ebd., 90)

³²⁷ Für weitere Auseinandersetzungen mit der Kontrapunktik siehe z. B. auch Arnheim (2004 [1934b]; 2004 [1934c]).

Auch hier also geht es um eine unteilbare und damit individuelle Verknüpfung von Bild und Ton. Nimmt man nun diese Idee einer – inhaltlich *sinnvollen* wie materiell *sinnhaften* – Individualität der medialen Audiovision als deren Grund und Ziel und verbindet sie mit der zuvor skizzierten Idee Arnheims, den Menschen wiederum nicht als Zentrum des Tonfilms als Kunst anzusehen, sondern lediglich als eines seiner Elemente, so scheint sich genau darin die Medialität des Tonfilms zu offenbaren. Obwohl dieser (auch) eine Sprech- und Sprachkunst sei wie Theater oder Literatur, sei er – anders als diese – eben nicht auf den Menschen als zentrales Motiv angewiesen; und dennoch kann er Individualitäten erzeugen, indem er anderweitig Bezüge herstellt – nämlich gleichermaßen zwischen Mensch, Ding und Medium in unterschiedlichsten *sinnhaften*, *sinnvollen* und dadurch künstlerisch einzigartigen Konstellationen.

Mit Ideen wie diesen arbeitet sich Arnheim also systematisch an der Frage ab, die er zu Beginn noch als offen formuliert hatte – inwieweit ist der »Tonfilm [...] eine eigene Kunstform ...?« Zehn Jahre nach dieser Frage von 1928 (und sechs Jahre nach den bisherigen Antworten aus *Film als Kunst*) folgt mit *Neuer Laokoon* (Arnheim 2004 [1938]) ein weiterer Text, der versucht finale Antworten zu geben. Dabei ist interessant zu beobachten, dass Arnheim hier von der Möglichkeit einer ausgeglichenen Bild-Ton-Kunst, die er zunächst noch angedacht hatte, abrückt. Als Einstieg in den Text formuliert er diesen Zweifel zunächst konkret:

»Die Anregung zu der folgenden Untersuchung gab ein Gefühl des Unbehagens, das den Verfasser immer wieder vor dem Sprechfilm befällt und das von keiner Gewöhnung zu stillen ist: der Eindruck, daß da etwas nicht in Ordnung sei; daß da etwas vorgeführt werde, was wegen grundsätzlicher innerer Widersprüche lebensunfähig bleiben müsse. Dieses von den Sprechfilmen erregte Unbehagen scheint dadurch verursacht zu sein, daß die Aufmerksamkeit des Zuschauers zu zwei entgegengesetzten Lagern hin auseinandergezogen wird, daß um den Zuschauer gekämpft statt gemeinsam geworben wird und daß im Bestreben, das gleiche Ding doppelt zu sagen, ein verwirrendes Nebeneinander zweier Stimmen entsteht, die jede ihre Sache nur halb sagen können, weil sie einander stören.« (Ebd., 377)

Arnheim rechnet in diesem Text also final mit dem Sprechfilm ab; alle potenziellen Gewöhnungsprozesse der vergangenen Jahre haben die Zweifel nicht ausräumen können, wie es scheint. Und: Die Frage nach Bild und Ton wird hier erneut ganz explizit als eine Frage nach Körper und Stimme verhandelt. Deshalb spricht Arnheim auch nicht von Bild und Ton, sondern von Bild und Wort, »deren Wettstreit der Sprechfilm zu keiner Einheit zu zwingen« wisse (ebd., 378). Erstaunlich findet Arnheim dies vor dem Hintergrund, »daß in der alltäglichen Lebenserfahrung Sichtbares und Hörbares ganz natürlich zu einem einheitlichen Weltbild ver-

schmolzen« seien (ebd.) So sei im Alltag das Sprechen keine Störung des Bildes (und umgekehrt), wohl aber auf der Projektionswand des Kinos. Dies erklärt er damit, dass die Verbindungen im Film Ergebnis artifizieller Prozesse seien, sodass man ihre Kombination ständig mit einer spezifischen Bedeutung auflade. Denn die Kunst bestehe gerade darin, dass sie die natürlich kombinierten Dinge der Welt in Einzelsphären aufzusplitten vermöge, wodurch in dieser einzelnen Sphäre eine künstlerische Ganzheit herzustellen sei.³²⁸ Deshalb könne es auch nicht Sinn einer Kunst sein, die durch sie produktiv zerstörten, selbstverständlichen Verbindungen wieder herzustellen:

»Optische und akustische Elemente, etwa gesprochene Worte und bewegte Körper, im Kunstwerk zu einer elementaren Einheit verschweißen zu wollen, in der gleichen Art wie sich Satz an Satz, Bewegung an Bewegung fügt, ist ein unsinniges, unvorstellbares Beginnen. Die gegenständliche Einheit, die ›im Leben‹ z.B. den Körper und die Stimme eines Menschen umschließt, würde im Kunstwerk eine viel elementarere Gleichartigkeit dieser Faktoren und viel elementarere Beziehungen zwischen ihnen voraussetzen.« (Ebd., 380)

Das bedeutet also, die Ganzheitlichkeit, die in der Einzelsphäre einer Kunst (Bild in der Malerei, Sprache in der Literatur etc.) generiert wird, würde durch unbedachte Verbindungen zu außerhalb dieser Sphäre liegenden Elementen prinzipiell zerstört oder zumindest gestört. Es gilt also zunächst die Scheidung der sinnlich differenzierten »Materialsphären« als das fundamentale Grundprinzip von Kunst anzuerkennen (»Materialstufe« 1), und erst dann könne man sich Gedanken über deren Verbindungen machen (»Materialstufe« 2 und ggf. höher) (ebd. 381). Die »Beziehung zwischen Menschenkörper und Menschenstimme« nun sei in ihrer irdischen Ausprägung aber einerseits zu natürlich, um die künstlerisch sphärische Trennung klar zu vollziehen; gleichzeitig sei sie im Film aber zu elementar unterschieden, um sinnvolle Beziehungen herstellen zu können (ebd., 383). Und genau das sei der Grund, warum sich Film und Ton in Bezug auf den Sprechfilm nicht angemessen zu einer einheitlichen Kunst verbinden könnten, denn die Sprache verweise immer unweigerlich auf den Menschen (quasi anthropozentrisch), während im Stummfilm der Mensch als Teil der Dingwelt erscheine (quasi anti-anthropozentrisch). Arnheim beschreibt den Dialog daher auch als eine Art Zwang, der dem Film im Sprechfilm auferlegt würde:

328 Hier ist auch der klare Bezug zu Lessings Laokoon zu sehen, der ebenfalls die verschiedenen Ausdruckssphären der Kunst im Hinblick auf bildende Kunst und Literatur hin untersucht (vgl. Lessing 1766).

