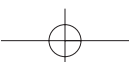
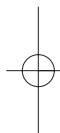


## MEDIEN/STIMMEN



**Mediologie**

Band 9

Eine Schriftenreihe des Kulturwissenschaftlichen Forschungskollegs

»Medien und kulturelle Kommunikation«

Herausgegeben von Ludwig Jäger

Cornelia Epping-Jäger / Erika Linz

# MEDIEN/STIMMEN

DuMont

Diese Publikation ist im Sonderforschungsbereich/Kulturwissenschaftlichen  
Forschungskolleg 427 »Medien und kulturelle Kommunikation«, Köln, entstanden  
und wurde auf seine Veranlassung unter Verwendung der ihm von der Deutschen  
Forschungsgemeinschaft zur Verfügung gestellten Mittel gedruckt.  
Der Band wurde durch das Ministerium für Schule, Wissenschaft und Forschung  
des Landes Nordrhein-Westfalen gefördert.

**Erste Auflage 2003**

© 2003 DuMont Literatur und Kunst Verlag, Köln

Alle Rechte vorbehalten

Ausstattung und Umschlag: Groothuis, Lohfert, Consorten (Hamburg)

Gesetzt aus der DTL Documenta und der DIN Mittelschrift

Gedruckt auf säurefreiem und chlorfrei gebleichtem Papier

Satz: Greiner & Reichel, Köln

Druck und Verarbeitung: B. o. s. s Druck und Medien GmbH, Kleve

Printed in Germany

ISBN 3-8321-7855-4

## INHALTSVERZEICHNIS

<b>Cornelia Epping-Jäger / Erika Linz</b> Einleitung	7
<b>I. STIMM-AKT</b>	
<b>Bernhard Waldenfels</b> Stimme am Leitfaden des Leibes	19
<b>Leander Scholz</b> Tierstimme/Menschenstimme: Medien der Kognition	36
<b>Erika Linz</b> Die Reflexivität der Stimme	50
<b>Sybille Krämer</b> Negative Semiologie der Stimme	65
<b>II. STIMM-WECHSEL</b>	
<b>Heike Behrend</b> Geisterstimmen in Afrika: Die Stimme als Medium der Fremdpräsenz	85
<b>Cornelia Epping-Jäger</b> »Eine einzige jubelnde Stimme«. Zur Etablierung des Dispositivs Laut/Sprecher in der politischen Kommunikation des National- sozialismus	100

<b>Michel Chion</b> Mabuse – Magie und Kräfte des <i>acousmètre</i> . Auszüge aus <i>Die Stimme im Kino</i>	124
<b>Gereon Blaseio</b> <i>Gendered voices</i> in der Filmsynchronisation. FIRST BLOOD versus RAMBO	160
<b>Ines Steiner</b> The voice behind the curtain. Zur Inszenierung der Stimme in SINGIN' IN THE RAIN	176
<b>III. STIMM-ART</b>	
<b>Daniel Gethmann</b> Das Zittern der Luft. Die Erfindung der mechanischen Stimme	211
<b>Marcus Erbe</b> Sprachliche und sprachlose Stimmen in der Elektroakustischen Musik	235
<b>Jin Hyun Kim</b> Die Singstimme als Ausdruckszeichen. Zur medialen Funktion der Stimme in der Musik	250
<b>Doris Kolesch</b> Die Spur der Stimme. Überlegungen zu einer performativen Ästhetik	267
<b>Anthony Moore / Siegfried Zielinski</b> Vox per →sonare. Ein Vortragsduett in acht Teilen	282
<b>Bildnachweise</b>	299
<b>Autorenverzeichnis</b>	301

Cornelia Epping-Jäger / Erika Linz

## EINLEITUNG

Die Stimme hat in den gegenwärtigen medientheoretischen Diskursen – insbesondere im Kontext der Performativitäts-Debatte – eine neue theoretische Aufmerksamkeit erlangt, die sie aufgrund verschiedener epistemisch-diskursiver Bewegungen lange verloren zu haben schien.<sup>1</sup> Versucht man den Gründen für die weitgehende Unsichtbarkeit der Stimme nachzugehen, so lassen sich verschiedene Ursachen namhaft machen:

Zunächst hatte gerade die Krise der »literacy hypothesis« auch zu einer Abwertung des Problemfeldes der mündlichen Rede und damit der Medialität der Stimme geführt. Dieser Umstand verdankte sich der spezifischen Rhetorik dieses Diskurses: Vor dem Hintergrund der Annahme, dass »die westliche Kultur in all ihren Erscheinungsformen [...] letztlich aus dem Geist der [Alphabet-]Schrift geboren wurde«,<sup>2</sup> argumentierte man lange Zeit strikt oppositiv. Auf der einen Seite verortete man die oralen Kulturen, charakterisiert durch »verkörperte Multimedialität, Zeremonie, Kollektivität«<sup>3</sup> sowie durch eine hochgradig transitorische Mündlichkeit, in welcher Kommunikation sich durch die Unmittelbarkeit auszeichnen sollte, in der sie sich vollziehe, – als »Anwesenheit und Ereignis«. Auf der anderen Seite stand die literale Kultur, »die Schreiber und Leser radikal vereinzelt«, in der Schrift »Reflexion und Distanz, Ich-Bildung und Individualisierung« erzeuge, womit allererst »im Menschen installiert [werde], was uns Subjekt heißt.«<sup>4</sup>

Nun konnten aber gerade ethnologische Studien zu nicht-europäischen Kulturen zeigen, dass viele der als spezifisch europäisch-literal reklamierten sprachlichen und kognitiven Eigenschaften auch die Diskurse weitgehend »mündlicher« Kulturen prägen.<sup>5</sup> Im Horizont dieser Forschungen erschien die »literacy hypothesis« ihren Kritikern als »great divide«<sup>6</sup>, und neben ihrer eurozentrischen Ausrichtung kritisierte man vor allem drei Momente: erstens die quasi ontologische Trennung von Oralität und Literalität, zweitens die aus dieser Dichotomie sich ableitende Tendenz, alle nicht voll literalisierten Kulturen als »defizitär« zu klassifizieren, und drittens schließlich die, an Walter Ongs Ausführungen orientierten, normativen Stereotypisierungen des Mündlichen und seiner strukturellen Eigenschaften. In dem Maße, in dem sich die Bestimmung der »Mündlichkeit« als »transitorisch«, »situativ«, »aktualesemantisch« und »syntagmatisch«<sup>7</sup> in dieser Generalisierung zunehmend als obsolet erwies, minimierte sich das medienhistorische und medientheoretische Interesse am kommunikativen Raum der Stimme.

Neben den Folgen der »great divide«-Debatte war es weiterhin vor allem Derridas Kritik des Logophonozentrismus, die zu einer skripturalistischen Verschiebung der medientheoretischen Aufmerksamkeit führte. In »Die Stimme und das Phänomen«<sup>8</sup> hatte Derrida den Husserlschen Versuch wirkungsmächtig dekonstruiert, über eine von jeder Signifikantenspur losgelösten inneren Stimme an den Nullpunkt eines sich selbst präsenten, logozentrischen Bewusstseins zu gelangen, in dem das Subjekt sich selbst begründe. Als absolute Präsenz, so Derrida, vermag der immer schon temporalisierte *Solilog* insofern nicht zu erscheinen, als er bereits – als »innere Stimme« in sich gespalten – durch die Spur des Signifikanten gezeichnet ist und damit das scheinbar ausgegrenzte Moment der Externalität immer schon in sich trägt. Gleichwohl ist es nicht die Externalität der phänomenalen Stimme, die sich als Signifikantenspur in das Bewusstsein einschreibt, sondern die Schrift, die als Urschrift jeder Rede zugrunde liegt und ihr vorausgeht. »Die Stimme, die in der Sprache zum Sprechen geworden ist, ist nicht die Sprache.«<sup>9</sup> Zwar realisiert sich die Stimme nur im Sprechen, die sprachlichen Zeichen aber gehen ihr, bevor sie sie spricht, als Schrift voraus. »Die Exteriorität des Signifikanten«, so formuliert Derrida, »ist die Exteriorität der Schrift im allgemeinen.«<sup>10</sup>

Ein dritter Umstand verstärkte die skripturalistische Entmächtigung der Stimme: die Tendenz nämlich der Medientheorie, den Raum des Medialen allererst jenseits der Stimme mit der Schrift und den technischen Medien beginnen zu lassen. Der skripturale Blick auf Stimme und Sprache erklärte diese geradezu zu Epiphänomenen der Schrift, während die Schrift selbst sich zum »Paradigma des Medialen« schlechthin erhob. Schrift und Medien werden in dieser Hinsicht, um mit Ludwig Jäger zu sprechen, »zu »Kunststücken«, denen die »Natürlichkeit« einer biologisch basierten und insofern gleichsam prämedialen Sprache« entgegengesetzt wird.<sup>11</sup> Die Tendenz, Medialität mit »Künstlichkeit« und »Technologie« kurzzuschließen, zeigt sich auch, wenn die humane Stimme in den Mittelpunkt einer technischen Medienaufmerksamkeit tritt: »Die Standards der technischen Medien haben mit dem sogenannten Menschen und seinem *sensus communis* nichts zu schaffen. [...] Um Stimmen und Geräusche auf Zelluloid zu bannen« – so formulierte jüngst Friedrich Kittler – »mußte die Metaphysik des Sich-selbst-Hörens gründlich vergessen, nämlich durch Schaltungstechnik ersetzt werden.«<sup>12</sup>

Schließlich kann sicher auch in der Medialitätsvergessenheit der Sprachtheorie eine weitere Ursache des medientheoretischen Niedergangs der Aufmerksamkeit für das Medium Stimme gesehen werden. Der psychologische und sprachwissenschaftliche Kognitivismus ist nicht nur der Auffassung, dass sprachliche Strukturen gegenüber gedanklichen durch Nachträglichkeit gekennzeichnet sind, sondern er geht auch von einer prinzipiellen Amedialität sprachlichen Wis-



sens aus. Die Stimme überlebt nur als phonologische Struktur im Jenseits des Mentalen. Aus einer solchen Kompetenzfixierung wiederum ergibt sich die durch den Mainstream-Kognitivismus vollzogene generelle Abwertung des Performativen und die damit einhergehende grundsätzliche Marginalisierung der Performativität der Stimme.

Machen wir eine letzte Ursache für das Unsichtbarwerden der Stimme namhaft: Die Gefahr ihrer Marginalisierung nistet nämlich selbst dort, wo sie in jüngerer Zeit explizit thematisch wurde, – etwa in den Diskursen der empirischen Psychologie und der Medizin. Vertritt man wie etwa Horst Gundermann die These, dass die Stimme »tiefer als die Sprache im endothyemen Grund des Menschen eingebettet«<sup>13</sup> liege, so ist das in einer doppelten Hinsicht problematisch: Zum einen ersteht damit das – von Derrida ja durchaus mit Recht nachhaltig dekonstruierte – Modell der abendländischen Metaphysik in einem platonischen Gestus insofern wieder auf, als die stimmliche Verlautbarung hier gleichsam als authentischste Sinnäußerung einer Seele verstanden wird.<sup>14</sup> Das Argument von der endothyemen Einbettung der Stimme verweist auf eine Tradition, der die »vokale Geste« als »entlarvender und beharrlicher« gilt als alle anderen Ausdrucksbewegungen, denn: »die Stimme lügt nicht.«<sup>15</sup> Bis in die Antike lässt sich dieses Verständnis der Stimme als »index mentis, als hörbares Zeichen von Gemüts- und Bewußtseinszuständen«<sup>16</sup> zurückverfolgen. Wie prekär dieses Argument ist, wird nicht erst in Zeiten digitaler Stimmbearbeitung offensichtlich: Denn bereits für die technisch nicht transformierte ›Normalstimme‹ gilt, dass es keine gesicherten Merkmale physiognomischer Stimmerkennung gibt, und dass auch keine »Sprechmerkmale existieren, die es erlaubten, direkt auf Unwahrhaftigkeit zu schließen.«<sup>17</sup> Selbst wenn es in forensischen Untersuchungen gelingt, mittels des Voice Print – des unverwechselbaren Stimmabdrucks – Täter zu identifizieren, so ist die Wahrheit von Aussagen auf diesem Wege gleichwohl nicht feststellbar. Darüber hinaus postuliert die These von der ursprünglichen Wahrhaftigkeit der vorsemantischen Stimme auch eine Nachträglichkeit sprachlichen Sinnausdrucks gegenüber der »Ausdruckswelt der Stimme«. Dass bereits die ›natürliche‹ Stimme medial konstruiert und hochgradig von kulturellen Artefakten und Inszenierungen bestimmt ist, gerät hier gar nicht in den Blick.

Für die gegenwärtige kulturwissenschaftliche Wiederentdeckung der Stimme ist es nun charakteristisch, daß sie den Stimm-Diskurs aus den oben skizzierten Abwertungs-Szenarien löst, um ihn in den Horizont medientheoretischer Überlegungen zu stellen. Vor diesem Hintergrund ist es das besondere Interesse des Bandes *Medien/Stimmen*, einerseits die Gefahren einer Anthropologisierung der

Debatte zu meiden, gleichwohl aber andererseits die vorprädikativen phänomenalen Bedingungen der Stimme theoretisch in Rechnung zu stellen. Auch wenn die Materialität der Stimme durchaus als Spur eines an der sprachlichen Sinn- und Selbstbildung beteiligten »Körper/Leibes« und dessen »Urskriptes«<sup>18</sup> verstanden werden kann, ist sie zugleich doch auch in die kommunikativen und technischen Formate der medialen Prozessierung der Stimme eingewoben. Bereits Hegel, dem die menschliche Stimme »als die Hauptweise« galt, »wie der Mensch sein Inneres kundtut«<sup>19</sup> und der sie, neben dem Lachen und Weinen, »noch ehe sie zur Sprache wird« zu jenen Entäußerungsformen zählte, die der »Wegschaffung der innerlichen Empfindungen dient«,<sup>20</sup> lässt sie nicht auf diesem anthropologischen »Standpunkte der natürlichen Seele« stehen, sondern schreibt ihr letztlich einen mediatisierten Status zu: »Der Mensch aber [...] schafft die artikulierte Sprache, durch welche die innerlichen Empfindungen zu Worte kommen.«<sup>21</sup> Und auch Bühler, der der *vox humana* in seinem Organonmodell mit den Funktionen ›Ausdruck‹ und ›Appell‹ neben der Darstellungsfunktion zentrale Orte anwies, ließ keinen Zweifel an ihrem medialen Status: »Das mediale Produkt des Lautes« tritt als Ausdruck und Appell in »je eine eigene Zeichenrelation« zu dem sich aussprechenden und andere adressierenden Sprachsubjekt.<sup>22</sup>

Die Stimme ist Teil des Körpers und als solche existiert sie immer nur eingebettet in symbolische Formen. »Vor der Trennung, in der Identität von Körper und Ton, gab es sie noch nicht, nach der Trennung in Verbindung von Ton und Symbol gibt es sie nicht mehr.«<sup>23</sup> Insbesondere aus dieser Perspektive öffnet sich der Raum für Thematisierungen der Stimme im Hinblick auf ihre sprachlichen bzw. sprachlich-technischen Mediatisierungen. Gleichwohl erheben gerade die im Horizont der neueren Performanzdebatte entfaltenen theoretischen Konzepte den Anspruch, eine Medialität der Körperstimme theoretisch auszufalten, die ihre Signifikanz nicht aus ihrer Bezogenheit auf die prädikative Semantik sprachlicher Prozesse bezieht. Nicht nur als Musical-, Synchron- und Lautsprecherstimme, sondern auch als Singstimme, Geisterstimme oder als Stimme des Leibes ist die *vox humana* immer schon Medien/Stimme: Eingeschrieben in die verschiedenen Formate ihrer anthropogenen und technischen Generierung organisiert sie jene kognitiv-semiotischen Prozesse, in denen sich die Diskursivität medialer Sprachspiele, ritueller Handlungen und ästhetischer Performanzen entfaltet.

Die Beiträge des vorliegenden Bandes folgen keinem integralen stimmtheoretischen Konzept, sondern sind bestimmt durch thematische Varianz und methodische Pluralität. Die methodische Spannweite reicht von empirischen bzw. an Einzelmedien orientierten Analysen über textliche ›Performances‹ bis hin zu theore-

tisch-systematischen Studien. Thematisch entfalten sich die Beiträge zwischen den Polen einer Analyse der Stimme im Horizont sprachlicher Prädikativität und technischer Mediatisierung einerseits sowie ihrer Fokussierung im vorprädikativen Raum leiblicher Bestimmtheit andererseits. Gemeinsam ist dabei allen Beiträgen die Frage nach der operativen Logik der medialen Inszenierung von Stimme, d. h. die Frage nach den medialen, technischen und anthropogenen Prozessierungsformen ihrer Performanz. Die Stimme, so scheint es, ist aufgrund ihrer fluiden Materialität in besonderem Maße ein performatives Medium, ein Medium, das sich nicht unabhängig von seinen Prozessierungsformen betrachten lässt. Bestimmend für die Beiträge des Bandes ist daher nicht mehr die lange vorherrschende Konzentration auf die Differenz von Stimme und Schrift, sondern die Analyse der Stimme im Dispositiv ihrer unterschiedlichen Generierungsbedingungen.

Die Sektion **Stimm-Akt** thematisiert die Funktionen der Stimme als Medium des Sprechens. In ihrem Zentrum steht die Frage nach der phonischen Materialität der Stimme und ihren Auswirkungen auf die Prozessualität des Sprechens. Alle Beiträge machen deutlich, dass sich eine solche Annäherung an die Stimme weder mit einer sprecherfixierten noch mit einer sprachfixierten Betrachtung begnügen darf. Sprechen ist trotz seiner linearen Sukzessivität ein mehrdimensionaler Prozess, in dem sich unterschiedliche semiotische und nicht-semiotische Ebenen überlagern. Eine allein auf den Sprecher konzentrierte Untersuchung der stimmlichen Artikulation übersieht zum einen, dass es sich auch beim monologischen Sprechen grundsätzlich um adressiertes Sprechen handelt, zum anderen, dass Sprechen immer auch Hören bedeutet und uns die Stimme, auch die eigene, zuerst als Fremdstimme begegnet.

Diese Interdependenz von Sprechen und Hören nimmt *Bernhard Waldenfels* zum Ausgangspunkt seines Beitrages, in dem er über eine phänomenologisch orientierte Analyse des Hörereignisses die Responsivität von Sprechen und Hören und deren Verankerung im »Urmedium« des Leibes aufzeigt. *Leander Scholz* legt über eine Analyse der Hegelschen Anthropologie frei, inwiefern sich die phonozentrische Privilegierung der Stimme wesentlich dem Paradigma des Lebendigen verdankt. Sein Beitrag macht deutlich, dass die Stimme in der Tradition der abendländischen Metaphysik nicht nur als »Medium des Lebendigen«, sondern ebenso als »lebendiges Medium« konzeptualisiert wird. Der Beitrag von *Erika Linz* diskutiert in einer Verbindung von zeichenphilosophischer Reflexion und psycholinguistischer Analyse der kinästhetischen, motorischen und auditiven Rückkopplungsschleifen im ›Sich-sprechen-Hören‹ die epistemische Spurfunktion der Stimme. *Sybille Krämer* führt in ihrem Beitrag vor, inwiefern sich die analo-

gisch-indexikalische und mehr noch die musikalische Dimension der Stimme im Sprechen einer zeichentheoretischen Erfassung entzieht und plädiert in diesem Sinne für eine sprachtheoretische Rehabilitation der Musikalität des Sprechens.

Die Sektion **Stimm-Wechsel** thematisiert die Einheit von Stimme und Körper in verschiedenen Facetten ihrer medialen Dissoziierung anhand des Zusammenspiels von technischen Formaten und medialen Inszenierungen. Im Vordergrund steht hier die Frage, wie die Trennung der Stimme vom Körper in unterschiedlichen medialen Dispositiven auf das Rezeptionsverhalten der Hörer rückwirkt und welche Effekte die differenten Formen ihrer Aufspaltung für die Konstruktion personaler und kollektiver Identitäten haben. Die Beiträge zeigen auf, dass die Veränderungen in der Eigen- und Fremdrezeption der Stimme durch ihre Entkopplung von einem sprechenden Körper nicht nur von den technischen Formaten abhängen, sondern weit mehr noch durch die jeweiligen medialen Inszenierungen und die damit einhergehenden Rezeptionserwartungen geprägt sind. Die Beispiele differenter kultureller Formen des Stimmwechsels eröffnen darüber hinaus einen neuen Blick auf die enge Bindung westlicher Personalitätskonzepte an die Vorstellung der *einen* personalen Körperstimme.

*Heike Behrend* untersucht die Stimmen afrikanischer Geistmedien im Kontext einer ›elektrischen‹ Kosmologie der Kraft. In einer Gegenüberstellung paganer und christlich geprägter Formen der Geistbesessenheit führt sie vor, wie sich die verschiedenen Formen der Mehrstimmigkeit von Geistbesessenen nicht nur in der stimmlichen Performanz niederschlagen, sondern auch Auswirkungen auf die Konzeptualisierungen der sozialen Person und des Körpers haben. *Cornelia Epping-Jäger* zeigt, dass in Deutschland zwischen 1932 und 1936 ein medialer Großversuch stattfand, in dessen Verlauf das Dispositiv Laut/Sprecher als eine Diskursapparatur etabliert wurde, die die Verinnerlichung der Macht als akustische Erlebnisform der Volksgemeinschaft organisierte. Der Beitrag von *Michel Chion* ist eine Übersetzung des ersten Teils seines bislang nur auf Englisch und Französisch vorliegenden Buchs »La voix au cinéma«, das bis heute als das Referenzwerk zur Position der Stimme im Kino gelten kann. Anhand des Begriffs des *acousmètre* diskutiert Chion die unterschiedlichen Wirkungen der körper- und gesichtslosen Stimme je nach medialem Kontext, in dem sie uns begegnet: Man sieht anderes, wenn man zugleich hört, und man hört anderes, wenn man zugleich sieht. *Gereon Blaseio* verdeutlicht am Beispiel der Transformation von *FIRST BLOOD* zur deutschsprachigen Fassung *RAMBO* die kulturelle Eigenständigkeit und Wirkmächtigkeit des synchronisierten Films. Seine Analyse führt die deutsche Synchronisation als Entstehungsprozeß vor, bei dem uneindeutige Gen-

der-Semantisierungen und politisch-kulturelle Differenzen ›behooben‹ werden. *Ines Steiner* fragt in einer filmischen Analyse des Musicals SINGIN' IN THE RAIN nach der Funktionsstelle der in diesem Film explizit ausgestellten Inszenierungsformen der Sing- und Sprechstimme. Sie zeigt, dass das Musical in einer selbstreflexiven Geste den Medienumbruch vom elaborierten Kode des Speechless Film zum Tonfilm thematisiert.

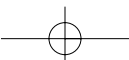
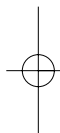
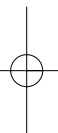
Die Sektion **Stimm-Art** thematisiert die Beziehung zwischen Körper und Stimme in ihren technisch-medialen Transformationen und ästhetischen Inszenierungen. Die Diskussion fokussiert hier insbesondere die artikulatorisch-produktive und die akustisch-rezeptive Dimension der Stimme als Gegenstand technischer Reproduktion und künstlerischer Performanz. Dabei wird die Frage nach der spezifischen Medialität der Stimme und ihrer körperlich-emotiven Expressivität von zwei gegensätzlichen Blickrichtungen aus beleuchtet. Der Rekonstruktion von Versuchen einer technischen Simulation der Stimme sowie ihrer musikalischen Transformation in entpersonifizierte, von körperlichen Spuren befreite Klänge stehen Analysen von musikalischen und theatralischen Ausdrucksformen gegenüber, in denen gerade die Präsenz des Körpers und der Person in der Stimme performativ in Szene gesetzt werden.

*Daniel Gethmann* setzt sich mit den Entwürfen unterschiedlicher Konzepte der Stimme und ihrer Werkzeuge im Nachbau der menschlichen Sprechfähigkeit auseinander. Anhand einer Analyse der mechanischen Stimmnachbildungen in den Sprechmaschinen von Kratzenstein, Kempelen u. a. kann er zeigen, dass die Konstruktion dieser Maschinen einen Bruch in der Epistemologie um 1800 verdeutlicht. *Jin Hyun Kim* fragt nach der Möglichkeit einer Kodierung von Emotionen durch die Singstimme und damit nach einer medialen Funktion von Stimme und Musik, die sich jenseits einer an der Darstellungsfunktion orientierten Semantik zeigt. *Marcus Erbe* stellt in seinem Beitrag dar, wie Strategien einer Entsemantisierung und Entpersonifizierung der sprachlichen Stimme durch ihre musikalische Transformation in der elektronischen Musik den Raum für eine textfreie Emotionssemantik eröffnen. *Doris Kolesch* diskutiert das Konzept einer performativen Ästhetik der Stimme: Aufgrund ihrer Nicht-Verortbarkeit, ihrer Atopie, wird die Stimme als sinnlich-sensuelles Organ rekonzeptualisiert, das mehr und anderes darstellt als ein bloßes Medium der Sprache. *Anthony Moore* und *Siegfried Zielinski* setzen in ihrem Vortragsduett, das auf einer multimedialen Performance basiert, die Stimme als medienphilosophischen Reflexionsgegenstand, als mikroskopiertes und vermessenes Anatomieobjekt und als musikalisches Vokalinstrument gegeneinander.

Der vorliegende Band ist aus einer Konferenz mit dem Titel »Medien/Stimmen« hervorgegangen, die das Kulturwissenschaftliche Forschungskolleg »Medien und Kulturelle Kommunikation« (SFB/FK 427) an der Universität zu Köln im November 2002 veranstaltet hat. Die Herausgeberinnen danken insbesondere Jan Hoppe, Reiner Rasche und Marcus Krause für ihre redaktionelle Mitarbeit an diesem Band.