»Und doch darf man nicht voreilig die Tonkunst und die Bildkunst in ihrem Verhältnis zum Wort gleichsetzen. Es liegt im Wesen des dramatischen Dialogs, daß er das Kunstwerk vollständig auf den handelnden und verhandelnden Menschen einschränkt. [...] Das Bild hingegen zählt, auch ohne das Wort, den Menschen bereits zu seinen Gegenständen. Nun nimmt aber der Mensch im Bilde nicht jene unbedingte Vormachtstellung ein wie etwa auf der Bühne. [...]. Auch dem bewegten Bilde, dem Film, lag die vom Menschen belebte Welt von vornherein näher als der aus der Welt gelöste Mensch. Und deshalb mußte ihm die vom Dialog erzwungene Beschränkung auf den Menschen unerträglich sein. Der Schauplatz der dramatischen Wechselrede ist die Menschenseele, seine einzige mögliche optische Ergänzung der von Wahrnehmung, Gefühl, Wille und Gedanke geprägte und gelenkte Menschenleib. Die Umgebung ist – bühnenhaft – als Hintergrund zugelassen, so wie im Dialog die Seele sich gelegentlich durch Hinweise auf die Außenwelt ausdrückt; aber die Verlagerung des Schwerpunktes auf den Menschen ist zwingend und damit eine wesentliche und gerade der Weltanschauung der Gegenwart wieder so vertraute Haltung des Bildes vernichtet.« (Ebd., 403)

Es mache also weder Sinn auf diesen automatischen Menschenbezug der Sprache mit einer umso stärkeren Betonung der Bilder zu reagieren, weil dadurch die Sphären nur umso mehr auseinanderfielen. Gleichzeitig bleibe man bei einer Dominanz der Sprache zu nah am Theater, um eine eigene Kunst zu sein, und bei einer Dominanz der Bilder könne man auch gleich beim Stummfilm bleiben (was Arnheim insgeheim zu befürworten scheint). Denn die Ding-Bezogenheit des Stummfilms sei es gewesen, die ihn überhaupt erst als Kunst neben den menschenzentrierten Künsten des Theaters, der Literatur etc. habe bestehen lassen. Durch die Option zur Sprache und zum Dialog aber, die der Tonfilm biete, sei diese Einzigartigkeit stets direkt gefährdet, weil die Aufmerksamkeit automatisch vom Ding auf den Menschen übergehe.

»Natürlich hat auch der Stummfilm den Schauspieler oft in Großaufnahme gezeigt. Aber was wichtiger ist, er hat eine Verbindung von stummen Menschen und stummen Dingen geschaffen, ebenso wie von der (hörbaren) nahen Person und jener (nicht hörbaren) weiter weg. In der allgemeinen Stummheit des Bildes konnten die Scherben einer zerbrochenen Vase genauso ›sprechen‹ wie ein Darsteller zu seinem Nachbarn, und eine Person, die sich auf der Straße nähert und am Horizont nur wie ein Punkt zu sehen ist, ›sprach‹ wie jemand in der Großaufnahme. Diese Gleichförmigkeit, die dem Theater vollkommen fremd, der Malerei hingegen vertraut ist, ist zerstört durch den Sprechfilm: Er stattet den Schauspieler mit Sprache aus, und seit nur dieser sie haben kann, sind alle anderen Dinge in den Hintergrund gerückt. Es gibt eine Grenze des sichtbaren Ausdrucks, der aus der

menschlichen Gestalt gezogen werden kann, insbesondere wenn das Bild Dialog zu begleiten hat.« (Ebd., 404)³²⁹

Der Sprechfilm als Kunst scheint somit eine Unmöglichkeit und die zuvor im *Film als Kunst* noch geschilderte Option eines Tonfilms auf der Mitte beider Ausprägungen (Sprechfilm auf der einen und Stummfilm auf der anderen Seite) ebenfalls kein Thema mehr zu sein. Doch der Ausgang dieser tonfilmkritischen Aushandlung ist dabei für die vorliegende Untersuchung weniger relevant. Zentral ist die Feststellung, dass Arnheim hier erneut die Beschaffenheit des Tonfilms (und zwar stärker noch als in seiner vorherigen Monografie) explizit im Hinblick auf seine möglichen Menschenmotive hin verhandelt – und vor allem problematisiert. Der Filmton bzw. das Wort (Arnheim scheint diese Begriffe nahezu gleichsetzen zu wollen) bedeute

»[f]ür das Bild [...] die Ersetzung des optisch ergiebigen handelnden durch den optisch unergiebigen sprechenden Menschen. Der Dialog schneidet also nicht nur das Bild des Menschen aus dem seiner Umwelt heraus, sondern lässt auch den Ausdruck dieses Menschenbildes selbst erstarren.« (405)

Der »optisch [...] sprechende Mensch« erscheint hier also auch für Arnheims Frage nach dem Tonfilm als eigenständiger Kunst³³⁰ »unergiebig«. Für die vorliegende Arbeit hingegen bestätigt sich in dieser Konstellation und der intensiven Verhandlung des (egal ob handelnden und/oder sprechenden) Menschen erneut die These von der Ergiebigkeit dieser Denkfigur innerhalb medientheoretischer Auseinandersetzungen.

Die Ding-Orientierung des Films und somit seine einzige Chance auf Eigenständigkeit erscheint bei Arnheim am Ton (bzw. am Konkurrenzverhältnis von Körper/Dingwelt und Stimme) final gescheitert. Dass dieses Urteil nicht so ne-

329 Und wieder wird analog zu Luhmann (1995, 126) sowie Wendler/Engell (2009, 39ff.) eine Vase der Anhaltspunkt, um Unterschiede und Parallelen zwischen menschlichem und dinglichem Individuum zu erklären (siehe auch Kapitel 2.1.3 und 2.2.1).

330 Zum Ende seines Artikels verweist Arnheim auf die einzigen Bereiche, in denen der Tonfilm – dann aber eher als Werkzeug denn als Kunst – sinnvoll sein könnte. Dies sei zum einen das Ausnutzen seines dokumentarischen Charakters für das »Festhalten von Natur und Mensch« (Arnheim 2004 [1938], 410) sowie zum anderen für die breite Zugänglichmachung von Kunst (wobei hier zum Ton des Films noch Farbe und Plastik hinzukommen müssten für die angemessene Illusion) (vgl. ebd.). Dabei seien Kunst und Dokumentarisches zwei widerstrebende Modi: »Die dokumentarische und die künstlerische Absicht führen also zu teilweise verschiedener Bewertung und Verwendung der Mittel. In gewissen Grenzen wird die Kunst durch mechanische Treue, die mechanische Treue durch Kunst gefördert. Darüber hinaus entsteht Feindschaft zwischen den beiden Haltungen« (ebd., 411).

gativ ausfallen muss, zeigt ein Blick auf ein weiteres Buch aus dem Jahre 1938, das als ein ergänzender Beleg dafür dienen kann, dass der (bei Arnheim missglückte) Versuch den Tonfilm jenseits anthropozentrischer, anthropomorphozentrischer und anthropophonozentrischer Ansätze zu definieren auch an anderer Stelle aufgegriffen wird, nämlich bei Bruno Rehlinger in seinem Buch *Der Begriff Filmisch* (1938).

Das Audioviduum und das ›Filmische‹

Rehlinger geht in seinem Versuch einer Bestimmung des ›Filmischen³³¹ davon aus, dass es für eine Definition des Tonfilms nicht zielführend sei, dessen Problem vorrangig im Wort zu sehen.

»Bedeutsamer als die Einwände gegen den Ton im allgemeinen sind die gegen die *Verwendung des Wortes* im Film. Auf diese verdichtet sich der größte Widerstand. Gern beruft man sich hierbei auf die Beobachtung, daß auf der Leinwand wohl das Bild des Menschen, aber nicht das Bild seiner Stimme, sondern diese Stimme selbst erscheine.« (Rehlinger 1938, 17; Herv. i. O.)

Es wird also deutlich, dass auch bei Rehlinger das Audioviduum bzw. seine problematisierte Vormachtstellung innerhalb der Tonfilmdebatte den Ausgangspunkt bildet, den es aber eigentlich zu überkommen gilt. Ziel seines Buches ist es daher noch einmal die »Unterschiede der filmischen Gestaltungsweise nicht nur im Vergleich zum Drama, sondern umfassender zur Wortkunst klarzustellen« (ebd., 1f.) und er stellt zudem fest: »Der reine Begriff ›Tonfilm‹ hat mit der landläufigen Vorstellung vom Sprech-, Gesangs-, also Spielfilm nichts zu tun« (ebd., 16). Von Anfang an ist somit klar, dass der Mensch (in Form von Sprache, Gesang und Spiel) für Rehlinger gerade nicht zu den zentralen Definitionsmitteln des Films gehört.

Dazu setzt er sich zunächst mit den Argumenten von Fedor Stepun (1932) auseinander. Den Unterschied, den dieser zwischen ›wirklicher Stimme‹ im Kontrast zum ›unwirklichen Bild‹ aufmache (und der sich im obigen Zitat schon angedeutet findet und ähnlich von Leisegang 2002 [1929] beschrieben wurde), hält Rehlinger (1938, 17) dabei nur für einen »Sonderfall«, der daher für eine Definition des Films nicht tauge. Zudem fordere Stepun die Ablösung der menschlichen Stimmen von ihren Körpern, was laut Rehlinger damit einhergehe, dass »das Wort [...] vornehmlich als Klang erfaßt und gewollt und in die Lautäußerungen der Dinge eingeschmolzen« werde (ebd., 17). Diese von Stepun als zentrales Merkmal des Tonfilms vorgestellte Verdinglichung von Stimme sieht Rehlinger jedoch als unnötig an, weil in diesem Moment das Wort für den Film keine eigentliche oder besonde-

³³¹ Laut Knut Hickethier sei Rehlinger damit einer der ersten, der »[e]ine Bestimmung der Mediälität des Films [...] vorgelegt« habe (Hickethier 2010, 251).

re Bedeutung mehr hätte und daher auch einfach ausgeschlossen werden könnte (statt es selbst in seiner Problematisierung wieder zentral zu stellen). Weder die »Naturferne« des Films, die Stepun einfordert, noch die fortgeführte Betonung des Wortes teilt Rehlinger also; und man könnte es als klare Spitze gegen Arnheim³³² werten, wenn er schreibt: »der Sprechfilm ist nur ein Randgebiet, freilich als solches besonders geeignet, Klärungsgriffen eine breite Fläche anzubieten« (ebd., 18). Rehlinger scheint also insofern der These der vorliegenden Arbeit zuzustimmen, als dass in all diesen Fällen (Stepun, Arnheim etc.) das Audioviduum als eigentliche Randerscheinung des Tonfilms zum Zentrum versuchter »Klärungsgriffe« gemacht wird.