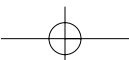
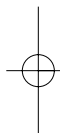
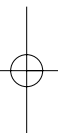
- 1 Einer der ersten, der dieses Desiderat aus kulturwissenschaftlicher Sicht erkannt hat, ist Karl-Heinz Göttert: *Geschichte der Stimme*, München 1998; aus medizinisch-psychologischer Perspektive: Horst Gundermann: *Phänomen Stimme*, München/Basel 1994; in jüngster Zeit: Reinhart Meyer-Kalkus: *Stimme und Sprechkünste im 20. Jahrhundert*, Berlin 2001; Friedrich Kittler/Thomas Macho/Sigrid Weigel (Hg.): *Zwischen Rauschen und Offenbarung. Zur Kultur- und Mediengeschichte der Stimme*, Berlin 2002. Eine Auswahlbibliographie zum Thema »Stimme« findet sich unter: [www.medienstimmen.de](http://www.medienstimmen.de).
- 2 So formulieren Aleida Assmann/Jan Assmann (Hg.): Einleitung, in: Eric Havelock: *Schriftlichkeit. Das griechische Alphabet als kulturelle Revolution*, Weinheim 1990, S. 1–35 (hier S. 3).
- 3 Vgl. die keineswegs kritisch kommentierte Gegenüberstellung der Implikationen von Oralität und Literalität bei Hans Georg Pott: *Die Wiederkehr der Stimme. Telekommunikation im Zeitalter der Postmoderne*, Wien 1995, S. 23.
- 4 Ebd.
- 5 Vgl. Jan Vansina: *Oral Tradition. A Study in Historical Methodology*, Middlesex/Victoria 1973; Richard Bauman/Joel Sherzer (Hg.): *Explorations in the Ethnography of Speaking*, Cambridge 1974; Sylvia Scribner/Michael Cole: *The Psychology of Literacy*, Cambridge 1981; Brian Street: *A critical look at Walter Ong and the »Great Divide«*, in: *Literary Research Center Newsletter* 4, Heft 1 (1988), S. 1 ff.
- 6 Der Terminus »great divide« als Charakterisierung der dichotomischen Gegenüberstellung von Oralität und Literalität geht zurück auf: Ruth Finnegan: *Literacy versus Non-Literacy. The Great Divide*, in: dies./Robin Horton (Hg.): *Modes of Thought*, London 1973, S. 112–144.
- 7 Vgl. Walter Ong: *Oralität und Literalität. Die Technologisierung des Wortes*, Opladen 1987; vgl. auch Cornelia Epping-Jäger: *Der Literalisierungsprozeß und die Ordnung des Wissens. Der deutsche Belial*, in: Ludwig Jäger/Erika Linz (Hg.): *Medialität und Mentalität. Theoretische und empirische Studien zum Verhältnis von Sprache, Subjektivität und Kognition*, München (im Druck).
- 8 Jacques Derrida: *Die Stimme und das Phänomen. Ein Essay über das Problem des Zeichens in der Philosophie Husserls*, Frankfurt/M. 1979.
- 9 Michael Wimmer: *Verstimmte Ohren und unerhörte Stimmen*, in: Dietmar Kamper/Christoph Wulf (Hg.): *Das Schwinden der Sinne*, Frankfurt/M. 1984, S. 115–139 (hier: S. 125).
- 10 Jacques Derrida: *Grammatologie*, Frankfurt/M. 1983, S. 29.
- 11 Zu der These, dass die Medialität eine konstitutive Eigenschaft bereits der präliteralen Sprache des Menschen darstellt, vgl. Ludwig Jäger: *Sprache als Medium. Über die Sprache als audio-visuelles Dispositiv des Medialen*, in: Horst Wenzel/Wilfried Seipel/Gotthart Wunberg (Hg.): *Audiovisualität vor und nach Gutenberg. Zur Kulturgeschichte der medialen Umbrüche*, Wien 2001, S. 19–42 (hier S. 22).
- 12 Friedrich Kittler: *Das Werk der Drei. Vom Stummfilm zum Tonfilm*, in: Kittler/Macho/Weigel: *Zwischen Rauschen und Offenbarung* (Anm. 1), S. 357–370 (hier: S. 358 und 365).
- 13 Horst Gundermann: *Das Phänomen Stimme*, in: ders. (Hg.): *Die Ausdruckswelt der Stimme. 1. Stuttgarter Stimmtage*, Heidelberg 1998, S. 3–7 (hier S. 4).
- 14 Dieser Vorstellung liegt das Modell einer Stimme zugrunde, die aus dem von Aristoteles so genannten »Luftschlag« entsteht: Die Atemluft schlägt auf die Außenluft auf, dabei entsteht hörbarer Schall. »[...] das Schlagende«, so Aristoteles, »muß belebt sein und eine bestimmte Vorstellung haben. Denn die Stimme ist ein Ton mit einer bestimmten Bedeutung.« (Aristoteles, *Von der Seele*, S. 308; zitiert nach: Karl-Heinz Göttert: *Vox – ein vernachlässigtes Kapitel der Rhetorik*, in: Heinrich Plett (Hg.): *Die Aktualität der Rhetorik*, München 1996, S. 57–66, hier S. 60f.) »Die menschliche Stimme«, so formuliert Karl-Heinz Göttert, dessen Ausführungen wir hier folgen, »entsteht in der Seele und ist damit irgendwie – *seelenvoll*. Dahinter steht die Vorstellung, daß die menschliche Stimme schon als artikulierte »schlägt«, und d. h., daß sie bereits in der Lunge das Gedachte verkörpert.« 1741 entdeckte Antoine Ferrein die Stimmbänder »und trägt damit die Theorie vom Luftschlag zu Grabe.« Vgl. ebd., S. 60f.
- 15 Gundermann: *Phänomen Stimme* (Anm. 1), S. 16 ff.

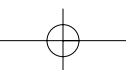
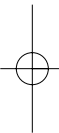
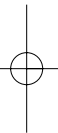
- 16 Meyer-Kalkus: Stimme und Sprechkünste (Anm. 1), S. 15.
- 17 Rudolf Fähmann: Die Deutung des Sprechausdrucks, Bonn 1960, zitiert nach Helmut K. Geißner: Über den Brustton der Überzeugung. Zur Sozialkritik des Imponiergehabes, in: Gundermann: Die Ausdruckswelt der Stimme (Anm. 13), S. 107.
- 18 Vgl. Bernhard Waldenfels: Stimme am Leitfaden des Leibes, in diesem Band.
- 19 Georg Wilhelm Friedrich Hegel: Enzyklopädie der Philosophischen Wissenschaften, Bd. III: Philosophie des Geistes, § 401, Zusatz, in: ders.: Werke, Bd. 10, Frankfurt/M. 1970, S. 109.
- 20 Ebd. S. 113.
- 21 Ebd. S. 116. Vgl. dazu die Beiträge von Leander Scholz und Erika Linz in diesem Band.
- 22 Karl Bühler: Sprachtheorie, Frankfurt/M. 1968, S. 31.
- 23 Wimmer: Verstimmte Ohren (Anm. 9), S. 25.





## I. STIMM-AKT





## Bernhard Waldenfels

### STIMME AM LEITFADEN DES LEIBES

Wenn wir den Leib zum Leitfaden nehmen, um uns im Labyrinth neuer Stimmmedien zu orientieren, so kommt es entscheidend darauf an, wie wir den Leib begreifen. Der Gedanke, dass der Leib selbst als Urmedium fungiert und dass Natürlichkeit und Künstlichkeit sich dabei durchdringen, könnte für eine fruchtbare Spannung zwischen Leib- und Medientheorie sorgen. Der zusätzliche Gedanke, dass uns die Stimme zunächst als gehörte Stimme begegnet, könnte technischen Allmachtsvisionen einen Dämpfer versetzen.<sup>1</sup>

#### 1. LAUTWERDEN ALS EREIGNIS

Die Stimme teilt das Schicksal anderer akustischer Phänomene, allzu schnell substantialisiert und personalisiert zu werden. Als »bedeutsamer Laut« (*φωνή σημαντική*) bezeichnet und bedeutet sie etwas; als kundgebender Laut bringt sie seelische Vorgänge und Zustände zum Ausdruck; als soziales Medium leistet sie ihren Beitrag zur Weitergabe von Sinn. Diese Semantisierung, Subjektivierung und Pragmatisierung der Stimme führt dazu, dass sie sogleich in Dienst genommen wird zu Zwecken der sachlichen Darstellung, des subjektiven Ausdrucks und der intersubjektiven Verständigung. Sie wird übertönt durch die Gebote sachlicher Stimmigkeit, durch subjektive Gestimmtheiten und durch die Forderung nach intersubjektiver Übereinstimmung. Sie findet ihren Platz in einer Welt von Sachen und Personen, in der jemand sich anderen gegenüber über etwas äußert. Man *hat* eine Stimme, sei es kraft Natur, sei es kraft Konvention, und so ist man ihr immer schon überlegen. Ihre Flüchtigkeit bildet einen zusätzlichen Makel, der durch dauerhafte Medien zu beheben ist. Doch es geht mir nicht in erster Linie um den Wettstreit zwischen Mündlichkeit und Schriftlichkeit,<sup>2</sup> sondern darum, dass die Stimme von vornherein einem Kommunikationsnetz angeschlossen ist wie ein Sprechapparat. Die Medialisierung, die inzwischen an der Tagesordnung ist, hat insofern etwas Zweideutiges, auch wenn sie eine bloße Instrumentalisierung von Sprache und Zeichen weit hinter sich lässt.

Um diese Mundanisierung, Subjektivierung und Sozialisierung der Stimme zu durchbrechen, bedarf es einer *akustischen Epoché*, die das Hörphänomen als solches freisetzt. Wenn Sartre bemerkt, dass der Leib im gewöhnlichen Leben mit

Schweigen übergangen wird, so können wir hinzufügen: Die Stimme wird als Stimme gewöhnlich überhört, und das gleiche trifft zu auf die alltägliche Geräusch- und Klangkulisse, aus der sich die Stimme heraushebt. Doch es geht nicht nur darum, dass die Laute hinter dem verlautbarten Sinn verschwinden. Das Hörphänomen wird verkannt, wenn man dahinter nicht mehr vermutet als die Verlautbarung von etwas, das im Stillen bereits da ist und nur auf seinen Auftritt wartet, besteht es doch in einem Lautwerden von etwas, das erst im Ertönen oder Erklingen zu dem wird, was es ist. Man denke an Klangbilder einer Großstadt, wie wir sie aus Radio-Features kennen. Hier wird nicht einer Stadt ein akustischer Mantel umgehängt, sondern die Stadt taucht auf *in statu audiendi*, und es wäre nicht unpassend, Italo Calvino's Anregung aufzugreifen und die hörbaren Städte durch unhörbare zu ergänzen. Die passende Grundformel, die uns davon abhält, der Lautwerdung vorweg etwas zu unterschieben, das den Laut trägt, oder jemanden, der ihn hervorbringt, lautet: »Es wird hörbar«, so wie wir sagen: »Es raschelt, es klingelt, es donnert«. Das Rätselhafte liegt in dem »es« dieser impersonalen Wendungen.<sup>3</sup> Wenn ich von einem Ereignis des Lautwerdens spreche, so ist nicht in erster Linie an pompöse und spektakuläre Großereignisse zu denken, sondern daran, dass immerzu etwas geschieht, das nicht schon als Akt oder Aktion zu begreifen ist und nicht vorweg den intentionalen und regelförmigen Beschreibungen und den subjektiven Zuschreibungen gehorcht, von denen die gewöhnlichen Sprechakt- und Handlungstheorien ausgehen. Dies gilt auch für die Stimme, die sich im Wechsel des Gesprächs oder im Ton der Aufforderung bemerkbar macht. Sie bedeutet mehr als einen Bedeutungsträger oder ein Ausdrucksmedium, und dieses Mehr gilt es aufzuzeigen.

## 2. HÖREN ALS AUFFÄLLIGWERDEN

Wie alles das, was als Bestimmtes auftritt, lässt sich auch das Lautwerden nur differentiell beschreiben, als Überschreiten einer Schwelle, die Hörbares von Unhörbarem scheidet, und als Abhebung des aktuell Hörbaren von einem Hintergrund des potentiell Hörbaren. Die Gesetze der Gestalttheorie gelten unter entsprechender Abwandlung nicht nur im optischen, sondern auch im akustischen Bereich. Die Auffälligkeit, die sich bis zur Aufdringlichkeit steigern kann, deren Skala vom Pianissimo bis zum Fortissimo reicht und die mit mancherlei akustischen Merkzeichen ihre Überraschungseffekte erzielt, ruft ein hörendes Selbst auf den Plan. Ich spreche von einem Selbst, nicht von einem Subjekt im üblichen Sinne; denn die Auffälligkeit, die traditionellerweise unter dem Titel der *Auf-*

*merksamkeit* verhandelt wird, reduziert sich keineswegs auf einen Akt des Aufmerkens, der von jemandem spontan oder gar verantwortlich vollzogen wird. Hörende sind durchaus im Spiel, aber zunächst als jemand, dem oder der etwas widerfährt, der oder die sich affiziert fühlt, also als jemand im Dativ oder im Akkusativ und nicht im Nominativ eines eigenständigen Akteurs, der als Seh- oder Hörsubjekt installiert ist. Die inzwischen bis in die neurologische Forschung hinein akzeptierte »Ich-Perspektive« tritt zunächst als »Mir-« oder »Mich-Perspektive« auf.<sup>4</sup> Hierbei sind sowohl auf Seiten der Stimmerzeugung wie auf Seiten des Stimmhörens geschlechtliche Differenzen zu beachten, die verschwinden, wenn die Stimmäußerung von vornherein Bedeutungsregeln und Geltungsansprüchen unterworfen und der quasi-rechtlichen Form einer Stimmabgabe angenähert wird. Doch selbst wenn wir auf der Ebene des Hörens bleiben, so ist zu beachten, dass das intentionale Hören *von etwas* anhebt als ein responsives Hören *auf etwas*, als Aufhorchen oder Hinhören, das geweckt wird, das auf Aufforderungen eingeht und nicht eigenmächtig erzeugt wird. Blicke es bei dem Hören und Sehen von etwas, so liefe jedes Hören und Sehen im Grunde auf ein Wiederhören und Wiedersehen hinaus, das im Bereich vorhandener und lediglich zu entfaltender Möglichkeiten verbliebe. Zwischen dem Wem-Status des »Patienten«, dem etwas zustößt oder zufällt, und dem Wer-Status dessen, der darauf eingeht, sich abschirmt, sich dagegen wehrt, aber so oder so antwortet, liegt ein Sprung. Dieser lässt sich durch keine Ziel- oder Regelordnung überbrücken, und er sprengt auch den Regelkreis, der sich im Feedback aus sich selbst speist. Das Hören beginnt anderswo mit einem Fremdlaut, einer Fremdstimme. Diese genuine Enteignung eines Selbstbewusstseins, das glaubt, bei sich selbst beginnen und zu sich selbst zurückkehren zu können, bedeutet freilich nicht, dass das Selbst, dem etwas auffällt und widerfährt, verschwindet. Eine Auffälligkeit ohne jemanden, dem etwas auffällt, wäre wie ein Schmerz ohne jemanden, der ihn spürt. Wer dieses Selbst auslöscht, verwandelt Klänge in bloßen physischen Schall, er ersetzt das Hörereignis durch physikalische Wellenbewegungen und neurophysiologische Prozesse und setzt voraus, was er zu erklären versucht. Die akustische Epoché, von der schon die Rede war, zielt dagegen nicht auf nicht-auditive Substrate und Subprozesse ab, sondern auf eine potenzierte Hörbarkeit, auf ein Hörbarwerden des Hörbaren, das dem Sichtbarwerden des Sichtbaren gleicht und das in der normalen Erfahrung unauffällig, also in diesem Sinne unhörbar und unsichtbar bleibt. Eine Phänomenologie des Hör- und Sichtbaren begegnet sich an dieser Stelle mit den Hör- und Sehkünsten, die den Boden des Schon-Gehörten und Schon-Gesehenen verlassen.<sup>5</sup> Das bloße Hören und Sehen wird durch sich selbst überboten, durch ein Andershören und Anderssehen, nicht durch ein Denken, das sich auf ein *cogito*

*me audire* zurückzieht oder am Ende – als Nachfahren des Geistes – das Gehirn hören und sehen lässt.

### 3. SICH-ANGESPROCHEN-FÜHLEN

Wir sind bisher von einer generellen Form der Hörbarkeit ausgegangen und haben eine Menge spezifischer Differenzen vernachlässigt. Dies geschah mit Bedacht. Würden wir schlichtweg mit den bekannten Differenzen beginnen, so würden wir in die Hörwelt Unterschiede einführen, die dem Hörbaren von außen auferlegt würden. Der »Logos der ästhetischen Welt«, den Husserl im Auge hat, ist auch als »Logos der akustischen Welt« zu verstehen, also als eine Ordnung des Hörbaren, die dem Hören selbst entwächst. Die gilt nicht nur für Unterscheidungen wie die von Sprachäußerung und Schrei, von Wortakzent und Tonfall, von Sprachartikulation und Sprachmelodie, sondern auch für die Sonderung von Sprechstimme, Gesangstimme und Instrumentenklang, von musikalischem Ton und Geräusch, mit der die moderne Musik in besonderem Maße experimentiert, um Tonsysteme und Klangtraditionen zu durchbrechen.<sup>6</sup> Doch wenn wir uns nun ausdrücklich dem Phänomen der Stimme zuwenden, so fragt es sich, worin ihre Eigenart besteht und wodurch sie sich vor anderen Hörphänomenen auszeichnet.

Das landläufige Kommunikationsschema legt es nahe, die Stimme in erster Linie einem Sprecher zuzuschreiben, der sprachliche Laute erzeugt, und dies im Gegensatz zu Geräuschen, die auf physische Dinge und Vorgänge zurückgehen. Das Geräusch hätte dann eine bloße Quelle, die Stimme einen zumindest möglichen Adressanten. Vokal- und Instrumentalmusik, die auf mannigfache Weise ineinander verschränkt sind, nähmen dann eine merkwürdige Mittellage ein. Doch halten wir uns an den Prototyp der Sprechstimme, so scheint diese mehr als alles andere jenem Logos verschwistert, der gemäß der alten Bestimmung das menschliche Lebewesen auszeichnet. Haben wir erst einmal diesen Bereich betreten, so können wir ohne weiteres vom Sprecher zum Hörer überwechseln. Die Stimme ist etwas, was der Hörer empfängt und prinzipiell versteht, im Gegensatz zu Geräuschen, die uns aufschrecken oder beruhigen, uns aber nichts zu verstehen geben. »Felder und Bäume wollen mich nichts lehren, wohl aber die Menschen in der Stadt«, versichert der lernbegierige Sokrates, »der krank ist an der Sucht, Reden anzuhören«.<sup>7</sup> Durch seinen Bedeutungs- und Mitteilungscharakter unterscheidet sich der *flatus vocis* von einem bloßen Windhauch. Mund und Ohr die-

nen dann als leibliche Organe, wörtlich: als Werkzeuge, mit Hilfe derer bedeutungsvolle Laute erzeugt und aufgenommen werden, und dies nur dann, wenn Mund und Ohr beseelt oder vergeistigt sind.

Das Problem liegt nun darin, dass eine solche kommunikative Vernunft zwar den Gegensatz von Viel- und Einstimmigkeit umgreift und auch die Kakophonie eines Stimmengewirrs zulässt, dass aber die Symphonie eines übergreifenden Logos die Fremdheit der Stimme zum Verstummen bringt. Wenn Hören jedoch damit beginnt, dass etwas zu Gehör kommt, dass etwas unser Ohr trifft und die Monotonie eines bloßen Dahinrauschens unterbricht, dann tritt die Stimme zunächst als *Fremdstimme* auf, als eine Stimme, die anderswoher kommt, gleich der Sprache, die uns zunächst als Sprache der Anderen, also als originäre Fremdsprache gefangen nimmt. Die Asymmetrie zwischen fremdem Anspruch und antwortendem Hören bildet keineswegs eine bloße Vorstufe auf dem Weg zu einem symmetrisch angelegten Dialog, sie gehört zur Genealogie eines Dialogs, dessen Werden nicht bereits auf einer vorgängigen Übereinstimmung, auf einer Homologie beruht. Wenn die Sprache aus dem Hörensagen erwächst, so ist die Stimme genau jene Instanz, in der dieses Hörensagen laut wird.

Es trifft zwar zu, dass die Stimme sich gegenüber dem bloßen Geräusch dadurch auszeichnet, dass in ihr eine *adressierte Aufforderung* zum Zuge kommt. Dies unterscheidet schlichte Äußerungen wie »Komm!« oder »Geh!« von einem Windstoß oder von dem Schwirren eines Geschosses. Stimmen sind auf andere Weise erweiterungsbedürftig als Geräusche, die uns anrühren, uns zu einem bestimmten Verhalten ermuntern, aber keinen Anspruch erheben.<sup>8</sup> Dies gilt mutatis mutandis auch für Gesänge, in die man einstimmen, denen man aber nicht zustimmen kann. Dennoch behält selbst die Stimme etwas Anarchisches, sofern das *Doppelereignis* von Sprechen und Hören sich nicht auf gesonderte, subjektiv verfügbare Sprech- und Hörakte zurückführen lässt. Die Frage »Wer spricht?« stellt sich eben dann, wenn wir der gehörten Stimme keinen autonomen Sprecher unterschieben. Sie findet ihre Antwort im Vernehmen der Stimme, das heißt aber, sie findet niemals eine vollständige und endgültige Antwort. Eben deshalb ist das Pathos des Angesprochenwerdens keineswegs als die Umkehrung eines aktiven Ansprechens zu betrachten. Der Logos erwächst aus dem Pathos in Form eines Sich-angesprochen-Fühlens, das sich niemals völlig in eine identifizierende Zuschreibung und Rollenverteilung überführen lässt. Entsprechend abgewandelt gilt dies auch für das Sich-angeblickt-Fühlen durch den fremden Blick. Pathologische Phänomene wie das halluzinatorische Hören von Stimmen oder der para-

noide Beobachtungs- und Verfolgungswahn sind nur dadurch zu erklären, dass auch normalerweise Stimme und Blick nicht völlig anzueignen und auf verschiedene Instanzen aufzuteilen sind. Wenn ein Wort das andere ergibt, so bedeutet dies mehr als eine Abfolge individueller Sprech- und Hörakte. Sprech- und Hörereignisse enthalten einen Kern an Anonymität, an Namenlosigkeit. »N'importe qui parle«, so Foucault im Anschluss an Beckett. Man mag einwenden, dass es unter Umständen, etwa im Falle eines Versprechens, durchaus darauf ankommt, wer spricht. Dennoch ist dies eine Frage, die mehr oder weniger offen bleibt; sie lässt Raum für eine wechselnde Verortung und für verschiedene Grade der Anonymität.

#### 4. SICH-SPRECHEN-HÖREN

Die Fremdheit der Stimme, die von Anderen ausgeht, findet ihren Widerhall in der Fremdheit der eigenen Stimme, die nur bis zu einem gewissen Grad als eigene Stimme bezeichnet werden kann. Was Derrida so minutiös in Form eines »Sich-sprechen-Hörens« dekonstruiert, betrifft in Wahrheit eine metaphysisch zu nennende Deutung der Stimme, nicht die ›Stimme selbst‹, die dazu angetan ist, von sich aus jede Metaphysik der Präsenz zu untergraben.<sup>9</sup> Die Tatsache, dass wir uns selbst sprechen hören, dass wir uns als Sprechende selbst affizieren und auch überraschen, führt zu einer Selbstspaltung zwischen mir als Sprechendem und mir als Hörendem. Wie Merleau-Ponty betont, gibt es eine »Reflexivität zwischen den Vorgängen der Lauterzeugung und dem Gehör«, aber keine Deckung; jene Vorgänge »schreiben sich klanglich ein, und jeder Stimmlaut weckt ein motorisches Echo in mir«; zwischen Lauterzeugung und Echo schiebt sich die Dichte des »Fleisches«.<sup>10</sup> Die eigene Stimme, in der eine eigentümliche Fremdheit anklingt, erleidet einen genuinen *Stimmbruch*, wenn sie nicht nur bei Anderen ein Echo weckt, sondern als »Klang« auftritt, »der in seinem eigenen Echo widerklingt«, so Levinas.<sup>11</sup> Die Re-sonanz der Stimme entspricht dem Re-flex der Spiegels. So wie dieser nicht nur Schon-Gesehenes abbildet, sondern das Selbst als Selbst sichtbar macht und mich mit mir selbst konfrontiert, so gibt das Echo nicht nur Schon-Gehörtes wieder, sondern macht mich mir als Selbst hörbar. Der Auftritt meiner selbst ist so besehen ein originärer Wiederauftritt, ich komme auf mich selbst zurück. Die Selbstspaltung, die mich von der eigenen Stimme trennt, hat zur Kehrseite eine Selbstverdoppelung im Anderen, eine Art Doppelgängertum. Dieses Motiv, das nicht nur in der Romantik eine solche Rolle spielt, sondern von Autoren wie Valéry und Merleau-Ponty zur Charakterisierung der Fremdheit des An-



deren herangezogen wird, ist nicht zu verwechseln mit dem trivialen Tatbestand, dass es noch ein weiteres Ego oder eine Pluralität von Subjekten gibt. Es besagt vielmehr, dass ich als leibliches Selbst dem fremden Blick und der fremden Stimme ausgesetzt bin als einer Resonanz, die mir nicht nur antwortet, sondern mir zuvorkommt, mich verblüfft.<sup>12</sup> Das Widertönen fremder Stimmen in der eigenen ist ein Emblem jener Zwischenleiblichkeit, in der Eigenes sich mit Fremdem verschränkt und überlagert.

Zugleich mit diesem Sich-sprechen-Hören und Andere(s)-in-sich-sprechen-Hören rückt auch der Gegensatz von Innen und Außen in ein anderes Licht. In der Tradition des englischen Empirismus pflegt man zu unterscheiden zwischen äußeren Bildern, die als Bildnisse oder Bildapparate abgelöst vom eigenen Leib existieren, und inneren Bildern, die nur in unserer Vorstellung vorkommen. Dem entsprechen Stimmen und Melodien, die uns von außen her oder auch aus unserer eigenen Kehle entgegenschallen, und solche, die wir gleichsam im Stillen auf der Bühne unseres Geistes aufführen. Nun ist zuzugeben, dass wir Klangformen und Tonabläufe hören können, ohne dass einer von uns den Mund auf tut oder etwas vor sich hin summt. Wir wären außerstande, nach einer Melodie oder auch nach dem Klang eines Namens zu suchen, wenn es keine vorausgreifenden wie auch zurückgreifenden Klang- und Tonmuster gäbe, und der Musiker, der seine Partitur noch einmal liest, bevor es zur Aufführung kommt, würde daraus keinen Nutzen ziehen, wenn er sich zwischen den Notenzeilen bewegen würde wie ein Algebräiker. Zweifellos gibt es Vor- und Nachklänge auf ähnliche Weise, wie es Vor- und Nachbilder gibt. Nicht zuletzt die Motorik, die jede Vorstellungstätigkeit, etwa auch das leise Lesen, begleitet, verweist darauf, dass die leibliche Stimme einer Körpergeschichte, speziell einer Hörgeschichte eingeschrieben ist. Die Neurologie bestätigt dies mit ihren externen Aufzeichnungsverfahren, die eine eigene Form der Graphie entwickeln. Die Definition der Explananda ist auch durch reduktionistische Analysen nicht zu erbringen.<sup>13</sup> Die Annahme, wir würden innere Bilder und innere Stimmen frei entwerfen, gehört zu den Fiktionen einer Subjektautonomie. Selbst Bekanntes muss uns wieder einfallen, ganz zu schweigen von Ungewohntem und Unerhörtem. Auch Höreinfälle, kommen, »wenn sie wollen, nicht wenn wir wollen«. Der Innenraum, in dem »innere Stimmen« oder »heimliche Melodien« erklingen, hat selbst sein Außen, das darin spürbar wird, dass sich uns eben das entzieht und verweigert, was wir doch angeblich selbst hervorbringen. Diese Entzugserfahrungen sind natürlich nicht frei von entsprechenden neurologischen Schaltungen, aber zu entdecken sind sie dort nicht.

## 5. HÖREN UND ÜBERHÖREN

Es bedürfte keiner akustischen Epoché, um die Stimme als solche zu thematisieren, würde diese nicht in der natürlich-kommunikativen Einstellung durchgängig überhört. Dieses Überhören ist nicht Folge einer bloßen Auslassung, sondern einer Überproduktion, die das stimmliche Hörereignis in ein Gehörtes verwandelt, das man registrieren, identifizieren, beschreiben, erklären und auch abändern kann. Die unvermeidliche *Normalisierung* der Erfahrung, ohne die es keine verlässliche Wiederkehr der Dinge gäbe, macht auch vor dem Stimmphänomen nicht halt. Dennoch ist die Stimme als Hörereignis *lautlos, unhörbar*, so wie der Blick als Sehereignis *bildlos, unsichtbar* bleibt. So wie das Sehen im Sehereignis seinen blinden Fleck hat, so hat das Hören im Hörereignis seine stumme Pause. Jede Phoné rührt an eine Aphonie, die dem beredten Schweigen gleicht. Um dieses Unhörbare zu erfassen, bedarf es keiner separaten »Ohren des Geistes«, sondern eines »dritten Ohrs«. Es bedarf eines *Andershörens*, das die phono-logische Ordnung durchbricht und der domestizierten, normalisierten Stimme die Untertöne einer wilden Stimme entlockt. Die »wilden Laute«,<sup>14</sup> die hier anklingen, machen sich auf ihre Weise bemerkbar in der zweideutigen Form von Störlauten, von Pausenzeichen oder im Stocken des Hörflusses.