Nachdem Rehlinger im Folgenden anderweitige Weiterentwicklungen des Stummfilms – wie Farbfilm, plastischen Film oder auch »Komplettfilm« (ebd., 19) – streift, nimmt er schließlich Vergleiche und Abgrenzungen zu anderen Künsten wie bildender Kunst (ebd., 23ff.), Musik (ebd., 29ff.), Wortkunst (37ff.) und Hörspiel (53ff.) vor. Auf diesem Wege kommt er zu dem Schluss, dass der Tonfilm sich von diesen anderen Künsten gerade darin unterscheide, dass er den Menschen weder in Bild noch Ton zentral stellen müsse³³³ sondern dass er ihn bis zur Unkenntlichkeit in der Dingwelt auflösen könne: »Der letzte Steigerungsgrad der Einordnung des Menschen in die Welt der Dinge ist erreicht, wenn der Mensch aus ihr ausgeschaltet wird« (ebd., 44) – so formuliert es Rehlinger und als Beispiel nennt er etwa Filme, die Mikroorganismen oder den Kosmos zeigen (in die der Mensch maximal noch als Beobachter – z. B. eines Ameisenhaufens – eingelassen sei) (vgl. ebd.). Ein »Grenzfall« sei auch generell der »Naturfilm«, in dem der Mensch in seiner Umwelt gezeigt werde; und »[j]enseits der Grenze [lägen] die ungegenständlichen Zeichenfilme und die abstrakten Linienspiele, die der menschlichen Sphäre von vornherein entrückt« seien (ebd., 44f.). Das Besondere des Films sei also, dass diesem »die Stufenleiter von der Monopolstellung des Menschen (als Star) bis zu seinem völligen Ausschluß aus der Welt bewegter Formen zur Verfügung« stehe (ebd.).

Diese Idee faltet Rehlinger in einer »Grundlegung der filmischen Gestaltung« (ebd., 57ff.) und aufbauend auf seinen Vorbeobachtungen sodann weiter aus. Zunächst unterliege der Film einem »Gesetz der Zweipoligkeit« (ebd., 57) durch seine Präsenz in Form von Bewegtbild und Ton, wobei diese in einem entsprechenden Verhältnis zueinander stehen sollten. Ähnlich wie Arnheim in *Film als Kunst*

332 Rehlinger bezieht sich in seiner Arbeit zumindest auf Arnheims *Film als Kunst. Da Neuer Laokoon* erstmals im August 1938 auf Italienisch erschien, wird Rehlinger diesen Text höchstwahrscheinlich (noch) nicht gekannt haben.

333 Hier verweist er vor allem auf Hans Richters Buch *Filmgegner von heute – Filmfreunde von morgen* (1981 [1929]), in dem dieser bildreich die Grundgestaltungsformen des Films illustriert, wobei der Mensch eine deutlich untergeordnete Rolle spielt.

spricht er dabei von einer Art ausgeglichener »Gewichtsverlagerung nach zwei Polen, dem optischen und dem akustischen, sodaß entweder der Ton oder das Bild die Aufgabe der Ergänzung übernimmt« (ebd., 57f.). Zudem heiße Tonfilm in diesem Sinne, dass *alle* Arten von Tönen damit gemeint sein könnten:

»Dem Bild steht die gesamte Tonwelt: Wort, Geräusch, Musik gegenüber. Die einseitige Bevorzugung eines dieser Tonmittel führt zu den Teil- oder Nebenformen des Sprech-, Geräusch- Gesangs-, Schlagerfilms, während in dem übergeordneten Begriff *Tonfilm* der Hinweis dafür gegeben ist, den Tonteil des Films als ein Ganzes zu betrachten.« (Ebd., 58; Herv. i. O.)

Im Zuge dessen kritisiert Rehlinger Arnheims Grenzziehung zwischen Sprechfilm und Stummfilm als zu einseitig; vielmehr sei die einzig zulässige Unterscheidung zwischen *Stumm-* und *Musikfilmen* zu treffen. Erstere seien Filme, in denen sich der Ton ganz nach dem Bild richte (oder verschwinde) und letztere solche, in denen sich das Bild ganz dem Ton (bzw. als Extremform der Musik) unterordne. Der Tonfilm pendle nun genau zwischen diesen Extremen der Ausrichtung (vgl. ebd.).

Doch am zentralsten für Rehlingers Argument (und damit ebenso konträr zu Arnheims *Laokoon*) ist die Idee, dass der Mensch auch im Tonfilm weiterhin – so wie es im Stummfilm schon der Fall war – verdinglicht werden könne. Das Besondere des Tonfilms sei also die »filmische Gestaltung der Welt« (ebd., 59), die er ermögliche. Um dies zu begründen, arbeitet er sich allerdings dennoch zunächst am Menschenbild ab. Dieses sei im Film – vor allem im Gegensatz zu Theater und Hörfunk – gekennzeichnet durch:

»einmal die relative Bedeutungslosigkeit des Menschen an sich im Film und damit die Notwendigkeit seiner Einordnung in die Welt der Dinge, sodann die Arbeit der Kamera, der gegenüber die Eigenleistung des Menschen nur zweiten Wert besitzt. Der Unterschied zum Bühnenschauspieler auf der einen Seite, zum Funk sprecher auf der anderen liegt auf der Hand. In jedem Fall wird hier nur die Leistung des Menschen gewertet. Nicht-Spieler, Nicht-Sprecher, vor allem aber Kinder, die im Film einen hinreichenden Anblick gewähren können, nehmen sich auf der Bühne immer, beim Funk zum großen Teil unorganisch und hilflos aus.« (Ebd., 59; Herv. i. O.)

Rehlinger sieht hier also im Film die menschliche Darstellung durch die Kamera und durch die Gleichstellung mit der Dingwelt überformt; der ‚Kamerablick‘ sei besser darin, den tatsächlichen Menschen zu zeigen, als sein Spiel. Aus diesem Grunde erklärt Rehlinger auch Nicht-Schauspieler (wie etwa Kinder) zum eigentlichen Gegenstand des Tonfilms, wodurch er ihn in seiner Bestimmung in die

Nähe des Dokumentarfilms rückt. Gerade die Unbewusstheit und Unbedarftheit des Kindes passe zum ehrlichen Modus des Films, ebenso wie dies bei »fremden Völkern« der Fall sei (ebd.), so stellt Rehlinger primitivisierend fest. Und er proklamiert: »Im Naturwesen haben wir die filmisch reinste Erscheinung des Menschen« (ebd., 61; Herv. i. O.).