## 6. DER LEIBKÖRPER ALS UREDIUM

Mit letzteren Bemerkungen nähern wir uns dem Problem einer Medialisierung der Stimme. Medialisierung kann auf zweierlei Weise verstanden werden, nämlich derart, dass die leibliche Stimme medial gebrochen, verstärkt oder erweitert wird durch Hörgerät, Mikrophon oder Telephon, oder dass sie medial ersetzt wird durch Apparaturen wie Grammophon, Tonband oder Klanglabor, in denen Stimmen künstlich erzeugt werden ohne Mitwirkung der leiblichen Stimmorgane. Ich spreche selbst am Telephon, wohingegen das Tonband ohne mein Zutun abläuft, selbst wenn es meine Stimme ist, die dort wiedergegeben wird. Bedeutet dies, wie manche meinen, dass die Leiblichkeit der Stimme mitsamt ihrem somatomorphen Zubehör an Techniken mehr und mehr abdankt zugunsten einer technomorphen Körperlichkeit unter Einschluss einer *artificial voice*? Die Beantwortung der Frage, ob es zu einer technologischen Kehre kommt, hängt davon ab, wie man die Leiblichkeit begreift. Für einen Leib, der im reinen Spüren seiner selbst beheimatet ist, fällt jeder Außenblick und jeder äußere Eingriff unter eine sekundäre Verkörperung, die das Arkanum des eigenen Leibes nur indirekt berührt.

Umgekehrt gilt, dass unter dem Andrang einer resoluten Körpertechnik dieses Arkanum zusammenschrumpft zu einem Reservat, einem Überbleibsel, dessen Tage gezählt sind. Anders sieht es aus, wenn das leibliche Selbst nicht fugenlos, sondern in sich selbst gespalten und verdoppelt ist und wenn der Leib als *Leibkörper* selbst der Sphäre der Sichtbarkeit und Hörbarkeit angehört, die er eröffnet. Um nochmals Merleau-Ponty zu zitieren: »Wie der Kristall, das Metall und viele andere Substanzen bin ich ein tönendes Wesen, aber meine eigene Vibration höre ich von innen her.«<sup>15</sup> Die Materialität der Stimme lässt sich nicht überspringen oder nachträglich hinzufügen, vielmehr hat die Stimme als Stimme ihre Materialität. Pascals berühmtes Diktum abwandelnd könnte man den Menschen nicht nur als denkendes, sondern auch tönendes Schilfrohr bezeichnen. Es stellt sich dann die Frage nach den leiblichen Ansatzstellen für eine medial und technisch verfasste Körperlichkeit.

Zunächst ist zu unterscheiden zwischen dem *fungierenden Leib* und der Thematisierung, Erforschung und Behandlung bestimmter *Körperglieder* und *Körperorgane*. Letztere kann bis zu dem Punkt gehen, wo der Leib als bloßes Körperding, leibliche Vollzüge als bloße Körperprozesse auftreten und der Leichnam dem lebendigen Leib den Rang ablauft. Die erwähnte Unterscheidung gilt auch für die Stimme. Sie fungiert als Stimme, bevor sie als rein akustisches Schallphänomen auftritt. Sie fungiert, indem sie als Stimme laut wird und als solche vernommen wird.

Können wir sagen, dass die Stimme als Medium fungiert? Verstehen wir unter einem Medium das, *worin* etwas sichtbar, hörbar oder fühlbar wird, so ist die Stimme ebenso wenig ein Medium wie der Blick. Ein Seh- oder Hörereignis, das darin besteht, dass etwas in den Blick oder zu Ohren kommt, ist kein Medium. Als Medien können wir dagegen jene Bild- bzw. Laut- und Klangformen betrachten, in denen Gesehenes und Gehörtes sich auf bestimmte Weise darstellt. Dazu gehören auf der einen Seite Perspektiven und Farbskalen, auf der anderen Seite Lautskalen, Tonarten und Klangfarben. Eine Sonate wird in A-Dur gespielt, wie eine Bitte auf Französisch vorgetragen wird. Wie wir ›in Bildern‹ sehen, bevor wir diese vor uns sehen, so hören wir ›in Klang- oder Tonformen‹, bevor wir diese als solche vernehmen. Dies gilt für sogenannte innere Bilder und Klänge ebensogut wie für abgesondert existierende Bilddinge und Klangkörper. In kosmologisch verankerten Wahrnehmungslehren wie der von Aristoteles stoßen wir auf natürliche Medien, so das Licht im Bereich des Sichtbaren, die Luft im Bereich des Hörbaren und das Fleisch im Bereich des Tastbaren.<sup>16</sup> Schon dieses »Zwischen« ist

weder ein dingliches Merkmal noch jemandes Erlebnis, es gehört zum *Wie des Erscheinens*, zu dessen Modalitäten. Die Moderne bleibt dabei nicht stehen; sie zeichnet sich aus durch die zunehmende Einsicht in die Kontingenz dieses Wie, durch die Entkanonisierung der Wahrnehmungsformen und durch ein entsprechendes Experimentieren mit den Seh- und Hörbedingungen. Der Mensch, der als »nicht festgestelltes Tier« wechselnden Widerfahrnissen ausgesetzt ist, sieht sich gezwungen, Antworten zu erfinden und künstliche Feststellungen zu treffen. Daraus entsteht eine kulturell variierende Geschichte des Leibes und der Sinne, und der Kontrast zwischen Natürlichkeit und Künstlichkeit fällt selbst in diese Geschichte. Die Tatsache, dass das, wovon wir affiziert werden und worauf wir zu antworten haben, nicht selbst etwas ist, was in den Blick und zu Gehör kommt, öffnet einen Spielraum für mediale und technische Erfindungen, die weder durch eine allgemeine Naturanlage noch durch eine allgemeine Geistesausstattung vorwegzunehmen sind. Reine Natur und reiner Geist sind bloße Spaltprodukte, denen wir uns unter extremen Bedingungen annähern.

Der Leib lässt sich nun als *Urmedium* betrachten. Dies besagt nicht, dass Medien sich auf einen medienfreien Urleib zurückbeziehen, sondern es besagt umgekehrt, dass schon der Leib mediale Züge annimmt. Die Medialität des Leibes ist originär, eben weil ein Sehen und Hören, das aus einem Seh- und Hörereignis hervorgeht und nicht aus selbstgesteuerten Seh- und Hörakten, sich selbst entleitet. Es findet seinen Halt nur in etwas Anderem. Ich höre mich selbst sprechen, singen oder schreien, doch zwischen den Leib, der diese Laute hervorbringt, und jenen, der sie vernimmt, schiebt sich jener Spalt, der im klangverdoppelnden Echo aufklafft auf ähnlich Weise, wie die bildverdoppelnde Spiegelung sehenden vom gesehenen Leib absondert. Hören und Sehen sind nicht für sich selbst transparent nach Art eines reinen Selbst- und Für-sich-selbst-Erscheinens, sie sind leiblich gebrochen wie im Falle des Vibrato, in dem der ganze Leib mitschwingt. Aufgrund seiner materialen Dichte fungiert der Leib als sein eigenes Medium, das zugleich verbindet und trennt, formt und verformt. So wie laut Marcel Mauss<sup>17</sup> der eigene Leib als erstes technisches Mittel in Kraft tritt, so wirkt er auch als Urmedium, speziell als Urbild und Urskript, aber eben auch als ursprüngliche Klangfläche und als ursprünglicher Klangraum, so dass Spuren des Hör- und Sichtbaren sich im Medium des Körpers ablagern. Dies gilt ebenso für die Subsprache mit ihren Modulationen und Rhythmen, die von Sprachgeräuschen durchsetzt sind.<sup>18</sup>

## 7. HÖRSZENEN

Zur leiblichen Situation gehört es, dass wir uns ausgehend von einem Hier und Jetzt in einem Umfeld bewegen und dass alles leibliche Verhalten sich in Szenarien und Sequenzen entfaltet. Dabei ist jedes spezifische Sensorium in eine übergreifende Koinaisthesis eingelassen und jede Aisthesis mit einer Kinesis zu einer Kinaisthesis verflochten. Auf diese Weise findet auch das Stimmereignis hier und jetzt statt, genauer noch: Es schafft sich wie jede leibliche Äußerung eine Stätte. Die Sinne sind auf je verschiedene Weise zeit- und raumbildend. Wenn Husserl bemerkt, dass das Handeln nicht eigentlich produziert, sondern inszeniert wird,<sup>19</sup> so trifft dies auf das leibliche Verhalten generell zu. Bei dieser Inszenierung ist weniger an ein Regietheater zu denken als an ein »Theater ohne Autor«. *Etwas setzt sich in Szene*. Dies bedeutet zunächst, dass Bewegungskräfte freigesetzt werden über alle Ziele und Regeln hinaus. Dies bedeutet ferner, dass im Laufe des Geschehens eine Bühne entsteht und sich fortlaufend verändert. Dies bedeutet schließlich, dass die Aufführung keiner einheitlichen Regie gehorcht, dass sie sich statt dessen auf verschiedene Bühnen verteilt, wozu auch der »andere Schauplatz« des Unbewussten gehört. Was neuerdings unter dem Begriff der Performanz erörtert wird, hat diverse Vorläufer, so etwa in den verschiedensten Feldtheorien, in E. Goffmans Rahmenanalyse und eben auch in der reichhaltigen Phänomenologie von Leib und Raum. Regelbefolgung und Kompetenzerwerb haben hier immer schon eine geringere Rolle gespielt als in sprachanalytischen, normorientierten und kognitivistischen Denkansätzen.

Vor diesem allgemeinen Hintergrund gewinnt die Hörszene ihr eigenes Profil. Zunächst ist bei akustischen Phänomenen der zeitlich orientierte *Ereignischarakter* besonders stark ausgebildet. Anders als Farb- oder Tastqualitäten haften Klänge und Geräusche niemals an Dingen. Es gibt zwar Tonträger und Tonquellen, auch einen materiellen Klangkörper, der im Falle der Stimme aus einem klangerzeugenden ›Sprechapparat‹ besteht, nämlich aus Atemapparat, Kehlkopf und Vokalkontrakt, doch sind Töne und Klänge nicht dingfest zu machen. Deshalb überwiegt bei Klangbeschreibungen auch die verbale Ausdrucksweise: Es klingt, es donnert, es klopft, es stürmt, es schreit gen Himmel. Unsere abendländische Ontologie sähe anders aus, wenn sie sich stärker an das Hören und weniger stark an das Sehen angelehnt hätte.

Daraus folgt keineswegs, dass die Stimme und auch die musikalische Welt der Töne einseitig mit der Zeit und speziell mit einer Seelenzeit im Bunde stehen. Die

Dichotomie von Seelenzeit und Dingraum entstammt mitsamt der Bevorzugung der ersteren einer metaphysischen Deutung der Phänomene, sie findet in diesen selbst keinen hinreichenden Rückhalt. Schon die aristotelische Definition der Stimme als »Laut eines beseelten Wesens«<sup>20</sup> hat etwas Schillerndes, weil die Seele nicht als Innerlichkeit gedacht wird, sondern als Sich-selbst-Bewegendes, so dass auch der Stimmlaut einen kinetisch-räumlichen Aspekt bewahrt. Gehen wir von der speziellen Form der leiblichen Kinästhesie aus, so haben wir es immerzu mit einem *Zeit-Raum* zu tun, und eben dieser nimmt als Hör- und Klangraum eine besondere Gestalt an. Klänge, Geräusche und Stimmen kommen aus einer bestimmten *Richtung*, aber sie begegnen uns niemals frontal wie das, was wir ins Auge fassen. Dem Hörbaren, das sich störend als Lärm bemerkbar macht, sind wir mit einer besonderen Schutzlosigkeit ausgeliefert; die Ohren lassen sich nicht spontan schließen wie die Augen. Stimmen haben ihren Herkunftsort, aber dieser lässt sich nur indirekt und immer nur mehr oder weniger bestimmt lokalisieren, so dass man geradezu von einer spezifisch akustischen Unschärferelation sprechen könnte. Daher bietet die Stimme besonderen Anlass zur Sublimierung und Mystifizierung. Ferner können wir in eine Klangwelt eintauchen wie in ein Element. Eine solche Immersion ist ein Spezifikum der Hörsphäre; darin steht diese der Geruchsatmosphäre näher als dem Anblick, den uns etwas oder jemand gewährt. Es ist bezeichnend, dass die bildenden Künste dort, wo sie die Farbgebung und Linienführung von der Bindung an vorgegebene Dinge abzulösen trachten, sich nicht selten an musikalischen Darstellungsweisen orientieren.

Noch eine weitere Eigenart wäre zu erwähnen. Während optische Phänomene einander verdecken und es gezielter Anstrengungen bedarf, um etwa mit Hilfe eines wässrigen Farbauftrags oder durch Einschaltung durchsichtiger Medien die Verdeckung abzumildern, kommt es im Hörbereich zu Überlagerungen. Die Gliederung in Ober- und Unterstimme und die Gleichwertigkeit der Stimmen in der Polyphonie schaffen eine Klangarchitektur, die ein Gegengewicht bildet zur Tonfolge. Ähnliches trifft zu auf die Ober- und Untertöne der Sprache, durch deren Einsatz diese auf verschiedenen Ebenen zugleich agiert, bis hin zur Möglichkeit einer Heterophonie, die von prädikativen wie auch von sogenannten performativen Widersprüchen wohl zu unterscheiden ist. Der Widerstreit der Stimmen, der zulässt, dass sich Fehlstimmen in das offizielle Sprechverhalten einschleichen, hat mit geltungsrelevanten Widersprüchen direkt nichts zu tun. Wer von »begleitigen« spricht, widerspricht sich nicht. Solch ein *lapsus linguae* setzt eine Stimmführung voraus, der die Zügel immer wieder zu entgleiten drohen. Damit öffnen sich Schleusen der *Stimmverführung*. Auch diese Verfüh-

rungskraft, mittels derer der Andere in das Selbst eindringt und immer schon eingedrungen ist,<sup>21</sup> gehorcht sinnesspezifischen Gesetzen. Der Klangzauber der Sirenen lässt sich nicht einfach durch einen Farbzauber ersetzen. Schon Höhe und Tiefe der Klänge, die sich mit Bewegungen des Absinkens und Auftauchens verbinden, eröffnen eine eigenen Dimension, die der Tiefe des Gesichtsraums verwandt, ihr aber nicht gleichzusetzen ist.

Bei aller Differenzierung ist allerdings zu beachten, dass es keine reine Hörszene gibt, sondern nur synästhetische und kinästhetische Szenerien, deren Ausgestaltung mehr oder weniger durch Hör- und speziell durch Stimmphänomene dominiert ist. Hinzu kommt das Wechselspiel der verschiedenen Sinne, das bis zu diffusen oder eigens komponierten Gesamtszenen führen kann. Schon die Situation, in der wir den Sprechenden nur hören, ohne ihm »aufs Maul« schauen zu können, stellt eine Sondersituation dar. Aber auch dann stimmt es nicht, dass man den Anderen nur hört, man hört anders, unsicherer, hilfloser, aber auch abgeschirmter, wenn man dem Sprechenden nicht in die Augen schauen und Bestätigungen einholen kann.<sup>22</sup> Tut man dies, so fungieren Mund und Auge als leibliche Medien; denn »aufs Maul schauen« und »in die Augen schauen« bedeutet mehr als den Mund oder die Augen sehen, es gehört zu dem, was wir als Aufmerken vom bloßen Registrieren und Rezipieren unterschieden haben. Die Maske (*persona*), durch die hindurch die Stimme ertönt, gleicht einer zweiten Haut, einer Klanghaut. Etwas Maskenhaftes hat das Gesicht stets, da die fremde Stimme und der fremde Blick leibkörperlich gebrochen in Erscheinung treten. Was aber geschieht, wenn der Anblick sich einem künstlichen Image, die Stimme sich einer künstlich fabrizierten Stimme annähert oder gar durch sie ersetzt wird? Bevor wir uns an Horrorszenen berauschen, sollten wir nochmals genauer hinschauen und hinhören.

## 8. MEDIALISIERTE STIMMEN

Dass Medien zur Welt- und Selbstbildung beitragen und nicht nur instrumentelle Hilfestellung leisten, indem sie Vehikel und Speicheranlagen für die Weitergabe und Aufbewahrung von Sinn bereitstellen, versteht sich von selbst, wenn wir dem Leib, der wir sind und den wir nicht handhaben wie ein Werkzeug, eine Urmedialität zuschreiben. Die Frage ist nur, wie der Beitrag der Medien aussieht, wieweit er reicht und wie er sich auf die organisch verankerte Leiblichkeit auswirkt. Werfen wir also einen Blick auf die Vielfalt der Medienlandschaft.<sup>23</sup> Die ständig anwachsende Phonotechnik reicht von den Aufzeichnungen des Gram-

mophons über die Schnitt- und Löschtechnik des Tonbandes bis zum HiFi-Verfahren, in dem mehrere Aufnahmekanäle eingesetzt, verschiedene Aufnahmen verschnitten und im Play-Back-Verfahren nachträglich Stimmen unterlegt werden – auf der Suche nach einer Stimme, die sich von den Schlacken einer natürlichen Stimmführung befreit.<sup>24</sup> In der Digitalisierung unterläuft sie die Differenz zwischen Optik und Akustik, die an die Wiedergabe qualitativ differierender Sinnesphänomene gebunden bleibt. Dieser Prozess, in dem die Stimme durch Verfahren der Dekontextualisierung, der Synthetisierung und der Perfektionierung künstlich erzeugt wird, erzielt in der Erzeugung einer künstlichen Natürlichkeit, etwa durch Einfügung hörgerechter Nebengeräusche, einen hypermodernen Pygmalioneffekt. Die Dekontextualisierung tastet den Ereignischarakter der Stimme selbst dann an, wenn er es sich um eine Wiedergabe handelt; das Gerät, das eine Stimme konserviert, lässt sich vor- und zurückdrehen, wie man in einem Buch vor- und zurückblättert. Auch der Aufforderungscharakter der fremden Stimme lässt sich bearbeiten und einsetzen wie ein Stimmköder oder eine Stimmfalle. Die Grenzen zwischen Wirklichkeit und Fiktion verwischen sich auf ungeahnte Weise, wenn der Life-Charakter selbst nicht mehr augenfällig und vernehmbar ist. Wenn es sprichwörtlich heißt, das Papier sei geduldig, so sind Bildschirm und Tonaufnahme noch weitaus geduldiger. Schreibt man den computergestützten Medien neuartige Kräfte einer »Welterzeugung« zu, die sich aus Verfahren der Digitalisierung, der Virtualisierung und telematischen Interaktion speisen,<sup>25</sup> so wäre die Stimmerzeugung Teil dieses Erzeugungsprozesses.

Allerdings hat es mit der Stimme ein besondere Bewandtnis, da sie an den Lebensnerv zu rühren scheint, wie es bei kognitiven und dezisionellen Leistungen so nicht der Fall ist. Lässt sich eine Stimme operationalisieren? Der Hinweis auf Anrufbeantworter oder auf das Gehirn als »Reizbeantwortungsmaschine«<sup>26</sup> scheint zu genügen, um diese Frage mit Ja zu beantworten. Doch inwiefern antwortet der Apparat?<sup>27</sup> Bevor ich mich in einer technologischen Kasuistik verliere, möchte ich die Frage nach der Künstlichkeit der Stimme zuspitzen, indem ich mich nochmals auf die aristotelische Definition der *viva vox* beziehe. Die Stimme ist der »Laut eines beseelten Lebewesens«, so heißt es dort, und dazu gehört auch die zeit-räumliche Situierung der Stimme. Eine Stimme, die niemandes Stimme wäre und wie ein Spukgeist in der Welt herumgeistern würde, wäre keine Stimme mehr. Was die Apparate hervorbringen, wären dann nichts weiter als *stimmartige* Phänomene, die der wahren, sprich: natürlichen Stimme keinen Abbruch tun. Doch diese Immunisierungsstrategie würde nur dann verfangen, wenn Natürlichkeit und Künstlichkeit sauberlich zu trennen wären. Außerdem



würde die Frage nach dem Umgang mit den neuen Medien so erst gar nicht berührt. Unsere bisherigen Überlegungen zur Leiblichkeit der Stimme weisen in eine andere Richtung.

Zunächst stellt sich die Sphäre der Leiblichkeit nicht als eine homogene Sphäre da, die man nach Belieben ergänzt oder vermindert. Es gibt mehr oder weniger zentrale und periphere Körperpartien und Körperprozesse, die sich durch eine wechselnde Nähe und Ferne zum leiblichen Selbst auszeichnen und uns mit wechselnder Stärke affizieren. Ich bin mehr oder weniger und niemals gänzlich ich selbst. Deshalb bedeutet eine Herztransplantation einen stärkeren Eingriff in die Leibsphäre als das Anlegen einer Prothese oder die Einnahme eines Medikaments, und selbst hier gibt es bekanntlich erhebliche Unterschiede. In diesem Sinne gehört die Stimme gewiss zu den zentralen Instanzen des Leibes und ist in besonderem Maße »beseelt«. Doch dies besagt keineswegs, dass sie eine reine »Seelenstimme« ist. Wenn der leibliche Selbstbezug mit einem Selbstentzug Hand in Hand geht und wenn der Selbstentzug sich auch auf die Materialität des Leibkörpers erstreckt, so folgt daraus, dass die Stimme niemals ganz und gar in unseren Besitz übergeht, sondern eine gewisse Fremdheit behält. Der »Laut des beseelten Lebewesens« ist nicht durch und durch »beseelt« und »belebt«. Er ist von den Klängen und Geräuschen »unbeseelter« Dinge ebensowenig durch einen Graben getrennt wie von technisch und medial erzeugten Kunststimmen. Natürlichkeit und Künstlichkeit durchdringen sich, und unweigerlich auftretende Probleme sind solche der Dosierung.

Dies führt uns zu einer weiteren Folgerung, die ebenfalls durch unsere Vorüberlegungen nahegelegt wird. Die Beurteilung technischer Errungenschaften orientiert sich allzu selbstverständlich an den Erzeugungsaspekten. Darin gleicht die Phonotechnik den Sprach- und Handlungstheorien, die samt und sonders dazu neigen, einseitig die Position des Sprechers oder des Täters hervorzukehren. Doch wenn Hör- und auch Stimmereignisse damit anheben, dass jemand von Stimme, Ton oder Geräusch überrascht und überfallen wird, so greift eine Analyse der Stimme, die von der Stimmerzeugung ausgeht, zu kurz. Auch auf der Ebene der Technik, nicht erst auf der Ebene der Ethik, gibt es eine Asymmetrie. Die Gebrauchstechnik fällt keineswegs zusammen mit einer Umkehr der Herstellungstechnik.<sup>28</sup> Einer wirklich perfekten Technik, die einem Pygmalion Ehre machen würde, müsste es gelingen, selbst das Pathos noch herzustellen und das Gehör in ein Produkt zu verwandeln. Dies wäre dann der fragwürdige Gipfel einer Biotechnik, die auch vor der Phonotechnik nicht haltmacht.

Dass die Technik ihrerseits pathogene Wirkungen nach sich zieht, steht außer Zweifel. Da geht es ihr nicht besser als den anderen Großbereichen der Kultur, denen Freud längst eine entsprechende Rechnung aufgemacht hat. Die Medialisierung nähme totalitäre Züge an, wenn es nichts mehr gäbe, was medialisiert, und niemanden, für den etwas medialisiert wird. Man wird mit Fug und Recht behaupten können, dass das Medium die Botschaft prägt,<sup>29</sup> doch die Annahme, dass das Medium selbst die Botschaft *ist*, gehört zu den Übertreibungen einer Medientheorie, die ein Problem zu lösen verspricht, das sie selbst verursacht.

- 1 Zu den allgemeinen Voraussetzungen der folgenden Ausführungen vgl. Bernhard Waldenfels: *Das leibliche Selbst*, Frankfurt/M. 2000, speziell zur Stimme: S. 379–393 bzw. ders.: *Antwortregister*, Frankfurt/M. 1994, S. 487–496.
- 2 Vgl. hierzu unter Berücksichtigung der stimmlichen Kundgabe Bernhard Waldenfels: *Ferne und Nähe in Rede und Schrift*, in: ders.: *Vielstimmigkeit der Rede*, Frankfurt/M. 1999, Kap. 1.
- 3 Zur Problemgeschichte dieser grammatischen Figur vgl. *Historisches Wörterbuch der Philosophie*, Basel 1971 ff., Bd. 4, 1976, Stichwort: *Impersonalien*, S. 258 f.
- 4 So weist Wolf Singer darauf hin, dass die Wahrnehmungsschwellen, die in den neuronalen Auswahlssystemen auftreten, von dem Vorhandensein einer Aufmerksamkeit abhängig sind, die wir umgangssprachlich als »hinhören« oder »hinschauen« zu bezeichnen pflegen. Vgl. ders.: *Der Beobachter im Gehirn. Essays zur Gehirnforschung*, Frankfurt/M. 2002, S. 220.
- 5 Zur potenzierten Sichtbarkeit vgl. Bernhard Waldenfels: *Das Rätsel der Sichtbarkeit*, in: ders.: *Der Stachel des Fremden*, Frankfurt/M. 1990, S. 210–217.
- 6 Dieses weitere Hörfeld wird im folgenden ausgeblendet. Vgl. dazu Bernhard Waldenfels: *Lebenswelt als Hörwelt*, in: ders.: *Sinnesschwellen*, Frankfurt/M. 1999, Kap. 8.
- 7 Platon: *Phaidros* 228 b, 230 d.
- 8 Zur Unterscheidung von adressierter und unadressierter Aufforderung vgl. meine ausführlichen Darlegungen in: Bernhard Waldenfels: *Bruchlinien der Erfahrung*, Frankfurt/M. 2002, S. 109–113.
- 9 Es sei verwiesen auf meine Ausführungen zu Derrida: *Sich-sprechen-Hören*. Zur Aufzeichnung der phänomenologischen Stimme, in: Bernhard Waldenfels: *Deutsch-Französische Gedankengänge*, Frankfurt/M. 1995, Kap. 6.
- 10 Maurice Merleau-Ponty: *Das Sichtbare und das Unsichtbare* [1964], München 1986, S. 189.
- 11 Emmanuel Levinas: *Jenseits des Seins oder anders als Sein geschieht* [1974], Freiburg/München 1992, S. 228.
- 12 Vgl. Waldenfels: *Bruchlinien der Erfahrung* (Anm. 8), S. 207–215.
- 13 Vgl. Singer: *Der Beobachter im Gehirn* (Anm. 4), S. 25.
- 14 Vgl. Roman Jakobson: *Kindersprache, Aphasie und allgemeine Lautgesetze*, Frankfurt/M. 1969, S. 26.
- 15 Merleau-Ponty: *Das Sichtbare und das Unsichtbare* (Anm. 10), S. 189.
- 16 Vgl. Aristoteles: *De anima* II, 7–12.
- 17 Marcel Mauss: *Techniken des Körpers*, in: ders.: *Soziologie und Anthropologie*, Bd. II, München 1974, S. 206.
- 18 Vgl. dazu Waldenfels: *Antwortregister* (Anm. 1), S. 286–289, ferner: *Vom Rhythmus der Sinne*, in: ders.: *Sinnesschwellen* (Anm. 6), Kap. 3. Auch Sybille Krämer unternimmt den Versuch, die Alternative von intentionalem Sinn und kausaler Außenwirkung zu unterlaufen, indem sie – ausgehend von Derrida – das Medium als Spur begreift; vgl. Sybille Krämer: *Das Medium als Spur und Apparat*, in: dies.: *Medien Computer Realität*, Frankfurt 1998, S. 78 ff.
- 19 Edmund Husserl: *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie*, 2. Buch, (*Husserliana* IV), Den Haag 1952, S. 98, 259.
- 20 Aristoteles: *De anima* II, 8, 420 b 5 f.
- 21 Vgl. Jean Laplanche: *Die unvollendete kopernikanische Revolution in der Psychoanalyse* [1992], Frankfurt/M. 1996.