Als zweites, einzigartiges Vermögen des Films nennt Rehlinger sodann die Inszenierung der Masse, die die russischen Formalisten im »Massengesicht« zur Perfektion geführt hätten (vgl. ebd.). Im Hinblick auf das Audioviduum scheint also gerade das Auflösen des menschlichen Individuums in der Masse ebenso wie die Darstellung »fremder« Völker und Kinder, denen man also (noch) keinen direkten (bürgerlichen) Individualitätsstatus zuerkennt, gleichermaßen das de-individuierende und ent-individuerte Potenzial des Films zu offenbaren. Es geht Rehlinger also um die Verbannung des (individuellen) Menschen aus dem Zentrum des filmischen Gestaltungskosmos.

Der Mensch ist dabei – in Form des »Naturwesens«, der »Masse« und erst an dritter Stelle in Form des »Schauspielers« (ebd., 62) dennoch Teil dieses Gestaltungsrahmens, aber eben nicht mehr dessen Zentrum. Und die (Ton)Film-Schauspieler:innen seien dabei auch noch anders beschaffen, als z. B. die auf dem Theater, weil sie in der Fläche anders agieren müssten und nicht im Rezeptionsraum anwesend seien:

»Die Tatsache, daß eine ›Dingwand‹ die Erlebnisse des Filmschauspielers und des Zuschauers nicht ineinander greifen läßt, ist so bedeutsam, daß von ihr aus der Schauspieler für den Film und das filmische Erlebnis nur als ›Notbehelf‹ angesehen werden kann.« (Ebd., 62)³³⁴

Ebenso wie der Sprechfilm bei Rehlinger zum »Randgebiet« (ebd., 18) wird, wird also der Schauspieler zum »Notbehelf« (ebd., 62); während Bühnendarsteller »Ausgestalter des dichterisch-dramatischen Wortes« seien, sei der Schauspieler des Films lediglich »optische und akustische Ausdrucksform« (ebd.).³³⁵

334 Im Zuge dessen beruft er sich erneut auf Richter (1929), der ebenfalls die Bedeutung des Schauspielers für den Film in Frage stellt: »Der Wert des Schauspielers im Film ist relativ nicht größer als der irgend eines anderen im Film wirkenden Objekts, wenn es nur zur Ausdruckssteigerung des Films beiträgt. [...] Der Schauspieler ist im Film problematisch. Er ist ebenso Mittel des Films wie Tintenfaß, Vogel, Baum. Wichtig ist er nur in dem Maße, als er am Rhythmus des Films teilnimmt. [...] Das beste Mienenspiel ... hat wenig zu bedeuten im Film, in dem richtige Einstellung und Beleuchtung das Gleiche weit stärker zum Ausdruck bringen (selbst ohne das Gesicht des Schauspielers zu zeigen)« (Richter 1929, zit. n. Rehlinger 1938, 62).

335 In Bezug auf die Bildebene führt er dies z. B. im Hinblick auf das »Filmgesicht« detaillierter aus, das dabei eben »nicht die Domäne des Filmschauspielers« sei (als eine Art gesteigerte, professionelle, filmische Art der Mimik), sondern »ganz einfach: das Antlitz des Menschen im Film«

Auch Rehlinger stellt hier in seiner Auseinandersetzung mit dem Film somit kurzzeitig den Menschen (z. B. als Schauspieler) ganz ins Zentrum seiner Betrachtung; dennoch aber wird klar, dass es ihm dabei gerade darum geht dessen Vormachtstellung innerhalb der bisherigen Tonfilmdebatten zugunsten einer egalisierenden Einbettung ins Filmbild zu relativieren. Und er kommt zu dem Schluss: »die Bühnenpantomime ist ohne den Menschen undenkbar, der Film aber kann – im Höchstfall – den Menschen völlig ausschalten und trotzdem filmisch sein« (ebd., 66).

Rehlinger stellt hier also quasi dem Anthropozentrismus des Theaters eine Art posthumanistische Dezentrierung des Menschen im Film entgegen.³³⁶ Auch wenn der Film in besonderer Weise das menschliche Antlitz als »*Filmgesicht*« (ebd., 63) zeigen könnte, so sei das nur eine von zahlreichen Optionen.

»In die Begriffsbestimmung des Filmkunstwerks als solches kann der Schauspieler als Wesensbestandteil nicht einbezogen werden. (Dies zum Unterschied vom dramatischen Wortkunstwerk, dem er wurzelhaft verbunden ist.)« (Ebd., 69; Herv. i. O.)

Erneut ist hier also das Argumentationsmuster zu finden, dass gerade in der Ausgrenzung des (ganzheitlich ebenso wie partiell gedachten) Audiovidiums aus dem Darstellungs- und Motivrepertoire des Films seine eigentliche Bestimmung und damit Definition als Medium begründet wird.

Zum Ende seines Buches wendet sich Rehlinger dann noch einmal dezidiert dem Tonfilm zu und den Gestaltungsoptionen, die dieser im Hinblick auf die Verdinglichung und damit Dezentrierung des Menschen liefere, wozu er »Wort«, »Geräusch« und »Musik« als seine potenziellen Ausprägungen einzeln untersucht (ebd., 71ff.). In Bezug auf das »Wort« sei dabei festzustellen,

(Rehlinger 1938, 63; Herv. i. O.) »Das Filmgesicht ist nicht an die Forderungen der Schönheit, Aufmachung, Schminke und maskenhaften Typisierung gebunden – so Rehlinger (ebd.); stattdessen sei »an die Stelle des berechenbaren ›Wunschkörpers‹ der Fabrikation die Wahrheit des ›Menschenbildes‹ zu setzen« (ebd.). Deswegen sei der »Nicht-Schauspieler« dem »Schauspieler« im Film »unendlich überlegen« (ebd., 64).

336 Sollten im Film dennoch Schauspieler zum Einsatz kommen, so »[ließe sich] als allgemeiner Grundsatz [...] die Feststellung herleiten: daß *im Film die schauspielerische Leistung hinter der menschlichen Intensität* zurücksteht. Es ist eine gute Beobachtung, wenn man [...] die Aufgabe des Filmdarstellers nicht in der Übertragung des schauspielerischen ›Ichs‹ auf die verkörperte Figur sieht, sondern in der Übertragung der Figur und ihrer notwendigen natürlichen Zufälligkeit und Einmaligkeit auf den menschlichen Grundgehalt als die ›Wiedergewinnung der Natur im Spiel‹« (ebd., 68; Herv. i. O.).