- 22 Vgl. bei Marcel Proust das Erstaunen des Erzählers, als er am Telephon die Stimme seiner Großmutter, befreit von der »aufgeschlagenen Partitur ihrer Züge« und »ohne die Maske des Gesichts«, wie zum ersten Mal hört und in ihr eine ungeahnte Zartheit und Brüchigkeit entdeckt. Marcel Proust: *À la recherche du temps perdu* (Pléiade), Paris 1954, Bd. II, S. 134f., deutsch: ders.: *Auf der Suche nach der verlorenen Zeit*, Bd. III, Frankfurt/M. 1955, S. 195.
- 23 Vgl. dazu den Versuch von Kristin Westphal, Phonomedien und Phonotechniken innerhalb der Erfahrung zu verorten: Kristin Westphal: *Wirklichkeiten von Stimme. Grundlegung einer Theorie der medialen Erfahrung*, Frankfurt/M. 2002. Die Autorin dieser phänomenologisch angelegten Untersuchung gibt den Medien ein gehöriges Gewicht, doch hütet sie sich davor, mediale Erfahrung in bloße Medialität umschlagen zu lassen.
- 24 Wie vergeblich und illusionär ein solcher Versuch ist, zeigt Matthias Fischer in einem Essay, in dem er den Abstand zwischen musikalischem Ereignis und technischer Reproduktion ausmisst und den Identitätsverheißungen der Phono-Industrie ihren »Platonismus für das Volk« vorhält: Die Platte soll dem flüchtigen Hörereignis Dauer verleihen, obwohl die eigentümliche Produktion der Reproduktion doch gerade der differierenden Wiederholung entspringt. Vgl. Matthias Fischer: *Die Stimme der Musik und die Schrift der Apparate*, in: ders.: Dietmar Holland, Bernhard Rzehulka (Hg.): *Gehörgänge. Zur Ästhetik der musikalischen Aufführung und ihrer technischen Reproduktion. Mit Beiträgen von Sergiu Celebidache und Glenn Gould*, München 1986 (hier: S. 26).
- 25 Vgl. in diesem Sinne Krämer: *Das Medium als Spur und Apparat* (Anm. 18), S. 85; die Autorin unterscheidet strikt zwischen Werkzeug und Apparat.
- 26 Vgl. Singer: *Der Beobachter im Gehirn* (Anm. 4), S. 26.
- 27 Ich verweise auf meine diesbezüglichen Ausführungen in: Waldenfels: *Antwortregister* (Anm. 1), S. 115–121.
- 28 Vgl. ebd., S. 289–292.
- 29 So Krämer: *Das Medium als Spur und Apparat* (Anm.18), S. 81.

**Leander Scholz**

**TIERSTIMME/MENSCHENSTIMME: MEDIEN DER KOGNITION**

Das Phänomen der Stimme ist in der abendländischen Metaphysik ein besonderer Schauplatz der Frage nach dem Menschen, da dem Lautcharakter der Stimme im Unterschied zum bloßen Klang eines rein materiellen Körpers der Selbstaussdruck des Lebendigen zugesprochen wird. Während Nichtlebendiges nur Geräusche produziert, wenn eine Fremdeinwirkung vorliegt, so wird der Stimme ein Selbst zugerechnet, das je nach der Konzeption dieses »Selbst« Ausdruck für eine Person, ein Subjekt oder ein Selbstbewegtes sein soll. In jedem Fall erscheint die Stimme nicht als eine Äußerung dieses Lebendigen unter anderen, sondern soll das Lebendige in hervorragender Weise selbst verkörpern. Sie wird daher nicht nur als ein Anzeichen für Lebendigkeit verstanden, sondern als ein Medium der Äußerung, das in gewisser Weise die Lebendigkeit des Lebendigen selbst noch bewahren kann. Dem Vernehmen einer Stimme muss demnach schon die Differenz innewohnen, die sie von einem nichtlebendigen Klang unterscheidet.

Natürlich kann sich ein Geräusch anhören wie ein menschlicher Schrei oder eine menschliche Stimme durch eine synthetische simuliert werden, aber das entscheidende Kriterium für eine solche Täuschung ist nicht die Zurechenbarkeit auf einen Sprecher, der eventuell nicht ausgemacht werden kann, sondern die Attribuierung, ob etwas mechanisch oder lebendig klingt. Was oder wer eine Stimme hat, und das kann auch wieder ein Nichtlebendiges sein, dem wird Lebendigkeit zugesprochen. Deswegen findet sich die Stimme in dieser Tradition als etwas konzeptionalisiert, dem die Selbstbewegung der lebendigen Bewegtheit derart innewohnt, dass die Stimme als Medium nicht nur ein Medium des Lebendigen, sondern umgekehrt ebenso ein lebendiges Medium ist, in dessen Verfasstheit sich dementsprechend im Unterschied zum Nichtlebendigen das Lebendige selbst äußern kann.

In den traditionellen philosophischen Diskursen über die Stimme kreuzen sich daher Naturphilosophie, Anthropologie, Epistemologie und praktische Philosophie auf eine signifikante Weise, so dass man beinahe von einem Stimm-Humanismus sprechen müsste. Dem Anthropozentrismus der abendländischen Metaphysik korrespondiert deshalb häufig ein Phonozentrismus, wie ihn Jacques Derrida herausgearbeitet hat, der nicht unbedingt in jedem Fall auch ein Logozentrismus sein muss. Dennoch spielt in der Thematisierung der Lebendigkeit der Stimme die entlang eines spezifisch menschlichen Logos getroffene Unterscheidung von Mensch und Tier eine wesentliche Rolle und zwar hinsichtlich der Art

der jeweiligen Selbstbezüglichkeit dieses Lebendigen, die sich vor allem im Phänomen des Sich-sprechen-Hören ausdrückt und damit verbunden in der als »flüchtig« verstandenen Materialität der Stimme. Diese thematisierte Flüchtigkeit scheint in besonderer Weise dazu geeignet zu sein, dem zu korrespondieren, was man als Lebendigkeit anspricht und daher als etwas, das über eine bloße Materialität hinausgeht; und zwar vor allem dann, wenn sich der Entzug dieses Flüchtigen nicht etwa wie bei einem Geräusch nur akustisch darstellen lässt, sondern als ein freiwilliger, selbstgewollter, als ein Entzug eben, der dem Lebendigen gerade darin entspricht, dass es sich über seine Aufzeichnung hinaus behauptet.

Das Faszinosum der Stimme scheint deshalb einen ähnlichen Beobachtungstypus zu provozieren wie die Beobachtung des Lebendigen, wenn zunächst einmal festgestellt werden muss, dass sich das Beobachtete dem herangetragenen Verständnis entgegenstellt, ihm vielleicht sogar widerspricht, oder, wie man so schön sagt, wenn es sich nicht auf den Begriff bringen lässt. Die Stimme als Medium, die man in der engsten Nähe zum tatsächlichen Körper dieses Lebendigen wähnt, lässt sich daher selbst als ein lebendiger Körper anschreiben, der in dem, was man als flüchtig beschreibt, auch ein lebendiges Verhalten nahe legt. Die Stimme lebt also vor allem deshalb, weil sie auch sterben kann. Und das ist etwas völlig anders als das bloße akustische Verklingen eines Geräuschs oder das Vergilben und Vergehen einer Handschrift oder das mitunter tragische Ende einer Festplatte, weil dieser Tod all die Rätsel auf sich ziehen kann, die der Tod des Menschen aufgibt. Da entlang des menschlichen Todes und dem des Tieres die in der abendländischen Anthropologie zentrale Unterscheidung und Abgrenzung von Mensch und Tier verhandelt wird, lohnt es sich, die Konzeptionalisierung dieses Unterschieds anhand der Stimme des Menschen und der des Tieres zu beobachten, um daran das stimmliche Paradigma der Lebendigkeit, jener berühmten *vox viva*, und der damit verbundenen Medialität dieser Stimme genauer nachzeichnen zu können.

#### DER LAUT EINES BESEELTEN WESENS

Die wohl einflussreichste Definition der Stimme findet sich in der Schrift *De anima* von Aristoteles, wo der Bereich des Lebendigen in Abgrenzung zur anorganischen Natur noch einmal unterteilt wird in Lebewesen mit Stimme und Lebewesen ohne Stimme. Der häufig zitierte Satz von Aristoteles, wonach die Stimme ein »gewisser Ton des beseelten Wesens« ist (II, 8, 420b),<sup>1</sup> schließt also zunächst sowohl die Tierseele als auch die Menschenseele ein. In der Stufenleiter der Ge-

schöpfe kennt Aristoteles drei Seelenformen, die jeweils der Pflanze (*anima vegetativa*), dem Tier (*anima sensitiva*) und dem Menschen (*anima rationalis*) eigentümlich aber nicht ausschließlich zugeordnet sind. Das Vermögen der Pflanzenseele ist die »Zeugung« und die »Nahrungsverwertung«, das der Tierseele ist die Leistung der »Sinne«, die sowohl Wahrnehmung einschließt als auch Affekte sowie Begehren und – unterschieden von der äußeren Wahrnehmung – ebenso »Vorstellungen«. Die Menschenseele umfasst alle mentalen Leistungen vom intuitiven Verstand bis zur diskursiven Zergliederung und zum praktischen Abwägen.

Das Phänomen des Schalls grenzt Aristoteles zunächst von der Stimme dadurch ab, dass Schall niemals dort entsteht, »wo nur ein Gegenstand ist«. Was »tönt«, entsteht durch einen »Schlag«, der eine räumliche Bewegung und mindestens zwei Gegenstände voraussetzt. Die Hervorbringung der Stimme hingegen ist insofern selbsttätig, da sie vom Anschlagen der eingeatmeten Luft an die Luftöhre herrührt, einem Organ, dem nach Aristoteles die »tätige Seele« innewohnt. Nicht jeder Laut, den ein Tier oder ein Mensch hervorbringt, ist in diesem Sinne der Stimme zuzuordnen, wie etwa das Husten, da die Tätigkeit vom »Beseelten« ausgehen und immer – das ist entscheidend – in Verbindung mit einer »Vorstellungstätigkeit« stattfinden muss (II, 8, 421a). »Vorstellungen«, die sich von der äußeren Wahrnehmung unterscheiden und in der philosophischen Tradition seit Augustinus dem inneren Sinn (*sensus interior*) zugerechnet werden, gesteht Aristoteles damit auch der Tierseele (*anima sensitiva*) zu. Wahrnehmungen ergeben sich anhand der elementaren Medien wie Äther, Luft oder Flüssigkeiten durch sensorische Einprägungen nach dem Modell von Siegelring und Wachs. Aber erst der »Gemeinsinn«, den Aristoteles im Herz lokalisiert, macht daraus bewusste Wahrnehmungen. Dementsprechend spielt für die Stufenleiter der Geschöpfe die Ausbildung der Sinne und die Unterscheidung von »Blutlosen« und »Bluttieren« eine wesentliche Rolle.<sup>2</sup>

Mit dieser medizinisch-physiologisch inspirierten Deutung überwindet Aristoteles die alte Gleichsetzung von »Hauch« und »Seele«, die etwa noch auf Platons Seelenlehre einen entscheidenden Einfluss hat. Wenn Aristoteles die Stimme einen »bedeutungsvollen Laut« nennt, dann ist das Pneuma der Rede wesentlich als physiologischer Agens gedacht, dem sich nicht einfach ein Phonozentrismus im Sinne einer adäquaten Verkörperung des Logos durch die Stimme unterziehen lässt, wie es Derrida in der unmittelbaren »Verwandtschaft« von Seele und Logos bei Aristoteles gegeben sieht, von der dann die Schrift nur noch ein sekundärer Nachhall wäre (auch wenn die Rezeption der aristotelischen Textstellen anders verlaufen sein mag).<sup>3</sup> Die Nähe von Stimme und Pneuma wird bei

Aristoteles vornehmlich physiologisch durch die Nähe der Organe »Luftröhre«, »Kehle« und »Herz« begründet: »Der Atmung bedarf aber auch die Gegend um das Herz in erster Linie. Deshalb muss mittels des Einatmens die Luft ins Innere eintreten.« (II, 8, 421a)

Entscheidend an dieser Nähe ist die Kopplung des Lebensbegriffs und der Auffassung der Stimme als ein Medium dieses Lebens, dem nicht unbedingt auch ein Logozentrismus in dem Sinne entspricht, dass die Stimme als ein unmittelbares Medium der Wahrheit angesprochen wird. Im Folgenden soll daher das Verhältnis von Stimme und Denken, von Kognition und Medium der Kognition im Hinblick auf diese Lebendigkeit thematisiert werden, und zwar als das zentrale Thema der unterstellten Nähe von Denken und Stimme. Denn in der Unterscheidung von Tierstimme und Menschenstimme spielt zumindest in der modernen Anthropologie immer auch eine dritte Unterscheidung mit, durch die jene Stimme des Menschen aufgrund der Selbstständigkeit der Apparate und der Maschinen bedroht ist – und irgendwann vielleicht sogar zum Verstummen gebracht wird. Spätestens dann nämlich, wenn Maschinen solche Maschinen herstellen und hervorbringen, die deshalb zu einer eigenen Stimme kommen, weil sie nicht mehr die Signatur des Menschen tragen. Was Lebendigkeit hat und eine Seele besitzt und damit den Bereich dessen steuert, dem nach Aristoteles auch eine Stimme zukommt, ist also nicht im vorhinein ausgemacht. Es könnte auch umgekehrt die Frage gestellt werden, ob jener Phonozentrismus der abendländischen Metaphysik, in dem sich die anthropozentrische Gestalt des Menschen als ein Monopol der Produktion aufrichtet, nicht einem Dasein einer Stimme unter anderen Stimmen weichen muss.

Der Unterscheidung von Tierstimme und Menschenstimme soll im Folgenden anhand der Anthropologie Hegels nachgegangen werden, weil diese als Dialektik das evolutionäre Muster abgibt, mit dem sich der Mensch nicht nur einfach als ein Wesen mit einer Stimme setzt, sondern zugleich erklärt und ausweist, warum er dazu die Fähigkeit und die Potenz besitzt. Dabei geht es nicht um die immer auch mitverhandelte christologische Tradition und die antijüdischen Tendenzen, die in Hegels Schriftbegriff eine nicht unwesentliche Rolle spielen, sondern um den von Derrida analysierten Phonozentrismus, der dem Denken, das sich der Stimme wie keinem anderen Medium anvertraut, eine Art Selbstpräsenz verspricht. Gerade diesen Verdacht, bei dem die Wahrheit, der Ursprung, die Autonomie des Bewusstseins mit der Idee der Präsenz verknüpft sind und sich auf diese Weise vor dem unendlichen Aufschub des sinnlichen Zeichens schützen, könnte die Hegelsche Philosophie weit von sich weisen. Denn keine Philosophie wie die Hegelsche hat den Gedanken der sinnlichen Endlichkeit so konsequent

eingeführt, bis hin zu einem völligen Desinteresse an dem, was man Erkenntnistheorie nennt, und den »Geist« ganz und gar als einen »praktischen« aufgefasst. In diesem System ist die Stimme des Menschen keineswegs die adäquate Verkörperung eines unmittelbar bei sich seienden Cogito.

Der »Geist« muss sich »entäußern«. Hegel bietet diesen Gedanken an wie eine Wette, wie einen Potlatsch beinahe: »[...] der Schein selbst ist dem Wesen wesentlich.«<sup>4</sup> Der »Geist« muss sich »preisgeben«, er muss sich »verlieren«, dem Außen vollkommen »anvertrauen«. Denn außerhalb seiner »Erscheinung« ist er nichts. Seine »Tiefe« ist nur so »tief« wie seine Fähigkeit, sein »Wagemut«, sich zu »entäußern«, bis hin zu dem, was man als Diskurs des Opfers und des Todesmutigen kennt oder auch als den der Souveränität, die keineswegs tatsächlich in den Kampf ziehen und den Tod in Kauf nehmen muss, sondern in den komplexen und wechselseitigen Bezügen von Ängsten und Erwartungen die größte Drohung auch gegen sich selbst auszusprechen hat, die rückhaltloseste Geste eben.<sup>5</sup> Schon hier drängt sich also ein ganz anderer Verdacht auf, nämlich auf welche Weise diese Endlichkeit in das System Hegels eingeführt wird, ein bloßer Zwischenschritt womöglich, zwischen zwei Unendlichkeiten, in der die Geschichte als Vermittlung zwischen »Natur« und »Geist« nur ein Übergang bleibt. Alles andere als diese »Tiefe« jedenfalls bleibt für Hegel ein nebulöser »Dünkel« oder eine leere und tautologische Identität.

Hegel wiederholt diesen Gedanken beinahe manisch, der besagt, ich vertraue mich an: meinen Handlungen, meinem Tun, meinen schriftlichen und mündlichen Äußerungen, dem »Objektiven«, und zwar so sehr, dass dieser »Geist« sogar »verwüestet« wird; ich vertraue mich der ganzen Natur an und auch der toten. Im Hegelschen Werk lassen sich Wendungen für diese »Entäußerung« finden, die von einer Herausforderung, einem extremen Schmerz zeugen, als ginge es darum, einen unmöglichen Beweis anzutreten, oder, wie gesagt, eine Wette anzunehmen, die es eben mit der gesamten Welt aufnehmen kann. Der »Geist«, könnte man schlussfolgern, muss in gewisser Weise die Sinnlichkeit und all ihre bedrohlichen Formen der Endlichkeit, die ihm immer wieder seinen Ursprung entziehen, schließlich beeindrucken können, oder genauer: in letzter Konsequenz dem Tod selber eine imponierende Geste entgegenhalten. Eine Geste, die diesen Tod zwar nicht leugnet, ihn aber in den Dienst und die Souveränität des Lebendigen stellt. Oder wie Nietzsche sagt, die ihn nur zu einem weiteren Zeichen unter anderen macht, also letztlich eine Steigerung darstellt für die Macht des Lebens, unter deren potentem Blick selbst der Tod sich als eine Art von Leben zeigen kann.<sup>6</sup> Alles das lässt sich in dem versammeln, was oben als das Lebendige der Stimme und ihrer besonderen Medialität beschrieben wurde.



## DAS HOHE VORRECHT DER STIMME

Das »Leben« oder die »Lebendigkeit« wird in der *Enzyklopädie der philosophischen Wissenschaften* als ein Zusammenhang definiert, bei dem »alle Glieder sich gegenseitig momentane Mittel wie momentane Zwecke sind« (§ 216).<sup>7</sup> Das Leben stellt eine solche Einheit dar, in der alle »Glieder« selbst wieder von dieser Einheit durchdrungen sind und deshalb den Reproduktionsprozess leisten können. Diese Einheit, die den lebenden Körper durchdringt, nennt Hegel mit Bezug auf Aristoteles die »Seele«. Die Seele ist in diesem physiologischen Sinne im Körper präsent, da sie das Leben und die Reproduktion dieses Körpers organisiert. Sie »beseelt« diesen Körper als Zusammenhang, so dass ein abgetrenntes Körperteil ein »totes« ist. Insofern ist sie nach Hegel aber zugleich im Körper verstreut, weil ihre »Gegenwärtigkeit«, ihr »Auseinander«, »für« sie selbst »keine Wahrheit« hat. Das Lebendige ist daher nur der »Prozess« der Erhaltung, der einer anorganischen Natur gegenüber steht. Zwar ist sich das Leben in diesem fortwährenden Assimilierungsprozess gegenüber der anorganischen Natur selbst »gewiss«, aber das Prinzip oder die Organisation dieser Erhaltung ist dem Leben selbst nicht durchsichtig. Ohne »Seele« geht der Körper wieder in den Zustand der anorganischen Natur über, deren »Mächte« »fortwährend auf dem Sprunge« sind, »ihren Prozess im organischen Leibe zu beginnen, und das Leben ist der beständige Kampf dagegen« (§ 219, Zusatz).

Dieser Kampf, den das Leben zum Zweck der Selbsterhaltung führt, zeigt sich nach Hegel in verschiedenen Formen der »Negation«, etwa im »Essen« oder in der »Verdauung«. Der Selbstzweck des Lebens, die Einheit stiftende Relation der Zwecke, ist ein Resultat dieser Negation, in dem sich das Leben als dauerhafter Prozess hervorbringt. Hegel spricht von einem »Zusammenschluss« des Lebens mit sich selbst oder, in den Begriffen der Logik, von einem »tätigen Schluss«. Die »Selbstbezüglichkeit« dieses Zusammenschließens ist daher eine »negative für sich seiende Einheit«. Sie ist »negativ«, weil sie als dauerndes Resultat aus der Negation der anorganischen Natur hervorgeht; sie ist »für sich«, weil die Relation der Zwecke im ganzen lebenden Körper präsent sein muss – jedes Moment des Lebens ist wiederum ein System aus Schlüssen. Insofern hat das Leben eine »Gewissheit«, aber nicht die »Wahrheit« dieser »Gewissheit«. Das Leben ist sich seiner eigenen Reproduktion gewiss, aber es kennt sie nicht. Diese Differenz ist für Hegel zugleich die Differenz von »Körper« und »Seele«, die zusammen »für« das Leben nur eine »unmittelbare« Einheit bilden, keine vermittelte, kontrollierbare Einheit, so dass diese Einheit und das mit ihr einhergehende Leben »endlich« ist und sich »Körper« und »Seele« auch trennen können. Das Sterben ist dem Leben

nicht äußerlich, sondern in der Negation angelegt, mit der es sich gegen die anorganische Natur erhält.

Der Kampf lässt sich also nicht auf Dauer führen, er hinterlässt Spuren, die Negation bestimmt ebenso das Negierte wie das Negierende. Im Hegelschen System ist diese Negation eine »einfache« Negation, deren praktisches Tun nicht noch einmal eingeholt werden kann. Erst eine »Negation der Negation« wäre ein solches Tun, bei dem die produzierte Differenz wiederum einer höheren Einheit angehören würde und von dieser her kontrolliert werden könnte. Dann wäre sich das einzelne Leben nicht nur seiner selbst gewiss, sondern könnte seinen Kampf und seine Erhaltung auf Dauer stellen und wäre auch einer selbstbestimmten Unendlichkeit fähig. Die Dialektik der Natur kündigt insofern schon eine andere Existenz an, nämlich die des »Geistes«, der im Hegelschen System zugleich der Natur vorgängig ist und aus der Natur hervorgeht. Auf die Interpretation der Logik und des logischen Urteils im Sinne von Ur-Teilung (und auch auf die historischen und geopolitischen Implikationen dieser Interpretation) kann hier allerdings nicht näher eingegangen werden. Die »Seele« jedenfalls ist für Hegel in der Tradition von Aristoteles das Organisationsprinzip der lebendigen Reproduktion, nach dessen Suche sich auch die Gentechnologie aufgemacht hat, und das, wenn es gewusst würde, zu einer Selbststeuerung dieses Lebens führen könnte. Da für Hegel die Natur dialektisch verfasst ist, also die Logik des Begriffs in ihr materiell schon gegeben ist, stellt das Leben die »unmittelbare Idee« dar, die sich im Verlauf der historischen Evolution zunehmend expliziert. Dieser Explikationsvorgang betrifft das, was eingangs als »Entäußerung« angedeutet wurde. Die »Seele« bleibt nicht einfach nur »Seele«, sondern »äußert« sich, und nur aufgrund dieser »Äußerung« können wir mehr über sie erfahren.

Nun nennt Hegel die Stimme ein »hohes Vorrecht« der Tiere, etwa im Unterschied zur Pflanzenwelt und natürlich auch zu Tieren ohne Stimme (§ 351, Zusatz). Die Stimme ist der Ausdruck der »Selbstbewegung« des Lebendigen als »Empfindung«, »Selbstgefühl«, »Subjektivität« oder »Seelenhaftigkeit« im Sinne der oben genannten »Selbstgewissheit«, insofern das Tier »in sich für sich selbst« ist. Auch hier argumentiert Hegel sowohl physiologisch als auch dialektisch. Dass das Tier »in sich für sich selbst« ist, äußert sich etwa in der »Wärme«, die der Organismus des Tieres im Unterschied zur »chemischen Wärme« beständig produziert. Eine weitere Äußerung dieses »Fürsichsein« ist die Bewegung im Raum, das »Sichsetzen als die reine eigene Negativität *dieses* Orts und *dieses* Orts usf.« – also die Selbstbewegung im konkreten Sinne. Dieses »Selbst« und diese »Bewegung« hängen für Hegel eng mit der Stimme des Tieres zusammen. Zunächst folgt er dem aristotelischen Denken, wenn er den Klang von der Stimme unterscheidet:

Metalle haben Klang, aber noch nicht Stimme; Stimme ist der geistig gewordene Mechanismus, der sich so selbst äußert. Das Unorganische zeigt seine spezifische Bestimmtheit erst, wenn es dazu sollizitiert, wenn es angeschlagen wird; das Animalische klingt aber aus sich selbst. (§ 351, Zusatz)

Das Tier hat also eine Stimme, weil es sich selbst bewegen kann, im Raum und – wie wir noch sehen werden – auch in der Zeit, wodurch überhaupt erst die »Subjektivität« oder die »Seelenhaftigkeit« entsteht. Der äußeren Bewegung im Raum entspricht nämlich eine innere Bewegung der Zeit. Die innere Selbstbewegung, die bei Aristoteles mit dem Anschlagen des Luftstroms an die Luftröhre beschrieben wurde, identifiziert Hegel mit dem Prozess der Zeit, da sich der »leibliche Organismus« selbst zum »Erzittern« bringt:

Diese Subjektivität für sich ist, ganz abstrakt, der reine Prozess der Zeit, der im konkreten Körper, als die sich realisierende Zeit, das Erzittern und der Ton ist. Der Ton kommt dem Tiere so zu, dass dessen Tätigkeit selbst das Erzitternmachen des leiblichen Organismus ist. Es wird aber dadurch äußerlich nichts verändert, es wird nur bewegt; und die hervorgebrachte Bewegung ist nur die abstrakte reine Erzitterung, wodurch nur Ortsveränderung hervorgebracht wird, die aber ebenso wieder aufgehoben ist, [...]. (§ 351, Zusatz)

Das Thema des »Erzitterns« durchzieht die gesamte Hegelsche Philosophie, häufig im Zusammenhang mit den Themen »Furcht«, »Angst« und »Tod« und markiert den Anfang einer Bewegung, eines Subjektivwerdens und einer Negation. Das Tier bringt demnach seine Stimme selbst hervor, aber so, dass diese Stimme zugleich »verhallt«. Das heißt, in der Stimme drückt sich das »Fürsichsein« des Tieres auf die gleiche Weise aus, nämlich als reine Negativität, wodurch das Lebendige des tierischen Organismus gekennzeichnet ist. Dass die Stimme der Laut eines beseelten Wesens ist, wird also zunächst keineswegs als adäquate Selbstpräsenz eines Cogito verstanden, die mit anderen Äußerungen konkurriert, sondern als physiologische Folge der oben beschriebenen Lebendigkeit. Der »geistig gewordene Mechanismus« ist ein Mechanismus, der sich gegen seine Umwelt erhält und in dieser Differenz zu seiner Umwelt ein »Fürsichsein« ausbildet, in gewisser Weise eine selbstgewisse Maschine. Dieses »Fürsichsein« oder diese »Subjektivität« ist dabei nicht mit dem »Ego« eines »Cogitare« und dessen Identitätsversprechen zu verwechseln, denn die das Lebendige kennzeichnende Negativität ist ebenso eine Selbstdifferenz.

Für die Frage nach dieser Selbstdifferenz und dem Umgang damit ist die Stimme für Hegel deshalb von besonderer Bedeutung, weil sich der Prozess der Selbsterhaltung in der Zeit vollzieht, der ebenso nur negativ wie bei der räumlichen Bewegung nicht »dieser« Ort nicht »diese« Zeit ist. Die Stimme als Laut eines beseelten Wesens fällt daher für Hegel mit der spezifischen Selbstdifferenz des Lebendigen zusammen. Wenn Hegel in diesem Zusammenhang die Stimme als das »Nächste zum Denken« bestimmt, weil hier »die reine Subjektivität gegenständig« wird, »nicht als eine besondere Wirklichkeit, als ein Zustand oder eine Empfindung, sondern im abstrakten Elemente von Raum und Zeit«, dann ist die »Subjektivität« des Tieres durch die tätige Verwandlung von Raum und Zeit gekennzeichnet, die »Herrschaft über die abstrakte Idealität von Zeit und Raum« (§ 351), und Denken als eine Praxis angesprochen, die diese Verwandlung und mit ihr die aufgerufenen Seiten der »Innerlichkeit« und der »Äußerlichkeit« auf eine andere Weise beherrschen kann als die Stimme des Tieres.

Die Selbstbewegung des Lebendigen wird mit dem »freien Erzittern« der Stimme »in sich selbst« und dieses mit der »Subjektivität« identifiziert, und das ist das »hohe Vorrecht« der Tiere. Dieses Vorrecht hängt aber noch mit einem anderen »Vorrecht höherer Naturen« zusammen, nämlich mit dem Schmerzempfinden.<sup>8</sup> »Je höher die Natur«, so Hegel, »desto mehr Unglück empfindet sie«, und desto größer wiederum ist das Bedürfnis, diesen Schmerz zu beseitigen und sich Befriedigung zu verschaffen (§ 359, Zusatz). Dieses Vorrecht, in einem »Widerspruch zu existieren«, ist insofern mit der Stimme verknüpft, als die Stimme des Tieres eine ist, die sich selbst zwar hört, sich also gewiss ist und diese Gewissheit sogar genießen kann, aber gerade im Vergehen auch selbst vergeht. Man könnte sagen, die Tierstimme ist im System Hegels eine, die sich selbst nicht lesen kann, sie bleibt in einem gewissen Sinne reine Negation. Hegel nennt die Stimme das »tätige Gehör«, das im Falle des Tieres sein »Selbst« zur Wahrnehmung bringen kann, etwa als Begierde, Schmerz, Freude oder Zufriedenheit, aber dieses Selbst nicht erfassen kann. Die Selbstbezüglichkeit, die dem Sich-sprechen-Hören beigegeben wird, ist für das Tier nur eine negative, so dass Hegel schließlich sagen kann: »Jedes Tier hat im gewaltsamen Tode eine Stimme, spricht sich als aufgehobenes Selbst aus.« (§ 358, Zusatz) Dem Sterben des Lebendigen korrespondiert in der Stimme auch der Tod dieses Lebendigen, und zwar indem die Negativität dieses Lebendigen eine selbstproduzierte ist, der Tod ein Resultat der Art der Selbstproduktion und damit das Verklingen der Stimme (im Unterschied zu anderen Aufzeichnungen und deren Vergehen) als ein selbstproduziertes und »freiwilliges« verstanden wird.

## DIE RÜCKKEHR DER STIMME

Was für die Tierstimme gilt, lässt sich natürlich auch für die Stimme des Menschen sagen. Auch hier wird die Medialität der Stimme insofern als eine besondere aufgefasst, als sie den »Widerspruch« gewissermaßen aushalten kann. Ihre Existenz, und eine solche wird der Stimme sowohl durch die physiologische Nähe als auch durch die attribuierte Lebendigkeit zugestanden, ist im Grunde dieser Widerspruch selbst. Hegel spricht von einer »unkörperlichen Körperlichkeit«, weil ihr Entstehen zugleich ihr Vergehen sei. In der Hegelschen Naturphilosophie ist die Stimme immer schon auf dem Weg in die »Idealität«. Anders aber als die »Idealität« des Lichts etwa, für Hegel eine abstrakte Identität, verkörpert die Stimme das Wesen der Dialektik, das Prinzip der Aufhebung, das sich im permanenten Übergang befindet und deren Negation restlos ist. Ein Widerspruch bewegt die Stimme, sie ist eine dauernde Selbstnegation, und aus diesem Widerspruch geht noch etwas anderes hervor. Sie ist eben nicht einfach nur Schall. Was also könnte die Arbeit der Dialektik besser darstellen als die Stimme?

Die Szene der Menschenstimme unterscheidet sich nicht von der Tierstimme dadurch, dass der Mensch ein denkendes Wesen ist und seine Stimme deshalb eine vernünftige ist. Auch die Menschenstimme ist durch das Paradigma des Lebendigen bestimmt. Während die Tierstimme verhallt, weil dieses eine Leben an der selben Subjektivität, die es von seiner Umwelt unterscheidet, auch zu Grunde geht, gehört die Stimme des Menschen einem Leben an, das sich selbst »erfassen« kann. Wenn die Stimme der Ort des Subjektiven ist, der Prozessualität und der Zeit, des Schachts und des Inneren, dann leistet die Stimme diese »Selbsterfassung«, indem sie das Wechselspiel von Subjektivem und Objektivem beherbergen kann. Für Hegel ist Bewusstsein nicht etwas, das sich in einem denkenden Subjekt vollzieht, sondern ebenso in den sozialen und historischen Kämpfen, in den objektiven Institutionen, den juristischen und den pädagogischen, in allem, was »objektiver Geist« genannt werden kann. Auch die Stimme des Menschen spricht das »Selbst« als »aufgehobenes« aus, also negativ, und das Ausgesprochene wird, wie Hegel sagt, zu einem »Fremden«. Darin unterscheidet sich die Stimme nicht von anderen Äußerungen, nicht vom Tun und auch nicht von der Schrift. Aber die Ökonomie dieses Fremden, das wie alles in der Hegelschen Philosophie wiederkehrt, ist eine signifikant andere. Denn die Wiederkehr ist in diesem Fall gewissermaßen in nuce schon in Echtzeit mit der »Entfremdung« gegeben. Die Stimme ist das Modell der Wiederkehr, von dem aus sich die Wiederkehr alles Unterbrochenen denken lässt, indem die Unterbrechung selbst verschwindet. Die vorübergehende Unterbrechung der Selbstpräsenz könnte also

von Anfang an von einer viel größeren Geste der Souveränität schon aus dem Weg geräumt sein.

In der frühen Abhandlung *Der Geist des Christentums und sein Schicksal* kommt Hegel im Zusammenhang mit der »Rückkehr« vom Objektiven ins Subjektive auf das christliche Abendmahl zu sprechen, auf das Essen des Brotes, das aus der »objektivierten« Liebe wieder eine »subjektive« mache. Diese Rückkehr vergleicht Hegel mit dem Lesen, das aus einem Toten, einem Objekt, einem Ding wieder ein Subjektives entstehen lässt. Die ursprüngliche Subjektivität ist unterbrochen, ein Tod hat sie beendet, aber das Tote kann »aufgelesen« werden, wieder belebt. Dieses »Essen« ist nicht mehr das bloße negative Tun der Selbsterhaltung, mit dem sich das Tier gegen seine Umwelt erhält, sondern aus diesem »Essen« geht etwas anderes hervor. Wo das Tier im Tod eine Stimme erhält, kann der Mensch einen Schritt weitergehen. Er kann seine Stimme lesen. Auflesen. Das Negierende negieren. »Die Vergleichung«, sagt Hegel, »wäre treffender, wenn das geschriebene Wort aufgelesen [würde], durch das Verstehen als Ding verschwände.«<sup>9</sup> Was also bleibt übrig, außer dem Fremden, das zur Negation der Negation, zur erneuten Aneignung reizt, wenn die Stimme nicht wie ein bloßer Klang verhallt?

Die Stimme lebt, weil sie unendlich wiederzubeleben ist. Aber das, was wiederbelebt werden kann, ist zugleich das, was vielleicht niemals tot war. Das Vertrauen und das Schutzdach, das Derrida anhand seiner Lektüre der Phänomenologie Husserls beschrieben hat,<sup>10</sup> steuert auch die Hegelsche Logik, und zwar dort, wo die Stimme eine Existenzweise ankündigt, die über den Tod der Natur hinausgeht. Wenn Hegel sagt, dass die Stimme ein »solches Materielles erzeugt, in dem die Innerlichkeit des Subjekts durchaus den Charakter der Innerlichkeit behält« (§ 401, Zusatz), dann ist damit nicht nur das zeitliche Verklingen des Tons angesprochen, das in der Hegelschen Ästhetik die Künste hierarchisiert, sondern die Übertragung und die Realisierung dieser Subjektivität im materiellen Prozess selbst. Die Nähe von Denken und Stimme besteht nicht in einem Cogito, das sich auch noch äußern muss, sondern in einer Identifikation des praktischen Tuns und des Prozesses beider. Das Erzittern, der Luftanschlag, die Innerlichkeit, das Subjektive erhält sich in der Stimme als ein Lebendiges, aber nur, wenn das Fremde, das Äußerliche, der Gegenstand auch wieder »aufgelesen«, wenn der Schmerz, das hohe Vorrecht, wieder angeeignet werden kann. Das Fremde muss deshalb so beschaffen sein, dass es der Wiederaneignung nicht im Weg steht. Und genau darin besteht die »höhere Kraft« der Stimme, wie Hegel sagt, noch bevor sie zur »artikulierten Stimme« und zur Sprache wird (§ 401, Zusatz). Diese Ökonomie der Stimme ist zugleich der Umschlagplatz für die gesamte Sprachphilosophie He-

gels, die deshalb in einer tiefen Komplizenschaft mit der anthropozentrischen Logik befangen bleibt. Die Schwierigkeiten, die Hegel mit der operativen Logik eines Leibniz hatte, mit den Hieroglyphen und dem Maschinenwesen, sind bekannt.<sup>11</sup>

Es ist kein Zufall, dass Hegel im Zusammenhang mit der »Innerlichkeit« der Stimme im Paragraphen 401 auf das Klagegeschrei und auf Begräbnislieder zu sprechen kommt, die dem Schmerz eine »Gestalt« geben, während das Tier nur bis zum »Schrei des Schmerzes oder der Freude« kommt. Der Mensch hat Religion, das Tier nicht. Das sich selbst erfassende Leben ist nicht in erster Linie eine Frage eines etwaigen subjektiven Reflexionsvermögens, sondern eine Frage der sozialen Gestalten von der Religion bis zum Staat, in denen sich der »objektive Geist« realisiert. Man hat keine Stimme, sondern man bekommt eine. Das Versprechen der Stimme, ihre restlose Aufhebung, ist das Versprechen, dass sich das Leben selbst erfassen kann in seinen objektiven Gestalten, dass es sie verstehen und auflesen kann, um sich durch die Kontrolle der eigenen Differenzproduktion im Idealfall mit sich selbst zusammenschließen zu können. Oder wie Hegel in der Vorrede zur Rechtsphilosophie ausgeführt hat, dass »Vernünftiges« und »Wirkliches« sich in diesem Zusammenschluss entsprechen. Das berühmte »im Nachhinein« dieses Erfassens, das apriorische Perfekt, wie Heidegger gesagt hat, macht die Stimme zum besonderen Modell dieses sich selbst erfassenden Lebens. Die Nähe des »Geistes« zu sich selbst, die Derrida anhand der Hegelschen Hierarchie der Geistmedien, der Hieroglyphe, dem Symbol, dem arbiträren Zeichen, analysiert hat, ist nicht als die Nähe irgendeiner Hirnaktivität zu verstehen oder als Nähe eines »Geistes«, der noch an den archaischen Hauch der Seele erinnern soll und sich in der sinnlichen Abwesenheit selbst transparent sein kann, sondern dieser »Geist« ist für Hegel durch und durch praktisch. Insofern kann es leicht in die Irre führen, wenn Derrida von dieser Medienhierarchie bei Hegel zugunsten der »Offenbarung« der »Freiheit« des »Geistes« spricht und nicht das »Leben« dieses »Geistes« thematisiert,<sup>12</sup> das schon im Hegelschen Denken der zukünftigen Biopolitik und ihrer Macht über die Reproduktion eine zentrale Stelle anweist und das damit die Stimme des Lebendigen auch erst ab einem bestimmten historischen Moment steuert. Man darf sich nicht von der christologischen Metaphorik leiten lassen, zu der Hegel aus der Konzeption seiner evolutionären Geschichtsphilosophie heraus gezwungen ist, aufgrund welcher der Philosoph das, was »ist«, nicht »leicht« zu einem »tausendjährigen Irrtum« erklären kann (§ 401, Zusatz). Das »Cogito« und dessen Medien der Äußerung bilden längst nicht mehr die Szene des Hegelschen Denkens. Kognition vollzieht sich für Hegel nur anhand jener Kopplung von »Vernunft« und »Wirklichkeit« und bedeutet nichts anderes

als die Ausbildung der sozialen Gestalten des Lebens und ihrer Kämpfe und Untergänge. Daraus wären auch die Konsequenzen für diejenige Stimme zu ziehen, welche die Maschinen und ihr Sozialsystem haben bzw. haben werden.

Umgekehrt bedeutet das Hegelsche Versprechen der Stimme und der restlosen Aufhebung implizit ein soziales Modell, eine soziale Kontrolle, die auf Inklusion abstellt und alles Fremde absorbieren kann. Das Auflesen der eigenen Stimme, das Sich-sprechen-Hören, ist in diesem Sinne kein Ausdruck einer Dualität von »Eigenem« und »Fremdem«, sondern die Stimme des Sozialen, bei der immer zuerst jemand sagen muss, dass diese Stimme kein Klang, sondern »deine« Stimme ist. Die vielen bekannten Wendungen dieser sozialen Stimme, »man redet mit einer Stimme«, »man hat eine Stimme«, »man gibt seine Stimme ab«, »man erhebt die Stimme«, verweisen also eher auf jenes zurechenbare Verhältnis von »Tat« und »Täter«, das Nietzsche in der *Genealogie der Moral* als die Genealogie des »Du sollst« beschrieben hat. Die Tierstimme »verhält« nicht deshalb, weil sie von keinem anderen Tier »gehört« wird, sondern weil ihr kein »Grab« korrespondiert und damit nicht jene erste Gestalt dieser sozialen Stimme, der Religion, die Hegel die »innerste Region« des Geistes nennt. Es ist also nicht lediglich ein denkendes Subjekt, das seine »fremde« Stimme liest und im Nachhinein als die »eigene« konstituiert, sondern was spricht, ist analog zur Stimme des Tieres jene soziale Organisation des Lebens, das in der Sprache und den Gestalten des »objektiven Geistes« nach Hegel ein »höheres Dasein« gefunden hat und sich ebenso analog wie die Reproduktion des animalischen Lebens reproduziert. Man hört sich also keineswegs am meisten selbst zu, »sich« selbst sprechen; sondern man hört die Lebendigkeit der Stimme in einem ganz konkreten biopolitischen Imperativ der Reproduktion. Was übrig bleibt, im Sinne des Ausgeschlossenen, ist Klang. Die Konzeption der »Flüchtigkeit« und der »unkörperlichen Körperlichkeit« der Stimme entspricht insofern sehr genau der Organisation dieses Lebens, aus dem das Außen getilgt ist und für das der Tod nur eine Steigerung seiner reproduktiven Macht darstellt.

Der Phonozentrismus beschränkt sich nicht auf die Idee der Präsenz, in der die abendländische Metaphysik den Modus der Wahrheit gedacht hat, sondern umfasst auch die Art und Weise, wie diese Wahrheit mit dem Begreifen des Lebens gekoppelt ist. Wenn Derrida gegen diesen Phonozentrismus einen allgemeinen Schriftbegriff eingeführt hat, der sich nicht mit dem deckt, was man als Einzelmedium ansprechen könnte, dann ist ein solches Fremdes gemeint, das als Rest erst dasjenige Fremde konstituiert, das sich nicht ohne weiteres aufheben und in einer Kreisbewegung wieder aneignen lässt, ein Rest, den man, verwirrend genug, die Schrift in der Stimme nennen könnte. Diese »irreduzible Nicht-Ge-



genwart«, wie Derrida den Aufschub der allgemeinen Schrift formuliert hat und der sich der »Selbsterfassung« entgegenstellt, lässt sich durchaus auch der Stimme attestieren.<sup>13</sup> Aber das ist dann nicht das Rätsel der Stimme, das diese in besonderer Weise vor anderen Aufzeichnungsmedien privilegiert. Natürlich ist die Stimme nicht »flüchtig«, sie entzieht sich nicht und markiert jederzeit die Differenz, die sich das Denken der Kontrolle und der Logik der Dualität des »Eigenen« und »Fremden« als unendlich klein und beherrschbar vorstellt. Vielleicht aber müsste man, statt dem Logozentrismus, den Derrida in der Privilegierung der Stimme am Werk gesehen hat, der *Aufhebung der Ökonomie* von Georges Bataille folgend, einem grundlegenden Biozentrismus nachfragen.

- 1 Aristoteles: Über die Seele, Übersetzung nach W. Theiler, Griechisch-Deutsch, hg. v. Horst Seidl, Hamburg 1995, S. 111 f.
- 2 Vgl. dazu ausführlich bei Hubertus Busche: Die Seele als System. Aristoteles' Wissenschaft von der Psyche, Hamburg 2001, S. 13–17.
- 3 Jacques Derrida: Grammatologie, Frankfurt/M. 1983, S. 24 f.
- 4 Georg Wilhelm Friedrich Hegel: Vorlesungen über die Ästhetik I, Bd. 13, Werkausgabe in 20 Bänden auf der Grundlage der Werke von 1832–1845 neu ed. Ausgabe, redegirt v. Eva Moldauer u. Karl Markus Michel, Frankfurt/M., S. 21.
- 5 Vgl. dazu die Interpretation der Souveränität im Unterschied zur Herrschaft bei Hobbes von Michel Foucault: In Verteidigung der Gesellschaft, Frankfurt/M. 2001, S. 37–57.
- 6 Vgl. dazu Josef Simon: Philosophie des Zeichens, Berlin/New York 1989, S. 206–211.
- 7 Im Folgenden wird Hegels *Enzyklopädie der philosophischen Wissenschaften* unter Angabe der Paragraphen zitiert und zwar aus der Werkausgabe Bd. 8–10 (Anm. 4).
- 8 Peter Singer hat darauf eine »praktische« Ethik zu gründen versucht; vgl. Peter Singer: Praktische Ethik, Stuttgart 1984, S. 26 ff.
- 9 Hegel: [Der Geist des Christentums und sein Schicksal], Werkausgabe Bd. 1 (Anm. 4), S. 367.
- 10 Jacques Derrida: Die Stimme und das Phänomen, Frankfurt/M. 1979, S. 51–67.
- 11 Jacques Derrida: Der Schacht und die Pyramide, in: ders.: Randgänge der Philosophie, Wien 1988, S. 85–118.
- 12 Ebd., S. 100.
- 13 Vgl. Bernhard Waldenfels: Deutsch-Französische Gedankengänge, Frankfurt/M. 1995, S. 104.

Erika Linz  
DIE REFLEXIVITÄT DER STIMME

Wir hören uns sprechen, aber wir lesen uns nicht schreiben. Der Sprachgebrauch macht uns hier auf eine Differenz zwischen der Stimme und der Schrift aufmerksam, die die unterschiedliche Kopplung von produktiven und rezeptiven Aspekten in den beiden Medien betrifft. Zwar können wir uns schreiben sehen, doch im konkreten Vollzug der Schreibbewegung lesen wir nicht gleichzeitig, was wir schreiben. Nur beim Sprechen schließt das Hören zugleich ein semantisierendes Zuhören mit ein. Die simultane Verschränkung von *sinngenerierender* Sprachproduktion und *sinnentnehmender* Sprachrezeption scheint allein dem Sprechen vorbehalten zu sein.

In der sprachphilosophischen Tradition hat eben diese Besonderheit der »rezeptiv-produktiven Zweiseitigkeit«<sup>1</sup> des Sprechens den über lange Zeit exponierten Status der Stimme gegenüber der Schrift begründet. Neben der materiellen Eigenschaft der zeitlichen Negativität des Tons, also seiner Flüchtigkeit und seiner linearen Sukzessivität, war es insbesondere die prozessuale Eigenschaft der Kopplung von Sprechen und Hören im stimmlichen Entäußerungsakt, durch die der Stimme spätestens mit dem Idealismus eine besondere erkenntnistheoretische Funktion verliehen wurde. Vor dem Hintergrund dieser Tradition und ihrer Dekonstruktion durch Derrida greifen die folgenden Ausführungen die Frage auf, inwieweit der Stimme in Bezug auf ihre prozessuale Reflexivität auch nach der Kritik am Phonozentrismus noch eine erkenntnistheoretische Relevanz beigemessen werden kann. Gerade aus kognitions- und medientheoretischer Perspektive könnte die Diskussion um mögliche epistemische Effekte der Stimme etwas vorschnell beiseite gelegt worden sein. Im philosophischen und kulturwissenschaftlichen Diskurs verlor sie mit der Derridaschen Kritik für viele ihre Legitimation. Im kognitions- und kommunikationswissenschaftlichen Diskurs hingegen wurde sie kaum erst zur Kenntnis genommen. Unverbrüchlich lebt in den kognitiven Wissenschaften – in Psycho- und Neurolinguistik ebenso wie in weiten Teilen der Sprach- und Kommunikationswissenschaften – die Auffassung von Stimme als Übertragungsmedium fort: Stimme wird als Instrument zur nachträglichen Kommunikation intern formulierter Redeinhalte verstanden. Entsprechend bleiben mögliche kognitive Rückwirkungen des Sprechens und seiner reflexiven Prozessualität weitgehend aus der Modellbildung ausgeblendet. Sprachproduktion und -rezeption sind immer noch zwei überwiegend getrennte Forschungsbereiche, die nicht systematisch in ihrer Wechselwirkung fokussiert werden.

Die im medientheoretischen Kontext zu beobachtenden Versuche einer Rehabilitation der Stimme setzen primär auf prädiskursiver Ebene an und thematisieren vor allem die nicht-semiotischen Aspekte der Stimme.<sup>2</sup> Wie aber steht es um die semiotische Relevanz der Stimme im Sprechen? Unter der Voraussetzung, dass das Problem der erkenntnistheoretischen Dimension des Sich-sprechen-Hörens und der selbstnegierenden Materialität der Stimme aus dem philosophiegeschichtlichen Diskurskontext um eine Hierarchie unterschiedlicher medialer Symbolisierungssysteme und damit verbundener ästhetischer Ausdrucksformen gelöst wird, lässt sich auch in semiotisch-prädikativer Hinsicht für eine Rehabilitation der Stimme und ihrer epistemischen Funktion argumentieren. Gerade der Blick auf die reflexive Prozessierungsform der Stimme kann – sofern er von den phonozentrischen Implikationen befreit wird – dazu beitragen, die Medialitätsabhängigkeit kognitiver Systeme ebenso wie aktueller Sinngenerierungsprozesse aufzuzeigen. Verdeutlicht werden soll dies anhand zweier Thesen: zum einen anhand der Annahme, dass das Sprechen aufgrund seiner Verschränkung von produktiven und rezeptiven Prozessen als metaleptisches Verfahren aufgefasst werden kann und zum anderen anhand der These, dass sich die Medialität der Stimme trotz ihrer Flüchtigkeit in das kognitive System einschreibt und sich somit auch auf der Ebene vermeintlich amedialer mentaler Prozesse manifestiert.

### 1. SICHSPRECHEN-HÖREN ALS METALEPTISCHES VERFAHREN

In seiner Husserl-Lektüre hat Derrida die phänomenologische »Macht der Stimme«<sup>3</sup> dahingehend bestimmt, dass sie der Selbstpräsenz »ein Schutzdach bietet«: »Das Subjekt muß nicht aus sich selbst heraustreten, um unmittelbar von seiner Ausdruckstätigkeit affiziert zu werden. Meine Worte sind ›lebendig‹, weil sie mich gar nicht erst zu verlassen scheinen: [...] nicht aufhören, mir zuzugehören und mir ›ohne Dazwischentreten‹ disponibel zu bleiben.«<sup>4</sup>

Versuche, dieser Vorstellung einer »absoluten Nähe« und einer »unmittelbaren Präsenz«<sup>5</sup> von Signifikat und Subjekt im Sprechen entgegenzuwirken, finden sich aber nicht erst seit Derrida. In einer Verkürzung der Derridaschen Phonozentrismus-Kritik ist häufig in Vergessenheit geraten, dass bereits Hegel – und mit ihm Humboldt – die Bedeutung medialer Entäußerungshandlungen für die Genese kognitiver Gehalte und eines subjektiven Bewusstseins, das sich auf sie bezieht, hervorgehoben hat. Die Stimme wird unter dem Aspekt, im Sprechen immer auch zugleich sich selber zu hören, von Hegel gerade nicht als Medium der unmittelbaren Selbstpräsenz herangezogen; sie wird vielmehr primär zu einem

Medium der Entzweiung des Subjektes mit sich selbst.<sup>6</sup> Über die stimmliche Entäußerung und die dadurch möglich werdende sinnliche Rezeption des Entäußerten wird ein Prozess der Vergegenständlichung initiiert, der allererst ein »Unterscheiden des Denkenden von dem Gedachten«<sup>7</sup> erlaubt: »Wir wissen von unseren Gedanken nur dann, haben nur dann bestimmte, wirkliche Gedanken, wenn wir ihnen die Form der *Gegenständlichkeit*, des *Unterschiedenseins* von unserer *Innerlichkeit*, also die Gestalt der *Äußerlichkeit* geben [...]«<sup>8</sup>

Noch dezidierter wird dieser Gedanke einer notwendigen Objektivierung mentaler Prozesse in der stimmlichen Verlautbarung von Humboldt entfaltet: »Die intellektuelle Thätigkeit« bleibt – so Humboldt – »gewissermassen spurlos vorübergehend«, solange sie nicht »durch den Laut in der Rede äusserlich und wahrnehmbar für die Sinne« wird.<sup>9</sup> Erst über den Umweg der Entäußerung in einem sinnlich wahrnehmbaren Medium verlieren mentale Prozesse ihren transitorischen Charakter und werden als distinkte Einheiten für ein Bewusstsein, das sich zugleich über seine Entäußerungshandlungen konstituiert, erfahrbar. Humboldt hat diese Bewegung näherhin als eine in jedem Sprechakt »stillschweigend immer vorgehende Versetzung« der subjektiven Vorstellung »in zum Subject zurückkehrende Objektivität« bestimmt, ohne die »die Bildung des Begriffs, mithin alles wahre Denken, unmöglich [ist]«.<sup>10</sup> Und gleichermaßen bildet auch für Hegel die stimmliche Mediatisierung die Voraussetzung dafür, dass kognitive Inhalte zu »fassbaren«, d. h. vom kognitiven Prozess geschiedenen Erkenntnisobjekten für ein erkennendes Bewusstsein werden:

Der für die bestimmten Vorstellungen sich weiter artikulierende Ton, die *Rede*, und ihr System, die *Sprache*, gibt den Empfindungen, Anschauungen, Vorstellungen, ein zweites, höheres als ihr unmittelbares Dasein, überhaupt eine Existenz, die im *Reiche des Vorstellens* gilt.<sup>11</sup>

Im Unterschied zu Kant ist für Hegel und Humboldt die Bildung distinkter Begriffe nicht allein an die Notwendigkeit der Versinnlichung gebunden, sondern bedarf darüber hinaus einer Manifestation in einem »äußerlichen Elemente«.<sup>12</sup> Durch die sensorische Wahrnehmung der eigenen Vorstellungen als äußere Objekte wird erst jener differenzgenerierende Akt vollzogen, ohne den »das Denken nicht zur Deutlichkeit gelangen, die Vorstellung nicht zum Begriff werden [kann].«<sup>13</sup>

Indem [...] das geistige Streben sich Bahn durch die Lippen bricht, kehrt das Erzeugniss desselben zum eignen Ohre zurück; die subjective innere Handlung wird als äussres Object wieder aufgenommen. [...] Denn keine

Gattung der Vorstellungen kann als ein reines Beschauen eines schon vorhandenen Gegenstandes betrachtet werden. Die Thätigkeit der Sinne muss sich mit der inneren Handlung des Geistes synthetisch verbinden, und aus dieser Verbindung reisst sich die Vorstellung los, wird, der subjectiven Kraft gegenüber, zum Object, und kehrt, als solches aufs neue wahrgenommen, in jene zurück. Ohne daher irgend auf die Mittheilung zwischen Menschen und Menschen zu sehn, ist das Sprechen eine nothwendige Bedingung des Denkens des Einzelnen in abgeschlossener Einsamkeit.<sup>14</sup>

Für Humboldt wie für Hegel wird also die doppelte Bewegung einer medialen Entäußerung kognitiver Prozesse und der damit parallel einhergehenden Selbstlektüre des Entäußerten zur Voraussetzung für die Konstituierung von Bewusstseinsinhalten. Das Sich-selber-Hören im Sprechen stellt somit auch für Hegel keine Erfahrung der unmittelbaren Selbstpräsenz, sondern einen immer schon vermittelten, d. h. mediatisierten Prozess dar. Erst durch den sich im Sprechen vollziehenden Akt einer medial induzierten Distanzierung des Subjektes von sich selbst verlieren mentale Prozesse ihre Unmittelbarkeit und können eine – wie Hegel sagt – distinkte ›Existenz im Reiche des Vorstellens‹ gewinnen. Mit Jäger lässt sich diese Angewiesenheit der »intellektuellen Thätigkeit« auf Mediatisierungsprozesse genereller als mediales »Spurtheorem« pointieren:

Erst dort, wo das Denken sich in Sprachzeichen auf die *Spur* seiner eigenen Aktivität zu richten vermag, entsteht semantisch distinkter Sinn und damit zugleich auch die Differenz von Selbst- und Fremdreferenz. [...] Das nur scheinbar ›interne‹ mentale System vermag sich nur über seine ›externe‹ Zeichenspur [...] als mentales System zu konstituieren.<sup>15</sup>

Inwieweit für diesen Prozess der medialen Spurgenerierung die Tatsache von besonderer Bedeutung ist, dass es sich in der stimmlichen Verlautbarung generell um höchst flüchtige Spuren handelt, wird noch zu klären sein. Zunächst lässt sich festhalten, dass vor dem Hintergrund dieser spurtheoretischen Externalisierungsbedingung dem Sprechen insofern eine wesentliche epistemische Funktion beigemessen werden kann, als es sich hierbei ohne Zweifel um eine der zentralen Prozessierungsformen handelt, in denen sich eine mediale Externalisierung und parallele Relektüre der externen Zeichenspuren vollzieht.