»daß sich das dramatische Wort mit der Bedingung des Bildwechsels und des Bilderflusses nicht in Einklang bringen läßt, weil die reale Zeit des Wortes der filmischen Zeit der Montage widerstreitet. Darum hat das Sprechen als dramatischer Ausdruck dem Film nicht nur ›bisher‹ mehr Schwierigkeiten und wenigen Gewinn gebracht [...], sondern es ist der Bildkunst des Films wesensfremd.« (Ebd., 72)

Jedoch:

»Das Wort nimmt im Tonfilm nicht, wie auf der Bühne und im Hörspiel, die beherrschende Stellung ein – außer im reinen Sprechfilm, den wir aber von der Filmkunst ausgeschlossen wissen –, sondern es ist ein akustisches Ausdrucksmittel und als solches den beiden anderen des Geräusches und der Musik durchaus gleich-, nicht überwertig.« (Ebd.)

Rehlinger muss an dieser Stelle also insofern einen argumentativ schmalen Grat beschreiten, als er nicht – wie die von ihm kritisierten Autor:innen – den Film als durch das Wort grundsätzlich zerstört definieren kann; stattdessen verweist er auf die einzigartig dezentrierte »Stellung« des Wortes innerhalb der Filmwelt. Er stimmt Arnheim also insofern zu, dass »Sprechfilm« niemals »Filmkunst« sein kann (ebd.), aber dennoch sei das Wort als »akustisches Ausdrucksmittel«, wie jedes andere, Musik und Geräusch, schlicht ebenbürtig an die Seite zu stellen. Während bei Arnheim also das Wort als unumgänglich anthropozentrisch definiert wird, schlägt Rehlinger vor, im Tonfilm die erste künstlerische Apparatur zu sehen, in der das Menschenwort gerade nicht anthropozentrisch aufgeladen, sondern dezentrierend und egalisiert zur Dingwelt darstellbar wird. Auch bei Rehlinger wird somit die Stimme zum Geräusch und zur Melodie, die weniger in ihrem inhaltlichen Informations- oder ästhetischen Ausdrucksgehalt, denn in ihrer materiell-klanglichen Präsenz Relevanz entfaltet. So werde z. B. »[d]as kaum Auszusprechende, halb gesagt, Geflüsterte, alles, was außerhalb des sinnvollen Satzgefüges liegt, [...] im Tonfilm bedeutungsvoll« (ebd., 73); und »[e]s kommt hier viel weniger auf das Was des Gesprochenen an als auf das Wie und darauf, ob überhaupt gesprochen wird« (ebd., 73).³³⁷ Sprache würde somit »nicht als Rede-

337 Dies führt er auch noch einmal dezidiert in Bezug auf die Frage der Synchronisation des Films für internationale Märkte aus. Dabei nennt er mit mit »Dubben«, »Spruchbändern«, dem »Dreh in verschiedenen Versionen« und dem »sparsamen Einsatz des Wortes«, sodass dieses »auch bei Nicht-Übersetzung oder Untertitelung verständlich« sei, vier Verfahren, mit denen diese Internationalisierbarkeit des Films herstellbar sei (Rehlinger 1938, 74). Von diesen Vieren sei das »Dubben« als »Fälschung der darstellerischen Leistung und des Originaltons [grundsätzlich abzulehnen]«, auch, weil »selbst bei größter Sorgfalt niemals die künstlerische Einheit der ursprünglichen Fassung [erreicht]« werden könne (ebd.). Untertitel aber würden ebenfalls zu sehr ablenken und die polyglotten Sprachversionen (siehe auch Wahl 2005) wären nicht ver-

inhalt, sondern als Ausdruck, Handlungsmotor, Farbakzent, ‚Lichtfleck‘ (ebd., 75) denkbar, die »Worte als Streiflichter in das optische geschehen hinein« beschreibbar (ebd.). Rehlinger bemüht hier also zunehmend (licht)bildlich verdinglichende Metaphern, um das Wort aus seiner menschlich-geistigen Zentriertheit zu lösen. Und genau damit »unterscheidet sich der Wortgebrauch im Tonfilm von der Wortkunst der Dichtung, der Bühne und des Hörspiels« (ebd.), so stellt Rehlinger fest.

In Bezug auf das Verhältnis von Stimme und Musik werden diese Argumente noch einmal im Hinblick auf die Dinglichkeit des Menschen im Film und den Vorrang des Bildes bzw. der Montage vor den erweiterten Bedeutungen des Wortes verdichtet:

»Die Grundlage des Tonfilms aber ist nicht der, wenn auch gefilmte, so doch singende Mensch, der Mensch ist es überhaupt nicht, sondern: die Grundlage des Tonfilms und damit der Filmkunst ist die Montage, d.h. in erweiterter Bedeutung die Auswertung der filmischen Formungsmittel in dem Sichtbar- und Hörbarmachen der gesamten Wirklichkeit, zu der neben den Dingen auch der Mensch zählt. Diese wesensverschiedene Stellung des Menschen macht den Unterschied des Tonfilms zur Bühnenoper aus, nicht die technisch veränderte Erscheinungsform desselben Menschen.« (Ebd., 87)

Mit all diesen anthropo-dezentrierenden Argumenten ist man der Frage, was in Rehlingers Hinsicht als »filmisch« zu bezeichnen ist, somit bereits sehr nahe. Auch wenn er in weiten Teilen seines Buches dennoch (anthropozentrierend, anthropomorphozentrierend und anthropophonozentrierend) die Menschendarstellung im Film verhandelt, wird genau in diesen argumentativen Bewegungen und Differenzsetzungen erkennbar, wie sich das Audioviduum zunächst (vor allem im Hinblick auf seine Relevanz) von der Vorstellung des ‚tatsächlichen‘ Menschen löst, um es dann vollständig hinter sich lassen zu können. In seinem Fazit fasst Rehlinger diese Idee noch einmal zusammen, um schließlich zu einer Art post-anthropozentrischen Definition des »Filmischen« zu gelangen. Das Filmische könne dabei ausschließlich aus den »Formmitteln des Films« selbst abgeleitet werden – und eben nicht aus seinen Inhalten und dem, was ihm zugrunde liege; sei dies nun eine künstlich geschaffene, geschauspielerte oder realweltliche

gleichbar, weshalb eigentlich nur die vierte Variante übrig bleibe. Der interessante Nebeneffekt sei aber so oder so, dass »[g]erade fremdsprachige Tonfilme zur Klärung des Wortproblems im Film bei[tragen]. Die Übersetzung der Gespräche ins Deutsche – wie umgekehrt aus dem Deutschen in eine andere Sprache – ist ein vortrefflicher Prüfstein für Menge, Wert oder Sinnlosigkeit des Gesprochenen«, so Rehlinger (1938, 74).