Darüber hinaus stellt sich die Frage, inwiefern die sinnkonstituierende Funktion von Selbstlektüreprozessen nicht allein in erkenntnislogischer Hinsicht

relevant ist, sondern ein generelles Prinzip der sprachlichen Performanz markiert, das auch im jedesmaligen Sprechen seine Wirkung entfaltet. Interessanterweise findet die Bedeutung eigenrezeptiver Prozesse in der Sprachproduktion selbst in der Psycholinguistik, in der Sprachrezeption und Sprachproduktion jahrzehntelang als vollständig separate Gegenstandsbereiche behandelt wurden, zunehmend größere Beachtung: »The person to whom we listen most is ourself«, so formuliert etwa der führende Psycholinguist Willem Levelt die dort neu entdeckte Einsicht in die Kopplung von produktiven und rezeptiven Momenten beim Sprechen.<sup>16</sup> Allerdings bleibt auch in den jüngeren Sprachproduktionsmodellen die Rolle rezeptiver Feedbackprozesse im wesentlichen noch auf eine Kontrollfunktion über die korrekte Ausführung vorgefertigter Redeintentionen beschränkt. Die psycholinguistischen Modelle gehen zwar selbst zunehmend von einer inkrementellen Verarbeitungsweise, d. h. von einer erst im Sprechverlauf voranschreitenden Ausformulierung der geplanten Äußerung, aus. Doch der Redeinhalt wird in diesen Modellen nach wie vor als vorab fertig geplanter seiner Äußerung vorangestellt. Entsprechend bezieht sich das sogenannte »Self-Monitoring« in der Rede vor allem auf die Überprüfung etwaiger Fehler in der sprachlichen Realisierung.<sup>17</sup>

Auf die Fragwürdigkeit solch eines zweiphasigen Sprachproduktionskonzeptes scheint schon Kleist mit seiner Abhandlung *Über die allmähliche Verfertigung des Gedankens beim Sprechen* aufmerksam gemacht zu haben. Wenn, wie bislang argumentiert wurde, semantisch distinkter Sinn erst über den Umweg einer Lektüre der eigenen materiellen Zeichenspuren generiert wird, kann der einzelne Sprechakt nicht länger als Enkodierung vorgefertigter Propositionen modelliert werden. Der Gedanke geht der Rede nicht voraus, sondern – wie Kleist es in parodistischer Abwandlung des Sprichwortes »l'appétit vient en mangeant« formuliert –: »l'idée vient en parlant«.<sup>18</sup> Das Sprechen dient demnach nicht nur der nachträglichen Kommunikation präsprachlich gebildeter Propositionen, sondern muss vielmehr als eine Prozessierungsform angesehen werden, die über die Entäußerung in der Rede und eine damit verbundene Relektüre allererst zur Konstituierung propositionaler Gehalte als Bewusstseinsinhalte beiträgt. Der hier vertretenen These zufolge fungiert die Stimme somit weniger als Medium der Selbstaffektion denn als Medium der Selbsttranskription. Transkriptiv verfährt das Sprechen insofern, als der Sprecher vermittelt über die Wahrnehmung der eigenen Rede seine scheinbar ursprüngliche Redeabsicht auch für sich selber erst in eine lesbare Semantik überführt.<sup>19</sup>

Unter Rückgriff auf die Tradition der Rhetorik lässt sich diese Struktur der nachträglichen Erzeugung einer vorgängigen Redeintention als metaleptisches Verfahren charakterisieren. Die rhetorische Figur der Metalepsis bezeichnet eine

Umkehrung der metonymischen Ursache-Wirkung-Relation: »sie stellt« – um mit Bettine Menke zu sprechen – »etwas als Wirkung dar, um eine Ursache vorzusetzen und nachträglich als vorausliegende und vorgängig begründende vorzustellen«, d. h. sie generiert die Ursache als »rhetorischen Effekt«.<sup>20</sup> Wesentlich für die Figur der Metalepsis ist zudem ihr prozessualer Charakter, denn sie bezieht sich weniger auf eine Relation zwischen zwei sprachlichen Zeichen als vielmehr auf einen »semantischen Prozeß«.<sup>21</sup> Als »Redeweise, die schrittweise zu dem vordringt, was sie zeigt«,<sup>22</sup> scheint die rhetorische Figur der Metalepsis geradezu auf die hier postulierte rekursive Prozessierungsform des Sprechens hinzuweisen.

Gestützt wird die Annahme einer metaleptischen Struktur des Sprechens auch durch Befunde aus dem Bereich der neurowissenschaftlichen Bewusstseinsforschung. In einer Vielzahl von Wahrnehmungsexperimenten konnte gezeigt werden, dass die subjektive Abfolge bewusster Erfahrungen nicht notwendigerweise mit der objektiven Reihenfolge der für die subjektiven Erfahrungen verantwortlichen Hirnprozesse korreliert.<sup>23</sup> Vielmehr deutet eine zunehmende Zahl von Befunden darauf hin, dass Bewusstheit ein nachrangiges Phänomen ist, d. h. dass Bewusstseinsinhalte durch zeitlich vorausliegende neuronale Prozesse erzeugt werden. Auch im Hinblick auf die Initiierung von Handlungen scheinen die als handlungsauslösende Impulse wahrgenommenen Erfahrungen metaleptische Effekte zu sein, die im subjektiven Bewusstsein nachträglich vordatiert werden. Veranschaulichen lässt sich dieser Effekt anhand eines illustrierenden Beispiels von Benjamin Libet, einem der führenden Neurophysiologen auf diesem Gebiet der zeitlichen Verarbeitung: Ein trainierter Sprinter startet 100 msec oder weniger nach dem Startschussignal und beginnt damit zu laufen, bevor der Startschuss (nach ca. 200–250 msec.) ins Bewusstsein dringt. Gleichwohl wird er aber hinterher berichten, das Schussignal gehört zu haben, bevor er sich in Bewegung setzte.<sup>24</sup> In einer Vielzahl von Experimenten, in denen u. a. die Methode der elektrischen Hirnstimulation eingesetzt wurde, konnten Libet et al. einerseits zeigen, dass das subjektive Zeiterleben nicht mit der deutlichen neuronalen Verzögerung vom Moment der Einwirkung eines sensorischen Reizes bis zu dessen bewusster Wahrnehmung übereinstimmt, sondern rückwirkend an den Zeitpunkt der Reizeinwirkung angepasst wird. Zum anderen lieferten die Experimente empirische Evidenz für die Annahme, dass der Beginn einer willkürlichen Handlung auf zerebraler Ebene deutlich vor der Bewusstwerdung der Handlungsintention liegt.<sup>25</sup>

Bezogen auf den hier interessierenden Akt des Sprechens haben etwa die Neuropsychologen Halligan und Oakley aus diesem Mechanismus einer subjektiven Umdeutung der Wirkung eines neuronalen Prozesses in dessen Initialimpuls die Annahme abgeleitet, dass »almost certainly you'll have to wait until you hear

your words spoken before you know what they are.«<sup>26</sup> Vor dem Hintergrund der bisherigen Argumentation kann ihre Annahme dahingehend zugespitzt werden, dass die nachträgliche Erschließung der eigenen Äußerungsabsicht nicht nur den sprachlichen Enkodierungsprozess, sondern ebenso die Generierung des propositionalen Äußerungsgehaltes betrifft. Auch der propositionale Gehalt wird erst im Verlaufe des Äußerungsaktes ausdifferenziert. »Erst denken, dann reden« – jene gängige Auffassung, nach der die Planung des Redeinhaltes dem Versprachlichungsprozess vorausgeht, ließe sich somit als metaleptischer Effekt verstehen, der in einer retrospektiven Interpretation des Ergebnisses eines Äußerungsaktes als dessen Auslöser gründet.

Wenn, wie hier vorgeschlagen, das Sprechen als transkriptives Verfahren bestimmt wird, so ist damit also die These verbunden, dass die als vorgängig erlebte Redeintention der Rede nicht vorausliegt, sondern ein Produkt der nachträglichen medialen Transkription darstellt. Erst über die Lektüre der selbst generierten Zeichenspuren wird die vorprädikative Redeabsicht in eine distinkte Semantik transkribiert.<sup>27</sup> Die selbstinduzierende Wirkung der Rede könnte dabei wesentlich darin bestehen, mit dem Sprechverlauf den Möglichkeitsraum einer Fortentwicklung der Rede stetig weiter zu einzugrenzen und den Sprecher dadurch auf eine Fokussierung zu zwingen, die in einer Art radialer Struktur immer enger auf bestimmte Fluchtpunkte hin zuläuft.<sup>28</sup> Der Determinierungszwang wird dabei noch insofern durch die fluide Medialität der Stimme verstärkt, als diese die Option des Verharrens und der Korrektur stark einschränkt und statt dessen eher einen vorantreibenden Effekt erzeugt.

Nun ist die bislang skizzierte epistemische Funktion des Sprechens nicht allein auf die gesprochene Rede beschränkt, denn weder die Verschränkung von produktiven und rezeptiven Prozessen noch die damit verbundene metaleptische Form der Sinngenerierung sind ein allein der Stimme zuschreibbares Privileg. Gleichwohl lassen sich gerade in Bezug auf diesen Prozessierungsaspekt markante Differenzen zwischen unterschiedlichen symbolischen Praktiken festmachen: Beispielsweise scheint ein wesentlicher Unterschied der mündlichen Äußerung etwa zum Schreiben oder zum Malen darin zu bestehen, dass die rekursive Kopplung von sinngenerierender Produktion und sinnentnehmender Rezeption im Sprechen sehr viel enger und kontinuierlicher erfolgt. So sind für das Schreiben zwar sowohl motorisches Feedback als auch die visuelle Wahrnehmung des Geschriebenen unabdingbar, nicht aber die Fähigkeit einer re-semantisierenden Lektüre. Der Erwerb der Lesekompetenz ist sowohl in ontogenetischer als auch in kulturhistorischer Hinsicht nicht direkt mit der Schreibkompetenz



verknüpft. Ähnliches gilt auch für andere mediale Praktiken wie die Malerei oder die Musik, bei der produktive und rezeptive Kompetenz, also die Fähigkeit zur Zeichenproduktion bzw. -interpretation, häufig völlig separat voneinander ausgebildet werden. Deutliche Unterschiede zeigen sich aber auch auf der performativen Ebene der jeweiligen Prozessierung. So vollziehen sich etwa beim Schreiben Lese- und Schreibphasen nicht simultan, sondern in der Regel alternierend. Im Gegensatz dazu ist das Sprechen von Beginn an durch eine rekursive Parallelführung von Produktions- und Lektüreprozess charakterisiert. In dieser engen Kopplung von Zeichenproduktion und sinngenerierender Zeichenlektüre sowohl im Erwerbsprozess als auch in der konkreten Prozessierung könnte eines der Differenzmerkmale des Sprechens zu anderen symbolischen Praktiken sprachlicher und nicht-sprachlicher Art liegen.

## 2. SICH-SPRECHEN-HÖREN ALS DIFFERENZWAHRNEHMUNG

Es ist eben diese enge Bindung beider Aspekte im Sich-sprechen-Hören, die – um noch einmal auf Derridas Husserl-Kritik zurückzugreifen – dazu verleitet hat, das Erleben einer Nähe zu sich selbst im Sprechen in »die absolute Reduktion des Raumes«<sup>29</sup> münden zu lassen: »Als reine Selbst-Affektion scheint« – so Derrida – »die Operation des Sich-sprechen-Hörens bis auf die Innenseite der Oberfläche des eigenen Körpers reduziert zu sein.«<sup>30</sup> Wie steht es also um die der Stimme zugewiesene besondere Disposition zur Selbstaffektion, zur unmittelbaren Präsenz, die immer wieder angeführt worden ist, um ihre Ausnahmequalität und ihr Primat vor anderen Ausdrucksformen unter Beweis zu stellen? Inwieweit ist in den Akt des Sich-sprechen-Hörens trotz seiner scheinbaren Innerlichkeit das Moment eines Außen, einer Differenz immer schon eingeschrieben?

Versucht man sich dieser Frage aus der Perspektive der Prozessierung zu nähern, so lassen sich mindestens drei Aspekte unterscheiden, über die sich eine epistemisch relevante Differenzwahrnehmung für das sprechende Subjekt vollzieht:

### 1. Zeitliche Differenz zwischen Sprechproduktion und -rezeption

Zunächst einmal erfolgen Lautproduktion und Wahrnehmung nicht simultan, sondern zeitverschoben. Der Schall erreicht das Innenohr mit einer Verzögerung von 2 msec. Für das Erkennen einer zeitlichen Aufeinanderfolge zweier unterschiedlicher akustischer Reize – die sogenannte zeitliche Ordnungsschwelle, eine unabdingbare Voraussetzung für die Rezeption gesprochener Sprache – bedarf es

bereits 30 bis 50 msec. Bis ein akustisches Signal Gegenstand einer bewussten Wahrnehmung wird, vergehen sogar zwischen 200 und 500 msec.<sup>31</sup> Sprechen und Sich-sprechen-hören sind also immer schon durch einen zeitlichen Aufschub voneinander geschieden.

## *2. Differenz zwischen kinästhetischer Innen- und auditiver Außenwahrnehmung der Sprechhandlung*

Wesentlicher in kognitiver Hinsicht ist aber noch die Kopplung zweier unterschiedlicher Feedbackschleifen im Sprechen. Die »Operation des Sich-sprechen-Hörens« ist deshalb nicht auf die »Innenseite der Oberfläche des eigenen Körpers«<sup>32</sup> reduziert, weil die Erkenntnis, dass ich es bin, die spricht, gerade die Kopplung eines internen kinästhetischen und eines externen sensorischen Feedbacks voraussetzt. Jede motorische Handlung ist von einer Vielzahl propriozeptiver und kinästhetischer Wahrnehmungen begleitet, die eine notwendige Bedingung der motorischen Koordination und die Grundlage der Ausbildung und Speicherung motorischer Programme darstellen. Beim Sprechen wird dieses Feedback nun durch eine zweite Rezeptionsschleife der über das Außenohr vermittelten akustischen Reize ergänzt.<sup>33</sup> Die Besonderheit der auditiven Außenschleife besteht, um mit Mead zu sprechen, darin, dass sie es erlaubt, »auf den eigenen Reiz so zu reagieren, wie auf den, der von anderen ausgelöst wird.«<sup>34</sup> Durch die Verknüpfung der beiden qualitativ differenten Rezeptionsprozesse wird es überhaupt erst möglich, eine Zuordnung zwischen der Sprechhandlung und dem Hörereignis herzustellen und das Gehörte als das selbst Gesprochene zu identifizieren. Die Kopplung von motorischem und auditivem Feedback bildet die Voraussetzung für die Rückbeziehung der auditiven Wahrnehmung auf die eigene Ausdrucksbewegung, also für die Ausbildung dessen, was Mead die »vokale Geste« genannt hat.

Wenn Derrida darauf hinweist, dass die »Selbstaffektion keine Erfahrungsmodalität [ist], die ein bereits zuvor als Selbst (autos) verfaßtes Seiendes charakterisierte«, so lässt sich somit aus der Perspektive der Prozessierung präzisieren, inwiefern die stimmliche Selbstaffektion »das Selbst als Beziehung zu sich in der Differenz mit sich, das Selbst als das Nicht-Identische hervor[bringt].«<sup>35</sup> Die Selbstaffektion der Stimme besteht eben darin, dass wir in der stimmlichen Verlautbarung in uns selbst jene Reaktionen hervorrufen, die wir auch in anderen Personen auslösen und »damit die Haltungen anderer Personen in unser eigenes Verhalten herein[nehmen].«<sup>36</sup> Es ist diese Differenz zwischen Innen- und Außenwahrnehmung der eigenen Sprechhandlung, die es dem Sprecher erlaubt, das Gehörte als von ihm selbst Produziertes zu erkennen und so das Fremde als das Eigene wahrzunehmen.

Es ist diese Differenz zwischen Innen- und Außenwahrnehmung der eigenen Sprechhandlung, die es dem Sprecher erlaubt, das Gehörte als von ihm selbst Produziertes zu erkennen und so das Fremde als das Eigene wahrzunehmen.

### 3. Differenz zwischen Eigen- und Fremdwahrnehmung

Bereits Humboldt hat darauf insistiert, dass die Außenwahrnehmung der eigenen Sprechhandlung zwar eine notwendige, aber noch keine hinreichende Differenzenerfahrung darstellt. Der Gedanke erhält vielmehr »Bestimmtheit und Klarheit« erst »durch das Zurückstrahlen aus einer fremden Denkkraft.«<sup>37</sup> Das Sichsprechen-Hören ist von Beginn an kein monologischer Akt, sondern ein Prozess, der immer auch den Adressaten der Rede in die Wahrnehmung der eigenen Rede mit einbezieht. Diese zweite Entäußerungsschleife über den anderen Hörer markiert nach Mead auch die notwendige Bewegung, über die allein sich der Übergang von der vokalen Geste zum signifikanten Symbol vollziehen kann. Es bedarf einer weiteren Kopplung, der der Selbstwahrnehmung mit der Reaktion des anderen, um die Reaktion als Reaktion auf die eigene Sprechhandlung deuten zu lernen und damit sich selbst in der Differenz zu anderen als sinnsetzendes Subjekt zu erkennen.<sup>38</sup>

Die vermeintliche Nähe des Subjektes zu sich selbst wird also bereits im Sprechen in mehrfacher Hinsicht durchbrochen. Wenn aber das Sprechen im Hinblick auf seine Prozessierung ebenso als eine differenzgenerierende mediale Externalisierungsform betrachtet werden muss wie etwa das Schreiben, bleibt gleichwohl die Frage offen, inwieweit die Rede aufgrund ihrer spezifischen fluiden, verlöschenden Materialität eine besondere Form der Innerlichkeit und der Unmittelbarkeit für sich beanspruchen kann.

### 3. DIE MEDIALITÄT DER ER-INNERUNG

Besonders Hegel hebt die reine Zeitlichkeit des Tons und die daraus resultierende »ideelle, [...] sozusagen unkörperliche Leiblichkeit«<sup>39</sup> der Stimme hervor. Die Stimme erzeugt »ein solches Materielles [...], in welchem die Innerlichkeit des Subjektes durchaus den Charakter der Innerlichkeit behält, [...] da das Sichverbreiten des Tones ebensosehr sein Verschwinden ist.«<sup>40</sup> In semiologischer Hinsicht erlaubt das »Wort als tönendes«<sup>41</sup> einerseits, die »Vorstellungen in einem äußerlichen Elemente zu manifestieren.«<sup>42</sup> Andererseits ermöglicht die »abstrakte,

d. h. nur *vernichtende* Negativität<sup>43</sup> des Tons, dass das reproduzierende Gedächtnis »im Namen [i. e. im Sprachzeichen, E. L.] die Sache und mit der Sache den Namen, ohne Anschauung und Bild [hat und erkennt]«. <sup>44</sup> Für Hegel ist der ›Inhalt in der Intelligenz‹ zwar an das Sprachzeichen als materielles Zeichen verwiesen; dennoch gilt das idealistische Anliegen hier noch der Suche nach einem medialen Format, das möglichst wenige sensorische Qualitäten, insbesondere keine ikonischen Bezüge aufweisen soll. <sup>45</sup> Die Stimme, deren – wie Humboldt es formuliert – »scheinbare Unkörperlichkeit dem Geiste auch sinnlich entspricht«<sup>46</sup> scheint diese Forderung deshalb in besonderer Weise zu erfüllen, weil »im ideellen sinnlichen Dasein des Tones [...] die sichtbare körperliche Materie negiert [wird]«. <sup>47</sup>

Dieser idealistischen Bestimmung der phonischen Substanz als ideelle und abstrakte Negativität, die das erkenntnistheoretische Primat der Stimme begründet, lässt sich nun nicht nur mit dem Blickwechsel auf die nicht-diskursive Dimension der Stimme begegnen, wie dies im gegenwärtigen medientheoretischen Diskurs zu beobachten ist. In der Stimme zeigt sich nicht nur die Spur des Körpers, sondern die Stimme hinterlässt ihre mediale Spur ebenso im Körper und das gerade auch in ihrer semiologischen Verfasstheit. Die mediale produktiv-rezeptive Entäußerungsschleife manifestiert sich auch auf der Ebene des vermeintlich ›internen‹ kognitiven Systems im perzeptuo-motorischen Format der kognitiv gespeicherten Zeichenfigur. Wie bereits Saussure hervorgehoben hat, sind sowohl die motorisch-produktive als auch die audio-rezeptive Dimension des artikulierten Zeichens in das kognitive Zeichendispositiv eingeschrieben. Denn für die »langage intérieur« ist alleine – und das sei betont – die synthetische Verbindung von »image acoustique« und »concept« konstitutiv. <sup>48</sup> Sogar hier, auf kognitiver Ebene, haben wir es somit nicht mit medialitätsfreien Signifikaten zu tun; die semantische Dimension bleibt vielmehr untrennbar an die mediale Dimension des »image acoustique« gebunden. Das »image acoustique« wiederum ist keine ausschließlich perzeptuelle Form, sondern bereits eine synthetische Verbindung aus »mécanème«, dem Bewegungsbild, und »acoustème«, dem Klangbild. <sup>49</sup> Eine solche perzeptuo-motorische Verknüpfung von motorisch-produktiver und akustisch-rezeptiver Dimension im kognitiven Zeichendispositiv wird auch durch neuroanatomische Befunde zu motorisch-auditiven Interface-Systemen gestützt, die bei Sprachproduktion und -rezeption gleichermaßen aktiv sind. <sup>50</sup> Es gibt also auch in Bezug auf das kognitive System – so ließe sich die hier vertretene These zusammenfassen – weder amediale ›Types‹ noch semantische Einheiten, die losgelöst von medialen Eigenschaften ihrer Prozessierung wären.

Die Medialitätsgebundenheit des scheinbar Mentalen zeigt sich aber nicht nur auf struktureller, sondern ebenso auf prozessualer Ebene etwa beim soge-

nannten »inneren Sprechen«,<sup>51</sup> also bei sprachlichen Denkvorgängen. Das innere Sprechen weist nicht nur dieselbe metaleptische Struktur auf wie das entäußerte Sprechen, es werden auch dieselben Hirnareale aktiviert.<sup>52</sup> Auch beim inneren Sprechen werden, selbst wenn keinerlei Muskelaktivität nachweisbar ist, dieselben motorischen und sensorischen Areale aktiviert wie beim lauten Sprechen. Die einzig erkennbare Differenz besteht in der fehlenden Aktivierung der Verarbeitungsinstanz, die für die konkrete Ausführung der motorischen Artikulation zuständig ist.<sup>53</sup>

Die epistemische Brisanz des Sich-sprechen-Hörens ist also nicht darin zu suchen, dass sich »das Subjekt hören und aussprechen, sich also vom selbstproduzierten Signifikanten affizieren lassen [kann], ohne den Umweg über die Instanz des Außen, der Welt, des Fremden nehmen zu müssen.«<sup>54</sup> Die Brisanz liegt im Gegenteil genau darin begründet, dass das Subjekt *gezwungen* ist, diesen Umweg zu nehmen und die Spuren dieser Außeninstanz auch nicht aus der Innerlichkeit zu verbannen vermag. Die Stimme gewährt in ihrer Flüchtigkeit nicht die Möglichkeit einer »Tilgung des sinnlichen Körpers und seiner Äußerlichkeit«<sup>55</sup> für ein Innen, sie gewinnt ihre epistemische Bedeutung vielmehr gerade durch die Generierung medialer Spuren. Der Blick auf das Medium Stimme, genauer gesagt auf den medialen Prozess des Sprechens, lässt exemplarisch die Unzulänglichkeit der – in weiten Bereichen etwa der Kognitionswissenschaften aber auch der Medienwissenschaften immer noch unangetasteten – Annahme eines inneren mentalen Systems und einer davon geschiedenen materiellen Außenwelt, die allein auch den Raum des Medialen beherbergt, deutlich werden. Das Innen konstituiert sich nicht nur über das Außen; es bleibt auch in seiner Aufrechterhaltung, in seinen unterschiedlichen Formen der performativen Sinngenerierung auf mediale Entäußerungshandlungen und deren Fremd- und Selbstlektüre verwiesen.

Der Beitrag entstand im Kontext des DFG-finanzierten Teilprojektes »Medialität und Sprachzeichen« des Kulturwissenschaftlichen Forschungskollegs (SFB/FK 427) »Medien und kulturelle Kommunikation« und greift vielfach auf Ergebnisse gemeinsamer Projektdiskussionen mit Ludwig Jäger, Gisela Fehrmann, Luise Springer, Meike Adam und Wiebke Iversen zurück. Ihnen möchte ich ebenso danken wie Cornelia Epping-Jäger für Hinweise, Kritik und Unterstützung.

1 Vgl. Bernhard Waldenfels: Antwortregister, Frankfurt/M. 1994, S. 492.

2 Vgl. dazu näher die Beiträge, insbesondere von Sybille Krämer und Doris Kolesch, in diesem Band.

3 Jacques Derrida: Die Stimme und das Phänomen. Ein Essay über das Problem des Zeichens in der Philosophie Husserls. Übersetzung u. Vorwort v. Jochen Hörisch, Frankfurt/M. 1979, S. 131.