Szenerie. Und darin äußert sich somit eine Verabschiedung des *Menschen als Motiv des Films als Motiv von Medientheorie*:

»Der Begriff filmisch muß etwas bezeichnen, das die Erscheinungsform Film zur Voraussetzung hat und nur mit ihr in organischer Verbindung gedacht werden kann [...]. Der Begriff filmisch kann also seiner Bestimmung nach durch keinen anderen Begriff ersetzt werden, weder durch dramatisch noch episch noch pantomimisch noch mimisch. Wäre filmisch = mimisch, dann unterschiede der Film sich in nichts von der Pantomime oder dem abgelösten mimisch-gestischen teil eines Schauspiels. Gegen beide aber haben wir die Grenzlinien gezogen. Trotz allem könnte die Gleichung ein willkommener Behelf sein, wenn die Mienen- und Gebärden sprache des Schauspielers im Film alles bedeutete. Aber selbst im Stummfilm ist das nicht der Fall, denn der Schauspieler bildet – wir sagten es schon – nur einen Teil des Filmkunstwerks. Wie sollte man dann die Teile nennen, in denen nicht ›gemimt‹ wird? Wären sie nicht filmisch? Es sei denn, man faßte unter den Begriff Mimik (der ja doch nichts anderes bedeutet als die Kunst der Mienen- und Gebärden sprache) auch das bloße Aussehen aller sichtbaren Dinge, ihr ›Gesicht‹ im weitesten Sinne. Daß damit der Begriff verwässert würde, liegt auf der Hand. Weiter umfaßt die Arbeit des Filmkünstlers wesentlich mehr als die Herausstellung des Mimischen. Der ein enges Gebiet absteckende Begriff mimisch, den man mit größerer Berechtigung dem Begriff dramatisch gleichgesetzt hat, ist demnach für den höheren Begriff filmisch ein unzureichender Ersatz. Endgültig bleibt uns kein anderer Weg als der unmittelbar von den Formmitteln des Films auszugehen. Diese Formmittel sind: Bildstreifen, Bild- und Tonkamera.« (Ebd., 98f.)

Der Mensch – als Motiv – wird somit aus der Legitimation des Films bzw. des Filmischen getilgt zugunsten einer individuellen Selbstbezüglichkeit des Films. In der Berufung auf seine ureigensten technischen »Formmittel« allein gelangt der Film so aus sich selbst heraus zu einer Bestimmung – und das ohne ›Hilfe‹ des Menschen. Wenn in Bezug auf die Definition des menschlichen Individuums an dessen Beginn immer der Moment der Selbsterkenntnis (als Individuum) steht, so lässt sich Rehlingers Beschreibung als eine ebensolche, endgültige Individuation des Films lesen.

Und nur wenn auch der Tonfilm dieser Rahmung des Filmischen als Filmisches Rechnung trägt und sich nicht (wie etwa im Sprechfilm) durch andere (menschliche) Individualitäten irritieren lässt, kann er, laut Rehlinger, als filmische Kunstform Anerkennung finden. Dies sei gesichert durch die Verschaltung von bildlicher und akustischer Dingwelt jenseits (anderweitig) zentrierender Logiken:

»Nur soweit die sichtbaren Dinge dieser Welt, gleichviel ob bewegt oder unbewegt, ob Mensch oder Tier, organische oder anorganische Natur, auf diese Weise von der Kamera erfaßt werden, erhalten sie den filmischen Ausdruck. Nur soweit Wort, Geräusch und Musik – unter sich von wertgleicher Bedeutsamkeit – als zwangsläufig-unmittelbare, begleitende oder kontrastierende Bindung an die optischen Gesetze sich zu erkennen geben, können wir von einem filmischen Gebrauch dieser Tonmittel sprechen.« (Ebd. 101)

Die stärkste Ausprägung erhalte diese Egalisierung mit dem abstrakten Trickfilm, in dem »das Filmische rein aktualisiert« sei (ebd.). Abseits davon gäbe es aber noch zahlreiche andere Arten und Weisen zu entdecken, wie das Filmische als solches konkret einzulösen sei. Dabei erhält in der Ablösung vom Audioviduum und damit vom Menschen die Individualität des Films zunehmend Kontur:

»In welcher Weise aber die Vielfalt der Einzelemente zur künstlerischen Einheit ‚Film‘ zusammenzuschließen und wie diese Einheit begrifflich zu fassen sei, ist damit noch nicht geklärt. Es darf indessen erwartet werden, daß diese Problematik nur in Fortsetzung des von uns begangenen Weges gelöst wird, da sich alle anderen Versuche als unbrauchbar herausstellten. Erst wenn die Gesetze des filmischen Aufbaus im Einzelfilmkunstwerk festgelegt sind, das sich als der auf der Leinwand sichtbar werdende Bildstreifen darstellt, lassen sie sich gültig beweisen. Nur aus dem Zusammenklang der zur Ganzheit Film gebundenen Einzelheiten können dann die Bedingungen der kontrapunktischen Führung, der leitmotivischen Verknüpfung, der Architektonik, Ornamentik, der besonderen filmischen Zeit und des besonderen filmischen Raumes abgelesen werden.« (Ebd., 102)

Nicht mehr der menschliche Körper tritt also auf der Leinwand hervor, um das Medium zu definieren, sondern »der auf der Leinwand sichtbar werdende Bildstreifen« als eigene, individuelle Ganzheit; nicht mehr die menschliche Stimme bestimmt den Ton des Films, sondern der »Zusammenklang der zur Ganzheit Film gebundenen Einzelheiten« (ebd.). Der Film wird zum Individuum; das Audioviduum verschwindet – wie ein Gesicht im Sand des Meeresufers. Somit mündet die Emergenz des Audiovidiums im Tonfilm bei Rehlinger in dessen finaler Irrelevanz.

3.3.4 | ZWISCHENFAZIT: Emergenz und Irrelevanz des Audiovidiums

Der Weg des Audiovidiums von seiner zunächst diskursiven, dann zusätzlich materiellen Emergenz, über seine theoriehistorische Zentralstellung, bis hin zur finalen Erklärung seiner Irrelevanz, hat also gezeigt, dass sich die Medientheoriegeschichte des Audiovisuellen entlang dieser argumentativen Denkfigur gewinnbringend beobachten lässt.

Verortet man nun die skizzierten Tonfilmdebatten der späten 1920er und 1930er Jahre im Kontext der zuvor nachgezeichneten Stummfilm- und Rundfunkdiskurse, so sind bemerkenswerte Parallelen, aber auch Wendungen zu erkennen. So bleibt es dabei, dass das Menschenmotiv des Films (und damit in Bezug auf den Tonfilm das ›tatsächliche‹ Audioviduum) eine zentrale Rolle im Versuch einer Bestimmung des Tonfilms als Medium einnimmt. In Bezug auf Stummfilm und Rundfunk wird das Menschenmotiv zunächst als allein ideelle Denk- und Vergleichsfigur herangezogen, mit der es möglich ist die ›Mängel‹ zu identifizieren, die das Medium z. B. im Vergleich zu anderen Medien(menschen) oder zu realweltlich gedachten, menschlichen Individuen aufweist, bevor es dann bewusst zurückgelassen werden kann, zugunsten quasi-separierter Video- und Audio-Individuen, die zur eigentlichen Bestimmung des jeweiligen Mediums (als Kunst) werden. Innerhalb der Stummfilmtheorie bereitet sich dabei allerdings spätestens mit Balázs (2001 [1924]) auch schon der Gedanke vor, dass die Menschen(motiv)zentriertheit des Films einer Dingzentriertheit weichen könnte, indem es am Schluss nicht mehr nur um eine »Sichtbarkeit« des Menschen allein geht, sondern auch um die individuelle Belebung der Landschaften und Dinge, in die er sich über die Fläche der Leinwand nahtlos eingliedern lässt. Bei Balázs ist dabei schön zu sehen, wie seine Auseinandersetzung mit dem Menschlichen des Films oszilliert zwischen einer Anthropomorphisierung der Dingwelt und einer Verdinglichung des Menschen, auf deren offener Grenze der filmische Anthropomorphozentrismus seines Ansatzes zu verorten ist. Bei Vertov (2001 [1924]) oder auch Tynjanov (2001 [1924]) sind schließlich der filmische Anthropozentrismus und Individualismus bereits tendenziell überholt und auch der Anthropomorphozentrismus (also die Fokussierung auf das vom ›Realmenschen‹ abgelöste, alleinig mediale Menschenmotiv) wird relativiert durch die Bedeutung der Dinge, sodass ihre Theorien – wie man spätestens in Bezug auf Arnheims und Rehlingers Tonfilmkonzepte feststellt – als Anfang einer medientheoretischen Umorientierung (noch stärker hin zum Medium selbst und weg vom Menschen[motiv]) gewertet werden können. Bei Arnheim und Rehlinger verdichten sich diese Argumente, doch auch in den Diskursen zuvor ist zunehmend spürbar, dass der Mensch nach und nach nur noch als Beispiel unter anderen dient und nicht mehr als zentralster Anhaltspunkt wie zuvor. Damit bestätigt sich die anfänglich entwickelte These, dass vor allem in den Vor- und Frühphasen von Medientheoriebildung das Menschenmotiv – und

somit in der Vor- und Frühphase der audiovisuellen Medien: das Audioviduum – eine zentrale Rolle spielt, um eben diesen Medien (im Angesicht menschlicher Vorstellungen von Individualität) selbst nach und nach Individualität zu verleihen.

Auch in Bezug auf die Rundfunkdebatten sind ähnliche Verhältnisbestimmungen möglich. Hier sind teilweise allerdings weniger Kontinuitäten, als Kontraste festzustellen: So wird der Tonfilm zwar vielfach mit dem Hörspiel verglichen, aber in der Regel nur, um deutlich zu machen, dass die Sprache weiterhin im Funk aufgehoben sein soll, während das Wort in seiner Funktion innerhalb des Tonfilms schwieriger zu verorten ist. Die Rundfunktheorie der 1920er und 1930er Jahre, die damit ja auch zeitlich näher an den Tonfilm und seine Etablierung angegliedert ist, macht es – so könnte man vermuten – noch wichtiger, dass jedes Medium sein Gebiet absteckt. So gibt es zwar auch in der Hörspiel- und Funktheorie Bestrebungen ein reines Klang- und Geräuschradio zu entwickeln; gerade im Kontext der NS-Propaganda bleibt aber die Stimme ein zu wichtiges Instrument, um es anderen Medien zu überlassen.

Wie somit gezeigt werden konnte, wird der Mensch innerhalb der Tonfilmdebatten nicht mehr ganz so intensiv ins Zentrum der neuen Kunst gestellt, wie dies innerhalb von Stummfilm- und Radiodiskurs der Fall war. Dies hängt vor allem damit zusammen, dass die Kunstfertigkeit der vorherigen (oder auch noch parallel) ›neuen‹ Medien jeweils schon zu sehr an ihrer medialen Eindimensionalität festgemacht wurde, sodass ihre neuartige Kopplung diese Kunstanansprüche zu stark revidieren würde. Stattdessen wird in vielen Fällen gegen ihre Verbindung andiskutiert oder aber ihre sinnvolle Vereinigung weiterhin als Fernziel entworfen. Mit dem tonfilmischen Audioviduum erfährt die Mensch-Medien-Kopplung somit eine erneute Wendung – vom Menschen weg zum Ding, mit Stimmen als Geräusch und Körpern als Formen, über die sich jeweils der Mensch endgültig im Medium aufzulösen scheint, um dessen Individuation abzuschließen.