4 Ebd., S. 132.

5 Vgl. ebd., S. 134 f.

6 Das hat in Grundzügen schon Derrida anerkannt, wenn er Hegel als den »Denker der irreduziblen

- Differenz« und den »erste[n] Denker der Schrift« hervorhebt in: Jacques Derrida: *Grammatologie*, Frankfurt/M. 1983, S. 48.
- 7 Wilhelm von Humboldt: *Gesammelte Werke (GS)*. Hg. von A. Leitzmann, Bd. VII, Berlin 1903–36, S. 581.
- 8 Georg Wilhelm Friedrich Hegel: *Werke*, Frankfurt/M. <sup>3</sup>1995, Bd. 10, § 462, Zusatz, S. 280. Vgl. auch ebd., § 401, Zusatz, S. 116: »Der Mensch [...] schafft die *artikulierte* Sprache, durch welche die innerlichen Empfindungen zu *Worte* kommen, in ihrer ganzen Bestimmtheit sich äußern, dem Subjekte gegenständlich und zugleich ihm äußerlich und fremd werden.«
- 9 Humboldt: *GS* (Anm. 7), Bd. VI, S. 152; ebenso Bd. VII, S. 53; Bd. V, S. 174.
- 10 Ebd., Bd. VII, S. 55. Vgl. dazu näher Ludwig Jäger: Über die Individualität von Rede und Verstehen. Aspekte einer hermeneutischen Semiologie bei Wilhelm von Humboldt, in: Manfred Frank (Hg.): *Poetik und Hermeneutik XIII. Individualität*, München 1988, S. 76–94.
- 11 Hegel: *Werke* (Anm. 8), Bd. 10, § 459, S. 271.
- 12 Ebd.
- 13 Humboldt: *GS* (Anm. 7), Bd. V, S. 375.
- 14 Ebd., S. 377.
- 15 Ludwig Jäger: *Zeichen/Spuren. Skizze zum Problem der Sprachzeichenmedialität*, in: Georg Stanitzek/Wilhelm Voßkamp (Hg.): *Schnittstelle. Medien und kulturelle Kommunikation*, Köln 2001, S. 17–31, (hier: S. 20).
- 16 Willem J. M. Levelt/Ardi Roelofs/Antje Meyer: A theory of lexical access in speech production, in: *Behavioral and Brain Sciences* 22 (1999), S. 1–38 (hier: S. 6).
- 17 Vgl. dazu etwa die Überblicksstudien von Peter Indefrey/Willem J. M. Levelt: The Neural Correlates of Language Production, in: Michael S. Gazzaniga (Hg.): *The New Cognitive Neurosciences*, Cambridge, MA/London <sup>2</sup>2002, S. 845–865 (hier: S. 848 und 862 f.); Willem J. M. Levelt/Peter Indefrey: The Speaking Mind/Brain: Where do spoken words come from?, in: Alec Marantz/Yasushi Miyashita/Wayne O'Neil (Hg.): *Image, Language, Brain. Papers from the First Mind Articulation Project Symposium*, Cambridge, MA/London 2000, S. 77–93 (hier: S. 83 und die Grafik S. 89).
- 18 Heinrich von Kleist: *Sämtliche Werke und Briefe in vier Bänden*. Hg. v. Klaus Müller-Salget, Bd. 3: *Erzählungen, Anekdoten, Gedichte, Schriften*, Frankfurt/M. 1990, S. 535.
- 19 Zur transkriptiven Logik des Sprechens vgl. auch Ludwig Jäger: *Transkriptivität. Zur medialen Logik der kulturellen Semantik*, in: ders./Georg Stanitzek: *Transkribieren. Medien/Lektüre*, München 2002, S. 19–41.
- 20 Bettine Menke: *Adressiert in der Abwesenheit. Zur romantischen Poetik und Akustik der Töne*, in: Stefan Andriopoulos/Gabriele Schabacher/Eckhard Schumacher (Hg.): *Die Adresse des Mediums*, Köln 2001, S. 100–120 (hier: S. 115); vgl. auch Paul de Man: *Allegorie des Lesens*, Frankfurt/M. 1988, insb. S. 151.
- 21 Wie Armin Burkhardt unter Bezug auf Susenbrot's Bestimmung der Metalepsis ausführt in seinem Wörterbuchartikel »Metalepsis«, in: *Historisches Wörterbuch der Rhetorik*, hg. von Gerd Ueding, Bd. 5, Tübingen 2001, S. 1087–1099 (hier: S. 1093).
- 22 »Metalepsis est dictio gradatim pergens ad id quod ostendit.« Donatus zitiert nach ebd., S. 1090.
- 23 Vgl. etwa Daniel C. Dennett: *Bewusstsein hat mehr mit Ruhm als mit Fernsehen zu tun*, in: Christa Maar/Ernst Pöppel/Thomas Christaller (Hg.): *Die Technik auf dem Weg zur Seele. Forschungen an der Schnittstelle Gehirn/Computer*, Reinbek 1996, S. 60–89 (hier: S. 84).
- 24 Benjamin Libet: *Neurophysiology of Consciousness. Selected Papers and New Essays*, Boston/ Basel/Berlin 1993, S. 389; vgl. auch Max Velmans: *Understanding Consciousness*, London/Philadelphia 2000, S. 215: »[W]hether we consider conscious forms of input analysis (speech perception and reading), information transformation (verbal thinking) or output (speech production), the conscious experience that we normally associate with such processing *follows* the processing to which it relates.«
- 25 Vgl. dazu näher Libet: *Neurophysiology of Consciousness* (Anm. 24); ders.: *Neural Processes in the Production of Conscious Experience*, in: Max Velmans (Hg.): *The Science of Consciousness. Psychological, Neuropsychological, and Clinical Reviews*, London 1996, S. 96–117. Die zeitliche Verzögerung erklärt sich zu einem großen Teil aus der parzellierten und distributiven Verarbeitungsweise des Gehirns, die zur Generierung kohärenter Einheiten eine zeitliche Synchronisation der verteilten neuronalen Aktivität erfordert, vgl. dazu näher Erika Linz: *Indiskrete Semantik. Kognitive Linguistik und neurowissenschaftliche Theoriebildung*, München 2002, S. 154 ff. Die Angaben über die Dauer der neuronalen Verzögerung variieren zwischen 100 und 1000 msec, wobei sich nach Libet

- überindividuell und stimulusunabhängig ein Richtwert von ca. 500 msec annehmen lässt; vgl. Libet: *Neurophysiology of Consciousness* (Anm. 24), S. 378. Obwohl die Methodenwahl von Libet teilweise auch innerhalb der Neurowissenschaften umstritten ist, findet die generelle Annahme einer Nachträglichkeit bewusster Prozesse und deren subjektiver Vordatierung inzwischen breite Zustimmung und empirische Stützung. Vgl. dazu u. a. die Kommentare zu ders.: *Unconscious cerebral initiative and the role of conscious will in voluntary action*, in: *The Behavioral and Brain Sciences* 8 (1985), S. 529–566 sowie die Fortführung der Debatte in: ebd. 10 (1987), S. 781–786 und ebd. 12 (1989), S. 181–187.
- 26 Peter Halligan/David Oakley: *The greatest myth of all. What do you mean when you talk about ›yourself?‹*, in: *New Scientist*, 11. November 2000, S. 35–39 (hier: S. 36). Vgl. dazu auch Max Velmans: *Understanding Consciousness* (Anm. 24), S. 215: »[T]here is a sense in which one is only conscious of what one wants to say *after one has said it!*« Dasselbe gilt Velmans zufolge auch für bewusstes »inneres Sprechen«; vgl. ebd.
- 27 Jäger spricht hier davon, dass in jede Rede »bereits auf Seiten des Produzenten ein auto-hermeneutisches transkriptives Moment« eingeschrieben ist und sich das Sprechen somit als Form der »Selbsttranskription« verstehen lässt; vgl. Jäger: *Transkriptivität* (Anm. 19), S. 40.
- 28 Ein ähnliches Prinzip hat Auer in Bezug auf die grammatische Satzplanung formuliert; vgl. Peter Auer: *On line-Syntax – Oder was es bedeuten könnte, die Zeitlichkeit der mündlichen Sprache ernst zu nehmen*, in: *Sprache und Literatur* (SuL) 85 (2000), S. 43–56.
- 29 Derrida: *Die Stimme und das Phänomen* (Anm. 3), S. 136.
- 30 Ebd.
- 31 Vgl. dazu Ernst Pöppel: *Zeitlose Zeiten. Das Gehirn als paradoxe Zeitmaschine*, in: Heinrich Meier/Detlef Ploog (Hg.): *Der Mensch und sein Gehirn. Die Folgen der Evolution*, München/Zürich 1997, S. 67–97 sowie Stuart Rosen: *Temporal information in speech. Acoustic, auditory and linguistic aspects*, in: *Philosophical Transactions of the Royal Society London, Series B* 336 (1992), S. 367–373.
- 32 Derrida: *Die Stimme und das Phänomen* (Anm. 3), S. 136.
- 33 Wie wichtig dieses auditive Feedback nicht nur für die ontogenetische Entwicklung, sondern auch für die Aufrechterhaltung der hochkomplexen sprechmotorischen Programme bleibt, belegen besonders eindringlich die Fälle von Spätertaubten, bei denen die sprechmotorische Artikulationsfähigkeit bereits nach kurzer Zeit stark vermindert ist. Insofern greift hier Linkes Argumentation für eine »Rückkoppelungsfreiheit des Sprechens«, die auf einer Einzelfalluntersuchung ohne Langzeitstudie basiert, sicherlich zu kurz und lässt sich in ihrer Generalisierung kaum mit dem neurologischen Forschungsstand vereinbaren; s. Detlef B. Linke: *Die Stimme im Mund*, in: Hans Belting/Dietmar Kamper/Martin Schulz (Hg.): *Quel Corps? Eine Frage der Repräsentation*, München 2002, S. 467–475.
- 34 George H. Mead: *Geist, Identität und Gesellschaft*, Frankfurt/M. <sup>7</sup>1988, S. 104.
- 35 Derrida: *Die Stimme und das Phänomen* (Anm. 3), S. 140.
- 36 Mead: *Geist, Identität und Gesellschaft* (Anm. 34), S. 108.
- 37 Vgl. Humboldt: *GS* (Anm. 7), Bd. V, S. 380 und S. 455.
- 38 Vgl. Mead: *Geist, Identität und Gesellschaft* (Anm. 34), S. 107 ff.
- 39 Hegel: *Werke* (Anm. 8), Bd. 10, § 401, Zusatz, S. 115; vgl. auch ebd., Bd. 14, S. 261.
- 40 Ebd., Bd. 10, § 401, Zusatz, S. 115.
- 41 Ebd., § 462, Zusatz, S. 279.
- 42 Ebd., § 459, S. 271; vgl. zur semiologischen Relevanz der Materialität des Tons insbesondere auch Ludwig Jäger: *Linearität und Zeichensynthese. Saussures Entfaltung des semiologischen Form-Substanz-Problems in der Tradition Hegels und Humboldts*, in: *Fugen. Deutsch-Französisches Jahrbuch für Text-Analytik* 1980, Olten 1981, S. 187–212.
- 43 Hegel: *Werke* (Anm. 8), Bd. 10, § 462, Zusatz, S. 280.
- 44 Ebd., § 462, S. 278.
- 45 Vgl. hierzu insbesondere Jacques Derrida: *Der Schacht und die Pyramide. Einführung in die Hegelsche Semiologie*, in: ders.: *Randgänge der Philosophie*, hg. v. Peter Engelmann, Wien <sup>2</sup>1999, S. 93–132.
- 46 Humboldt: *GS* (Anm. 7), Bd. V, S. 377.
- 47 Jäger: *Linearität und Zeichensynthese* (Anm. 42), S. 193. Vgl. dazu näher Gisela Fehrmann/Erika Linz: *Resistenz und Transparenz der Zeichen. Der verdeckte Mentalismus in der Sprach- und Medientheorie*, in: Jürgen Fohrmann/Erhard Schüttelpelz (Hg.): *Die Kommunikation der Medien*, Tübingen 2003, S. 79–102 (hier insb.: S. 94 ff.).

Erika Linz

64

- 48 Ferdinand de Saussure: *Cours de linguistique générale*, édition critique par Rudolf Engler, Tome 1, Wiesbaden 1968, 149, 1100 f., III C 279; ebd. 43, 172, III C 14, 256; ebd. 43, 265, III C 272; ebd. 43, 261, III C 14; vgl. auch Ludwig Jäger: *Neurosemiologie. Das transdisziplinäre Fundament der saussureschen Sprachidee*, in: *Cahiers Ferdinand de Saussure* 54 (2001), S. 289–337 (hier: S. 318 f. sowie Anm. 149, S. 315).
- 49 Ferdinand de Saussure: *Cours de linguistique générale*, édition critique par Rudolf Engler, Tome 2 : Appendice, *Notes de Ferdinand de Saussure sur la linguistique générale*, Wiesbaden 1974, 32, N 14c [3305.8]; dazu näher auch Jäger: *Neurosemiologie* (Anm. 48), S. 319 ff. sowie Fehrmann/Linz: *Resistenz und Transparenz der Zeichen* (Anm. 47), S. 87 ff.
- 50 Vgl. etwa Gregory Hickock/David Poeppel: *Towards a functional neuroanatomy of speech perception*, in: *Trends in Cognitive Sciences* 4 (2000), S. 131–138.
- 51 Vgl. insb. Lew Semjonowitsch Wygotski: *Denken und Sprechen. Mit einer Einleitung v. Thomas Luckmann*. Hrsg. v. Johannes Helm, Stuttgart 1969.
- 52 Vgl. dazu Velmans: *Understanding Consciousness* (Anm. 24), S. 215.
- 53 Gregorsz Dogil/Herrmann Ackermann/Wolfgang Grodd et al.: *The speaking brain. A tutorial introduction to fMRI experiments in the production of speech, prosody and syntax*, in: *Journal of Neuro-linguistics* 15 (2002), S. 59–90.
- 54 Derrida: *Die Stimme und das Phänomen* (Anm. 3), S. 135.
- 55 Ebd., S. 134.



**Sybille Krämer**  
**NEGATIVE SEMIOLOGIE DER STIMME**

»... il n'y a pas de voix neutre« –  
 »... es gibt keine neutrale Stimme«<sup>1</sup>  
 Roland Barthes

**1. ZUR MARGINALISIERUNG DER STIMMLICHKEIT**

Ist das nicht eine Merkwürdigkeit: Zu sprechen gilt den Philosophen als *das* grundlegende symbolische Vermögen der Menschen. Schon Platon und Aristoteles ließen keinen Zweifel daran, dass die *mündliche* Sprache gegenüber allen Formen der verschrifteten Sprache primär sei. Mit Wilhelm von Humboldt wurde dann die ›dialogische Natur‹ des Sprechens fast zum philosophischen Gemeingut; und das 20. Jahrhundert mit seiner sprachkritischen Wende rückte nicht nur die Umgangssprache – statt die am Vorbild formaler, also geschriebener Sprachen gewonnene Idealsprache – ins Zentrum philosophischer Aufmerksamkeit, sondern entdeckte und explizierte überdies die Handlungsdimension des alltäglichen Sprechens in Gestalt der Sprechakttheorie.

Kaum ein anderes Sujet ist überdies so zum Kreuzungspunkt philosophischer Schulbildungen geworden wie das Themenfeld der Sprache. Nur: All dies ändert nichts an dem schlichten Tatbestand, dass die Sprache, welche die Philosophen interessiert und die sie in ihren Reflexionen Kontur gewinnen lassen, eine weitgehend *lautlose* Sprache ist: Denn die für das Sprechen konstitutive Stimmlichkeit ist – allerdings gibt es Ausnahmen<sup>2</sup> – kein Gegenstand philosophischer Beachtung.

Stellen wir uns vor, dass Worte, die wir sprechen, nicht vergehen, sondern stehen (bleiben); so dass also mit jedem ausgesprochenen Satz die Welt von ein paar Entitäten mehr bevölkert würde. Schon diese Vorstellung zeigt: dass Sprache überhaupt so fundamental wie allgegenwärtig sein kann für die Belange des Menschlichen, ist mit ihrer fluiden Materialität durchaus verschwistert. Der Laut ist ein ›Stoff‹, der die bemerkenswerte Eigenschaft hat, sich im Augenblick seines Entstehens zu verflüchtigen.

Warum aber bleiben die dem Sprechen eigenen lautlich-auditiven Merkmale so randständig und bleibt die Stimme als Organon und Medium des Sprechens so unthematisiert in der Sprachphilosophie?

Mit ihrer Neutralisierung der Lautlichkeit der Sprache steht die Philosophie allerdings nicht alleine; in der Sprachwissenschaft zeigt sich ein verwandtes Phä-

nomen: Zwar ist die Einsicht, dass die Sprachwissenschaft sich zur Sprechwissenschaft zu gestalten habe, derweilen in der Linguistik nicht nur anerkannt, sondern ein Stück weit auch praktiziert.<sup>3</sup> Gleichwohl ist es signifikant, dass der Gründungsvater der modernen Sprachwissenschaft, Ferdinand de Saussure, und mit ihm die phonologische Linguistik, den Sprachlaut als ein *hörbares* Ereignis aus dem Gegenstandsbereich der Linguistik ausgeschlossen haben.<sup>4</sup> In der strukturalistischen Perspektive kommt das Phonem nur ›in den Blick‹, insofern es spezifizierbar ist in einem System negativer differentieller Oppositionen, d. h. aber durch seine Relation zu dem, was mit dem Auftreten des je artikulierten Lautes gerade ausgeschlossen, was also im ausgesprochenen Worte abwesend ist. Das Phonem ist eine rein theoretische, somit geräuschlose Entität: Das, was es ist, verdankt es seinen formalen Charakteristika.

Die Entkörperung der Sprache als Gegenstand der Linguistik radikalisierte dann Noam Chomsky und weist eben dadurch der kognitiven Linguistik den Weg: In seiner Unterscheidung von ›interner‹ und ›externer Sprache‹ ist die wirklich geäußerte Sprache, die artikulierte Rede, nur ein Epiphänomen jener internen, mental und hirnhysiologisch instantiierten – also unhörbaren – Sprache, die das eigentliche Objekt der Linguistik repräsentiert.<sup>5</sup>

Doch zurück zur Philosophie: Nehmen wir einmal an, die Diagnose einer ›stimmlosen Sprache‹ als Bezugspunkt philosophischer Reflexion trifft zu. Aber warum soll(t)en wir diesen Umstand als eine *Vernachlässigung* deuten? Können wir nicht einräumen, dass es gute Gründe geben kann, wenn Philosophen die Sprachlichkeit gerade unabhängig von der Stimmlichkeit begreifen und erörtern? Um diesen Gründen auf die Spur zu kommen, wollen wir uns einige Eigenarten des Stimmlichen vergegenwärtigen.

## 2. EINE KLEINE PHÄNOMENOLOGIE DER STIMMLICHKEIT

In einer phänomenalen Perspektive treten zwei Dimensionen hervor, in denen Eigenarten der Stimmlichkeit Gestalt gewinnen: Einmal geht es um eine Perspektive, die mit den Begriffen Materialität, Leiblichkeit, Prozessualität und Ereignishaftigkeit des Lautlichen charakterisierbar ist; zum andern um eine Dimension, in welcher die Begriffe der Intersubjektivität, des Appellcharakters, der Affektivität und der Autorität des Stimmlichen eine Rolle spielen.

### 3. MATERIALITÄT, LEIBLICHKEIT, PROZESSUALITÄT UND EREIGNISHAFTIGKEIT:

#### ÜBER DIE FLÜCHTIGE MATERIE DER STIMME

Gehen wir aus von der Eskalation der Stimme im Schrei. Die erste menschliche Lebensäußerung ist der Schrei. Symptom dafür, dass wir atmen, also leben. »Leblose Dinge haben keine Stimme«, stellt Aristoteles fest.<sup>6</sup> In der phonischen Aktivität wird unser Körper zum ›Stimmapparat‹,<sup>7</sup> zum Instrument: Ein durch Lungen und Bronchien geformter Luftstrom versetzt unsere Stimmbänder in Schwingungen und erzeugt durch deren Vibration Töne. Die Stimme beruht auf einer elementaren Motorik des Körpers: ihr liegt ein Wechselspiel zugrunde zwischen der ›Immaterialität‹ des Atemhauchs, der Resonanzfähigkeit höhlenartiger Organe und der beweglichen Widerständigkeit der Stimmbänder. Welche Stelle auch eine Stimme auf der Skala zwischen Geräusch und artikuliertem Laut einnimmt: Ihre Klangfarbe, ihre Tonlage, ihre Lautstärke zeugen von der Verfassung des Körpers, der zum Organon der Verlautbarung wird.<sup>8</sup> Die Stimme ist weniger Gegenstand und Zustand, sondern sie ist Bewegung, ist Prozessualität. *Der ›Ort‹ der Stimme ist die Aktivität des sie erzeugenden Leibes.* Die Stimme ist die Spur des Körpers im Sprechen. Das aber heißt: Der Körper *zeigt sich* in der Stimme.

Während also die Stimme da ist, wo die Motorik des Leibes sich vollzieht, haben die durch sie erzeugten Laute keinen Ort. Waldenfels vermerkt: »Töne und Geräusche werden nie seßhaft.«<sup>9</sup> Denn Laute – wie alles Akustische – gelten gewöhnlich nicht als Entitäten für sich, sondern werden identifiziert mit dem, *was* sie erzeugt.<sup>10</sup> Sobald wir ein akustisches Phänomen einem Ding oder einem Lebewesen zurechnen und das Geräusch somit als ein ›Geräusch von etwas‹ identifizieren, hören wir keineswegs mehr ›Geräusche‹ oder ›Laute‹, sondern dasjenige, was ertönt: Wir hören ›die Uhr ticken‹, doch gewöhnlich ist es wenig sinnvoll zu sagen: wir hören ›das Ticken der Uhr‹. Allerdings ist bei der Identifikation des Akustischen mit seiner Quelle zu beachten, dass Geräusche dann weniger als Eigenschaften der sie verursachenden Dinge gelten, denn als deren aktivitätsbetonte *Äußerung*: Etwas hat eine Farbe, aber es *macht* ein Geräusch.<sup>11</sup>

Diese Ortlosigkeit und Prozessualität des Lautlich-Akustischen – und nur darum ist es uns hier zu tun – verweisen auf seinen Ereignischarakter. ›Ereignis‹ hier in einem durchaus unkomplizierten Sinne verstanden, als das, was singulär, einmalig, unwiederholbar ist. Kaum entstanden, ist der Laut auch schon verklungen. Eine Materie, die ganz und gar Bewegung ist, deren ›Sein‹ in ihrem ›Verschwunden-Sein‹ besteht, deren Anwesenheit nur Passage zur Abwesenheit ist. Der Laut ist eher Fluxus und Fluidum denn Stoff. Die verlautende Stimme verkörpert eine Ereignishaftigkeit par excellence.<sup>12</sup>

#### 4. APPELL, AFFEKTIVITÄT, AUTORITÄT: DAS STIMMLICHE ALS ›URSTIFTUNG‹ VON INTERSUBJEKTIVITÄT.

Kommen wir zurück zum Phänomen des ersten Schreis. Gibt es etwas Auffordernderes, Eindringlicheres als der Schrei des Säuglings? Ungerichtet, ohne Bedeutung im semantischen Sinne, jenseits aller Konvention, ganz und gar Anzeichen, unwillkürliche Spur eines Körpers, dessen Muskeln sich in seiner Bedürftigkeit anspannen, ist dieser Laut doch überaus signifikant, als Appell unfehlbar bedeutsam: Ein Schrei rüttelt den auf, der ihn hört. Vor dem »Notstand der Rede«,<sup>13</sup> der den Übergang vom Sprechen zum Schreien provoziert, steht die elementare Notsituation des Körpers, die sich im Schrei Bahn bricht. Die Intersubjektivität findet in der Stimmllichkeit ihre basale, vielleicht sogar ihre nachhaltigste, aber auch ihre machthaltigste Form.

Was im Schrei des Säuglings kulminiert: dass diese Stimme Fürsorge evoziert, oder – so sie nicht zum Schweigen kommt – uns nervlich zerrüttet, das gilt für jedwede menschliche Verlautbarung: Die Stimme kann anziehen oder sie kann abstoßen, kann binden oder entzweien; doch stets eignet ihr eine Kraft, die intersubjektiv wirksam wird. Stimmen berühren. »Das Wahrnehmen von Stimmen (ist) immer ein affektiver Prozess.«<sup>14</sup> Auch beim Hören, wie schon bei der Stimmproduktion, arbeiten Physiologie, Physik und Phänomenalität zusammen: Wie alles Akustische haben Stimmen etwas Eindringliches. Gleich dem Auge ist das Ohr zwar ein Fernsinn. Doch anders als das Auge ist das Ohr nicht verschließbar: Töne, nicht aber Visuelles, können uns aus dem Tiefschlaf wecken.<sup>15</sup> Dem Akustischen eignet eine Distanzlosigkeit: Was wir hören, ist ›innen‹ und ›außen‹ zugleich. Ist es diese ›Aufdringlichkeit‹, welche die Stimme prädestiniert, auch zur Elementarerfahrung von Macht und Ohnmacht zu werden? Aggressivität und Defensivität, Imponiergehabe und Selbstverkleinerung zeigen sich – das dokumentieren biophonetische Untersuchungen<sup>16</sup> – im Mischungsverhältnis der uns zu Gebote stehenden Register des ›Brusttons‹ und des ›Kopftons‹. Und was bedeutet die etymologisch verbürgte Nähe zwischen ›Gehör‹ und ›Gehorsam‹, ›Horchen‹ und ›Gehorchen‹? Warum verlieren Hitlers Reden, transkribiert zum Text, ihr Faszinosum und ihre Verführungskraft?

Jedenfalls: Die Lautgebung ist die basale Form appellativer und affektiver Bezugnahme auf den anderen. Sie ist praktizierte Intersubjektivität, und zwar vor aller symbolischen, legitimatorischen Bezugnahme auf den anderen in Gestalt von Geltungsansprüchen, die wir mit der Rede erheben oder zurückweisen können.

Doch selbst da, wo die Stimme weniger appellative, denn informative Funktionen erfüllt, eignet ihr eine Kraft der Überzeugung, die keineswegs – oder bes-

ser: noch gar nicht – zehrt von der Kraft des Arguments. Gebunden an die leibliche Präsenz des Sprechers – darauf haben Vernant und im Anschluss an ihn Mersch<sup>17</sup> verwiesen – verbürgt die Anwesenheit des Redners einen Wahrheitsbezug der Rede. Eine in der körperlichen Anwesenheit wurzelnde Zeugenschaft<sup>18</sup> desjenigen, der spricht, für das, was er spricht. Beruht die Performanz der mündlichen Rede, ihre besondere Autorität, gerade darin, dass das gesprochene Wort sich als Aussage noch gar nicht abtrennen lässt von der Person? So dass in der mündlichen Rede de facto Wahrhaftigkeit und Wahrheit zusammenfallen?

Wir nannten es eine ›Merkwürdigkeit‹, dass die Philosophie so nachdrücklich die Sprache, nicht aber die Stimme zum Objekt ihres Nachdenkens macht. Vielleicht sind wir den Gründen für diese ›Neutralisierung der Stimmlichkeit‹ nun ein Stück näher gekommen.

##### 5. DAS INTELLEKTUALISTISCHE SPRACHKONZEPT: WO ABER LIEGT DIE ALTERNATIVE?

Worauf achten Philosophen, wenn sie sich mit der Sprache auseinandersetzen?<sup>19</sup> ›Logos‹ meint ursprünglich ›Sprache‹. Doch seit ›Sprache‹ und ›Vernunft‹ – jedenfalls dem Worte nach – unterscheidbar werden, gibt es in der Philosophie ein Bestreben, deren Einheit wieder herzustellen. Und dies in zweierlei Versionen: Einmal ist das die Idee, dass die Sprache ein privilegiertes Organ der Repräsentation und Übermittlung von Gedanken, Wissen und Erkenntnis ist. Sei das nun in Gestalt der Auszeichnung der apophantischen Aussage vor allen anderen Redeformen oder als Konstruktion von Idealsprachen zwecks Repräsentation einer ihnen als vorgängig angenommenen logifizierbaren Ordnung der Welt. In der zweiten Version geht es um die Idee, dass die Sprache das Vernünftige nicht nur repräsentiert, sondern hervorbringt. Dass also im und durch Sprachgebrauch das Vernünftige sich nicht nur darstellt, vielmehr Vernunft kraft der Selbstreflexivität der Sprache im Sprechen entsteht.

Das theoretische Umfeld dieser beiden Versionen von der ›Logosauszeichnung der Sprache‹<sup>20</sup> birgt eine Fülle weiterer Ideen:<sup>21</sup> So die Annahme von der Existenz syntaktischer, semantischer und pragmatischer Universalien; von der Fundierung des Sprechkönnens in einem Sprachwissen, dessen Kern die Kenntnis von grammatischen und pragmatischen Regelsystemen ausmachen; von der Idealisierung von Sprache und Sprechen als einem notwendigen Verfahren um sprachliche Universalien zu rekonstruieren usw. usw. Der Preis für ein logosorientiertes Sprachkonzept ist allerdings die in ihm angelegte Virtualisierung

von Sprache und Kommunikation; es geht um die *Möglichkeit* von Sprache, um eine Kommunikation ›sub specie aeternitatis‹, nicht aber um die *Wirklichkeit* unseres alltäglichen Sprechens und Kommunizierens. Doch dies bedeutet keineswegs, dass das logosorientierte Sprachkonzept eine bloße Fiktion ist; denn es hat ein ›fundamentum in re‹. Diese nicht-fiktionale Verankerung im Realen liegt darin, auf einer Extrapolation und Stilisierung von Sprachspielen zu beruhen, die – bei näherem Hinsehen – *Schriftgebräuche* sind, sich also stützen auf Modalitäten des Umganges mit *Texten*. Das intellektualistische Sprachkonzept ist – so jedenfalls die Vermutung – einem impliziten Skriptizismus verpflichtet; ihm steht die Schrift – auf verschwiegene Weise – Modell. Zumindest ex negativo hat unsere ›kleine Phänomenologie der Stimmlichkeit‹ auch zeigen können, warum das so ist.

Rekapitulieren wir noch einmal die Besonderheiten der Stimmlichkeit:

- Im Fluxus des Sprechens verkörpert die Stimme Ereignishaftigkeit; die Aisthesis des gesprochenen Wortes ist von irreduzibler Singularität.
- Als vollzogener Machtgestus oder als Reflex gefühlter Ohnmacht, als eindringlicher und aufdringlicher Anspruch an den anderen, ist das Register der Stimme verwoben mit einem Typus der Zwischenmenschlichkeit, dessen Nährboden weniger das Argumentieren, denn der Affekt ist.
- Als Teil der elementaren wie existentialen Leiblichkeit der Sprecherzeugt die Stimme immer auch von unserer Bedürftigkeit, der ein Begehren eigen ist, das sich an den anderen richtet. In unserer Stimmlichkeit werden wir nicht nur als ›Essenzen‹, vielmehr als konkret situierte leibliche ›Existenzen‹ offenbar.

Wenn aber in der Stimme sich eine macht- und ohnmachtbezogene Form von Intersubjektivität, ein gefühlsmäßiges Gestimmtsein jenseits kognitiver Dispositionen, eine ästhetische Kraft, die nicht dem Logos zu Diensten ist, artikuliert, dann bricht sich in der Lautlichkeit Bahn, was der ›Logosauszeichnung‹ von Sprache und Kommunikation gerade nicht subsumierbar, durch sie nicht instrumentalisierbar ist. So ist es kaum verwunderlich, dass ein Sprach- und Kommunikationskonzept, für welches das Sprechen nicht alleine durch Regelbeschreibung rationalisierbar ist, sondern überdies noch zur Springquelle von Rationalität und rationalem Verhalten avanciert, dass ein solches Konzept ohne die Reflexion der Stimmlichkeit bestens auskommt. Wie umgekehrt: Die Stimme einzubeziehen bedeutet dann, sich an einem nicht-intellektualistischen Sprachkonzept zu orien-

tieren. Wir haben auch schon einen Wink, wie dabei methodologisch zu verfahren ist: Eine Alternative zum kognitivistischen Sprachbild zu entwerfen, heißt zuerst einmal, eine ›performative Revision‹ seiner methodologischen Prämissen einzuleiten.

Das ›Bauprinzip‹ intellektualistischer Sprachtheorien ist die Unterscheidung zwischen Schema (System, Regelwerk) und Gebrauch (Aktualisierung, Realisierung).<sup>22</sup> Gemäß dieser methodologischen Prämisse gehört das Lautliche nicht zur Sprache, sondern ist ›nur‹ das Medium zum Vollzug von Sprache, die selbst dabei als medienindifferent konzipiert ist. Die theoretische Gelenkstelle einer ›performativen Revision‹ ist es nun, dass in ihrem Rahmen Medien eben nicht mehr marginal, vielmehr konstitutiv sind. ›Konstitutiv‹, insofern im aktualisierenden Vollzug eines Schemas dieses Schema immer auch transformiert, untermindert oder überstiegen wird. Die Richtung, die der kreative Überschuss des Vollzuges gegenüber dem darin realisierten Muster jeweils annehmen kann, ist aber als Potential im Medium selbst angelegt.

Aber sind die ›Weichen‹ eines solchen Alternativprogramms nicht schon längst gestellt? Spätestens mit der im poststrukturalistischen Diskurs – einsetzend mit Lacan – üblich gewordenen Aufwertung des Signifikanten gegenüber dem Signifikat ist doch eine Perspektive gewonnen, in der auch die Stimme als genuiner Bestandteil sprachlichen Geschehens rehabilitierbar ist. Das ist dann der Fall, wenn wir die Stimme als materiellen Signifikanten und den Aussagegehalt der Rede als deren Signifikat deuten. Wir brauchen doch nur ›was ein Medium ist‹ zu identifizieren mit dem materiellen Zeichenträger – und schon ist ein bequemer Zugang gewonnen, um nahezu alle geistes- und kulturwissenschaftlichen Objekte, die doch im weitesten Sinne als ›Zeichendinge‹ bzw. ›symbolische Formen‹ qualifizierbar sind, in eine medientheoretische Perspektive zu rücken. Die Stimme wäre damit nicht länger als ein Außersprachliches marginalisiert, vielmehr als unverzichtbarer materieller Signifikant sprachlicher Semiosis rehabilitiert. Und doch: Dieser Weg führt noch nicht weit genug.

Der Grund dafür ist, dass, was immer ein Medium ist, nicht aufgeht in dem, wozu ein ›Signifikant‹ im Rahmen der semiotischen Beziehung zwischen Zeichenträger und Zeichenbedeutung dient. Oder, falls man doch das Medium irgendwie auf Seiten des Signifikanten lokalisieren möchte: Das Medium ist dann gerade jene Dimension am Signifikanten, »die nicht zur Signifikation beiträgt.«<sup>23</sup> Das Medium durchbricht also das Modell der Semiosis.<sup>24</sup> Medientheoretische Reflexionen erweisen sich damit als eine Möglichkeit, die Grenzen des semiotischen Paradigmas für Untersuchung und Reflexion kultureller Gegenstände auszuloten.

Wie dieses »Aufbrechen der Signifikanz«<sup>25</sup> zu verstehen ist, wollen wir nun

exemplarisch am Beispiel der Stimme zu klären versuchen. Zuerst gehen wir dabei einen indirekten Weg: Wir fragen nicht nach der Stimme als *Medium*, vielmehr nach der Zeichenfunktion der Stimme.

#### 6. DIE STIMME ›NUR‹ ALS ZEICHEN BETRACHTET

Wir verstehen unter einem Zeichen – ganz unambitioniert, aber dies ist keine zeichentheoretische Abhandlung! – eine sinnlich sichtbare Struktur, die für Interpretieren etwas ›Sinnhaftes‹ bzw. ›Bedeutungsvolles‹ vergegenwärtigt. Wenn eine Lautfolge als Sprachäußerung wahrgenommen wird, so deuten wir ein akustisches Phänomen als semantisches Phänomen: Wir können dann sagen, dass die Lautfolge – je nach sprachtheoretischer Position – einen Sinn, eine Satzbedeutung, eine Proposition oder einfach einen Gedanken zum Ausdruck bringt. Aber was ermöglicht diese Metamorphose eines physischen Ereignisses in ein semiotisches Geschehen? Von Worten sagen wir, dass sie arbiträre Bedeutungen haben. ›Arbiträr‹ auch hier wieder im unambitionierten Sinne einer durch Konventionen geregelten und/oder durch alltägliche Praktiken eingeschliffenen, also nicht in Ähnlichkeiten fundierten ›Zuordnung‹ von Wort und Sinn. Doch wenn unsere Stimme artikuliert, so ist das Verhältnis ihrer lautlich-tonalen Merkmale zu dem, was ausgedrückt wird, nicht einfach arbiträr. Laut und Sinn gehen in der mündlichen Äußerung eine intimere Verbindung ein: Empirische Untersuchungen haben gezeigt, dass beim Reden der Atemrhythmus, die ›Inspiration‹, Sinnschritte im Gesagten markiert.<sup>26</sup> Die Pausen beim Sprechen folgen nicht syntaktischen Zäsuren, sondern die Einatmung setzt und markiert *semantische* Segmentierungen.<sup>27</sup> Wie die Atemführung sich vollzieht, legt nah, *was* überhaupt als Sinneinheit zu gelten hat. Die Stimme *zeigt*, was sie sagt. D. h. aber auch: die Semiosis des Lautlichen ist fundiert im Analogieprinzip. Die Stimmlichkeit durchbricht die reine Diskursivität des Sprachlichen.

Was nun für die semantisierende Funktion des Lautlichen gilt, das gilt für die expressive Rolle der Stimme erst recht. Stimme und Lautgebung zeigen und zeugen von der Person, die spricht: ›Per-sonare‹, meint das Ertönen der Stimme durch das Mundstück der Maske beim antiken griechischen Schauspieler. ›Persona‹: das ist die Rolle, die sich im Sprechen artikuliert. Dass die Stimme die ›persona‹ verkörpert, gilt für die Merkmale unserer Geschlechtlichkeit ebenso, wie für alle übrigen körperlichen und seelischen Dispositionen: Gesteuert durch das vegetative Nervensystem prägt der leibliche Tonus den Ton.<sup>28</sup> Die Stimme – soviel jedenfalls zeichnet sich ab – zeigt die Gestimmtheit.



Fassen wir diese Überlegungen zusammen: Die semiotische Wirkung der Stimme beruht weniger auf kodierter Zeichengebung und mehr auf unwillkürlicher Indexikalität. Die Stimme ist nicht einfach Symbol, sondern Spur von etwas; sie fungiert nicht einfach als Zeichen, vielmehr als Anzeichen. In dieser ihrer Indexikalität ist es begründet, dass die Stimme nicht nur spricht, sondern zeigt.

Gemäß einer traditionellen Schematisierung unserer symbolischen Vermögen stehen uns zwei grundlegende Register der Zeichengebung zu Gebote: das ist das Diskursive und das Ikonische, das Sagen und das Zeigen, auch explizierbar als das Digitale und das Analoge. Wir sind gewohnt, die Sprache mit dem Diskursiven, dem Sagbaren, das Bild aber mit dem Ikonischen, dem Zeigbaren zu identifizieren. Wenn aber die Behauptung, dass mit der Stimme ein analogisch-indexikalisches Prinzip beim Sprechen wirksam wird, zutrifft, ergibt sich ein bemerkenswerter Umstand: Jenes physisch-psychische Substrat des zum Laut geformten Schalls, das wie kein anderes in seiner zeitlichen Sukzessivität und Fluidität geeignet ist, sprachliche Materialität zu stiften, also der Sprache ›einen Körper zu geben‹, wird als ein nicht-diskursives Anzeichen zur Bedingung der Möglichkeit zeichenhafter Diskursivität. Die Signifikanz der Lautsprache ist verwoben mit dem Signalcharakter des Lautlichen.

Gleichwohl: In einer semiotischen Perspektive ist das, was die Stimme leistet, im weitesten Sinne als ein Beitrag zur ›Zeichengebung‹ zu begreifen; sei das nun, indem sie zur Akzentuierung des Sinns einer Äußerung beiträgt, sei dies, indem sie zum individuierenden Ausdruck der Person wird. Und doch ist dies erst die ›halbe Wahrheit‹. Denn die Wirkkraft der Stimmlichkeit geht in ihrer Rolle für die sprachliche Semiosis keineswegs auf.

## 7. ›NEGATIVE SEMIOLOGIE‹ DER STIMME ODER: DAS MUSIKALISCHE IM UND AM SPRECHEN

Was es heißt, dass die Stimme nicht nur beiträgt zur Signifikanz der Rede, sondern diese auch durchbricht, erschließt sich erst einer Einstellung, die wir hier als ›Negative Semiologie‹ kennzeichnen können. Die Maxime der Negativen Semiologie ist: Was die Stimme als *Medium* von Sprache und Kommunikation bewirkt, ist in *zeichentheoretischen* Termini hinreichend nicht mehr beschreibbar. Die Medienperspektive einzunehmen ist also ein Versuch, die in der Semiosis nicht aufgehenden Dimensionen der Lautlichkeit zutage treten zu lassen. Daher läuft die Medienperspektive – in letzter Konsequenz – darauf hinaus, die uns selbstverständliche Idee, sprachliches Tun mit einem Zeichenhandeln zu identifizieren, zu relativieren.

Gehen wir noch einmal aus von der Materialität der mündlichen Sprache: Es gibt eine mit und seit Aristoteles ›definitiv‹ gewordene begriffliche Trias zur Kennzeichnung des Akustischen: *psophos* (lat. *sonus*) bedeutet Schall oder Geräusch; *phoné* (lat. *vox*) meint den sprachlichen Laut; *phthongos* (lat. *sonus musicus*) bezieht sich auf den musikalischen Ton.<sup>29</sup> Die Unterscheidung zwischen sprachlichem *Laut* und musikalischem *Ton* ist also ein lang tradiertes Kulturgut. Aber ist in dieser Tradition nicht auch etwas verlorengegangen? Die Kunstpraxis der altgriechischen *musiké* vollzog und verstand sich als Einheit von Musik, Sprache und Tanz, kristallisiert im Bindeglied des Rhythmus als Ordnung einer Bewegung in der Zeit.<sup>30</sup> Kann nun die konzeptuelle Aufspaltung der *musiké* in der Unterscheidung von Sprache und Musik auch als ein Reflex auf die Literalisierung der mündlichen Sprache durch das Alphabet gedeutet werden? Denn die phonetische Schrift mit ihrem erstmals durch die griechische Erfindung von Buchstaben für Vokale realisierten Anspruch, die mündliche Sprache vollständig in nicht weiter zerlegbare ›bedeutungslose‹ Elemente aufzuspalten, mithin die Sprache als eine Art von System aufzufassen, löst die lautsprachliche Schicht heraus aus einer kommunikativen Konstellation, in der Gestik, Mimik, Prosodie, Verbalität und Situationsbezug der Rede in holistischer Weise zusammen wirken. Auskristallisiert in einem allein zu den Augen sprechenden Schriftbild, kann die Sprache überhaupt erst zu einem Objekt von Beobachtung, Untersuchung und Reflexion und damit auch zu einem isolierbaren, solitären Medium der Kommunikation werden.

Wenn es aber so ist, dass die phonetische Schrift zur Modellbildnerin ›der‹ Sprache wird, – dann werden sich gerade im Begriff des ›Lautes‹ bzw. des ›Phonems‹ von Anfang an Merkmale eingeschrieben haben, die nicht diejenigen der akustischen Stimme, vielmehr der visuellen Schrift sind. Die abendländische Konzeption vom Sprachlaut – das jedenfalls ist die Vermutung – ist geprägt von einem impliziten Skriptizismus. Wir können diesen Gedanken hier nicht weiter verfolgen.

Uns kommt es nur auf eine Facette an, die darin besteht, dass mit der Skripturalisierung des Sprachlautes seine *Entmusikalisierung* eingeleitet ist. Mit der Dazwischenkunft der phonetischen Schrift wird die Sprache ihrer musikalischen Dimension entkleidet. In dieser Perspektive zeugt die Marginalisierung der Stimme daher auch von einer Eskamotierung – oder sollten wir sagen: Verdrängung? – des Musikalischen aus dem mündlichen Sprachgebrauch. Und umgekehrt: Eine Rehabilitierung der Stimmlichkeit heißt dann, die musikalische Dimension am und im Sprechen wiederzugewinnen. Das also, worin die ›negative Semiotik der Stimme‹ die Semiotik der Sprache aufbricht, liegt in der Musikalität des Sprechens.

Aber hier drängt sich ein Einwand auf: Ist nicht das Musikalische selbst ein symbolisches Vermögen? Verkörpert die Musik nicht gerade eine dem Sprachlichen verwandte Semiosis? Doch wenn Musik wie eine Sprache verstanden wird, verführt diese Semiotisierung des Musikalischen allzu leicht dazu, dem musikalischen Geschehen eine außermusikalische Bedeutung zuzuweisen. Doch gerade die kulturellen Praktiken der europäischen Instrumentalmusik, die wir signifikanterweise als ›absolute Musik‹ kennzeichnen, beruhen darauf, dass Musik nichts außerhalb ihrer selbst denotiert, dass ihre ›Bedeutung‹ allein im internen Beziehungsgefüge der Töne aufeinander liegt. Wenn also Musik in Analogie zur Sprache zu betrachten ist, dann ist das fruchtbar nur, wenn damit zugleich eine Umorientierung in der Sprachauffassung selbst angezielt wird. Musik ist dann ›Sprache minus Bedeutung‹, bzw. eine ›Signifikantenbewegung ohne Signifikat‹. Aber ist das dann noch ›Sprache‹ und ist das dann noch ein ›Signifikant‹?

Eine Umwendung der Blickrichtung gewinnt nun Kontur: Kann, statt dass die Musik wie eine Sprache gedeutet wird, die Sprache in Analogie zur Musik betrachtet werden? Rudolf Fietz hat in einer wegweisenden Studie gezeigt, dass es Nietzsche gewesen ist, der diese Umwendung des Blicks vollzogen hat und so zu einem bemerkenswerten Sprachbild gelangt, das verankert ist im ›Tonuntergrund‹ der Sprache und damit in der für das Sprechen bedeutsamen Stimmlichkeit.<sup>31</sup> Fietz allerdings ist auch der Auffassung, dass bei der Betrachtung der Sprache nach dem Vorbild der Musik semiotische Eigenschaften von der Musik auf die Sprache zu übertragen sind. Uns dagegen interessiert, dass die Musikalität des Sprechens gerade in der Gestalt, die Nietzsche ihm gab, das semiotische Paradigma nicht einfach von der Sprache auf die Musik verschiebt und erweitert, vielmehr *sprengt*.

## 8. NIETZSCHE UND DER ›TONUNTERGRUND‹ DER SPRACHE

»Das Verständlichste an der Sprache ist nicht das Wort selber, sondern Ton, Stärke, Modulation, Tempo, mit denen eine Reihe von Worten gesprochen wird, kurz die Musik hinter den Worten.«<sup>32</sup> Wie haben wir diese Äußerung von Nietzsche zu verstehen?

Nietzsche hat keinen Einzeltext zur Sprache in musikalischer Perspektive verfasst. Doch die musikästhetischen Ideen des jungen Nietzsche haben wegweisende sprachphilosophische Spuren hinterlassen.<sup>33</sup> Eine dieser Spuren ist Nietzsches Überblendung der kunsttheoretischen Kategorien des ›Apollinischen‹ und des ›Dionysischen‹ mit den sprachtheoretischen Begriffen ›Geberdensprache‹ und ›Tonsprache‹.

Für Nietzsche erfüllt die Sprache zwar auch begriffliche Funktionen, doch es bleibt im Sprachgebrauch immer ein im Begriffssprachlichen »unauflösbarer Rest«. <sup>34</sup> Dieser Rest vollzieht sich nicht bewusstseinsgesteuert, sondern instinktiv, doch gleichwohl »zweckmäßig wirkend«. <sup>35</sup> Und er realisiert sich in zwei Modalitäten: als Gebärdensprache und als Tonsprache. Allerdings bezieht sich Nietzsches Differenz zwischen Gebärde und Ton keineswegs auf diejenige zwischen Gestik und Stimme; vielmehr geht es um eine Doppelung, die im *gesprochenen* Wort selbst wirksam wird. In der mündlichen Sprache sieht Nietzsche die Zweifaltigkeit von Bild und Musik, »uranfänglich vorgebildet«. <sup>36</sup> Die Gebärde ist mit einer »begleitenden Vorstellung« verknüpft, ist damit dem Bildlichen zugehörig. Nietzsche nennt sie auch: »Mundgeberde(n)«. <sup>37</sup> Und so rechnet er die Konsonanten und Vokale zu dieser Gebärdensymbolik: <sup>38</sup> ohne Ton gedacht, sind diese bloß Stellungen der Sprechorgane. Allerdings erzeugen sie in dieser Eigenschaft verständliche Symbole. Der Ton dagegen ist nicht mit Vorstellungen, sondern mit »Strebungen des Willens« <sup>39</sup> verknüpft, die sich als Lust und Unlust äußern. <sup>40</sup> Aus dem Zusammenspiel von Mundgebärde und Tonalität geht dann die Lautsprache hervor. »Die innigste und häufigste Verschmelzung von einer Art Geberdensymbolik und dem Ton nennt man *Sprache*. Im Wort wird durch den Ton und seinen Fall, die Stärke und den Rhythmus seines Erklingens das Wesen des Dinges symbolisiert, durch die Mundgeberde die begleitende Vorstellung, das Bild, die Erscheinung des Wesens.« <sup>41</sup> Dass Nietzsche hier von ›Wesen‹ und ›Erscheinung‹ redet, zeigt zugleich, dass für ihn das Verhältnis von Ton und Gebärde nicht symmetrisch ist, vielmehr der ›Tonuntergrund‹ der Sprache – mit den Worten von Corina Caduff – zum »alleinigen und universellen Ursprungsort« <sup>42</sup> wird.

Für uns aufschlussreich ist nun, dass das Doppelverhältnis von Bild und Musik zugleich Nietzsches Unterscheidung zwischen ›dionysisch‹ und ›apollinisch‹ Pate steht: Er betont, dass ihm Tonsprache und Gebärdensprache zum »Vorbild« werden für Unterscheidbarkeit und Zusammenspiel der dionysisch-apollonischen Künste. <sup>43</sup>

Diese Überkreuzung von Bild/Gebärdensprache/apollinisch einerseits mit Musik/Tonsprache/dionysisch andererseits, wirft ein aufschlussreiches Licht auf das Sprachverständnis von Nietzsche. Mit ›apollinisch‹ und ›dionysisch‹ artikuliert er zwei Ordnungsformen des Dynamischen in der künstlerischen Entfaltung. <sup>44</sup> Die apollinische Energie realisiert sich im Akt der Grenzziehung zwischen Objekten; überdies schafft sie Distanz zwischen dem Akteur und seinem Werk, lässt so erst ein (Künstler) Subjekt entstehen. Das Apollinische ist assoziiert mit Maß, also mit der ›Ratio‹. Zugleich stiftet es Individualität: es ist das ›principium individuationis‹. Anders nun das Dionysische. Die dionysische Energie bewirkt

das Einreißen der Grenzen, ihr geht das Maßvolle im Exzess abhanden, sie überspringt und negiert Distanz und lässt Individualität verloren gehen im gemeinschaftlichen Erleben; sie entfaltet Macht. Das Dionysische ist nicht assoziiert mit dem Bildlichen, vielmehr mit der Musik, mit der immer auch »erschütternden Macht des Tons«. <sup>45</sup>

Doch genau so, wie Nietzsche in der Sprache Gebärde und Ton zusammenwirken sieht, so fallen auch künstlerische Phänomene nicht *entweder* unter das ›Apollinische‹ oder das ›Dionysische‹. Vielmehr macht das komplexe Verhältnis beider erst aus, was uns als Kunst begegnet.

Für unsere Überlegung zur ›negativen Semiologie der Stimme‹ heißt dies: In der Musikalität des Lautlichen ereignet sich etwas, das mit dem Dionysischen verwandt ist, also zu tun hat mit Phänomenen von ›Macht‹, ›Distanzverlust‹ und der ›Sprengung von Individualität‹. Doch das, was sich da ereignet, vollzieht sich nur im Zusammenspiel mit seinem anderen, also mit dem Apollinischen an und in der Sprache. Damit finden wir bei Nietzsche eine spannungsreiche Differenz, eine Heterogenität zwischen Bild und Musik, zwischen Ratio und Macht, zwischen Verführung und Begriff angelegt, und zwar in Funktionsweise und Wirkung unserer Sprachlichkeit selbst.

Diese Unterscheidung und die Dynamik, die damit in den Blick kommt, begegnet uns wieder bei drei zeitgenössischen Autoren, wenn auch in anderer Terminologie. Es geht um die Differenz zwischen dem ›Symbolischen‹ und ›Semiotischen‹ bei Julia Kristeva, zwischen ›Phänogeesang‹ und ›Genogeesang‹ bei Roland Barthes sowie um Paul Zumthors Unterscheidung zwischen ›Text‹ und ›Werk‹. Alle drei Begriffspaare – das ist hier die Hypothese – können als Versionen von dem gedeutet werden, was Nietzsche mit dem Unterschied von Tonsprache und Gebärdensprache, von Dionysischem und Apollinischem meint. In dieser Perspektive bergen die Überlegungen Kristevas, Barthes' und Zumthors Kommentare zur subversiven Rolle der Stimme als Medium gegenüber der Sprache als System und als Kode.

## 9. ÜBER DIE SUBVERSIVITÄT DER STIMME ALS MEDIUM DER SPRACHE:

### KRISTEVA, BARTHES, ZUMTHOR

Für Julia Kristeva entfaltet sich das Sprachgeschehen aus dem Wechselverhältnis zweier Modalitäten: einmal geht es um eine präsymbolische Dimension, in der sich komplexe, gleichwohl unbestimmte Triebregungen noch vor allem ›Ausdruck‹ artikulieren und die sie ›das Semiotische‹ nennt; und zum andern geht es

um eine auf Repräsentation, Referenz und das Thetische gerichtete Dimension, die sie als ›das Symbolische‹ kennzeichnet. Das Semiotische ist dem Musikalischen und dem Stimmlichen verwandt. So »duldet« das Semiotische »keine andere Analogie als den Rhythmus von Stimme und Geste.«<sup>46</sup> Worauf es Kristeva ankommt ist, dass *jedes* Zeichensystem sowohl ›semiotisch‹ wie ›symbolisch‹ verfasst ist. Und das gilt erst recht für die Sprache: »Beide Modalitäten sind vom Prozess der Sinngebung, durch den sich Sprache konstituiert, nicht zu trennen.«<sup>47</sup> Allerdings agiert in diesem ›dialektischen Verhältnis‹ das Semiotische gegenüber dem Symbolischen auf bemerkenswerte Weise: denn einmal ist »die semiotische Funktionsweise Bedingung des Symbolischen«,<sup>48</sup> zum andern »muss das Semiotische als Negativität definiert werden, die in das Symbolische eingeschleust wird und seine Ordnung verletzt.«<sup>49</sup> Kristeva charakterisiert diese Negativität auch als ›Überschreitung‹ und ›Übertretung‹. Deren »Kunst‹ besteht gerade darin, mit der Negativität das Thetische zu überschreiten und aufzureiben und es dennoch nicht aufzugeben.<sup>50</sup> Der Differenz zwischen Semiotischem und Symbolischem gibt Kristeva auch eine auf Texte bezogene Version und bildet dafür die Begriffe »Genotext« und »Phänotext«.<sup>51</sup> Doch auch hier begegnet uns das Verhältnis von Negativität und Konstitution: Wieder verkörpert der Genotext Triebenergien, das Prozesshafte, das Flüchtige, die nicht-signifikante Strukturierung, während der Phänotext repräsentiert und zur Geltung bringt, was am Text der Kommunikation dient; und wieder bildet der Genotext einerseits die Grundlage für den Phänotext, verkörpert aber zugleich auch dessen Überschreitung.

Während bei Kristeva die Bezugnahme auf die Stimme eine noch vage, unausgearbeitete Assoziation bleibt, wendet Roland Barthes Kristevas Unterscheidung von ›Genotext‹ und ›Phänotext‹ unmittelbar auf die Erklärung der Stimme an.<sup>52</sup> Er versteht unter ›Genogesang‹ dasjenige an der Stimme, das »nichts mit der Kommunikation, der Darstellung (von Gefühlen) und dem Ausdruck zu tun hat«, dafür aber eine Art »signifikantes Spiel« ist.<sup>53</sup> Hier geht es um Volumen, um ›Rauheit‹, um die Physis der Stimme; ›Physis‹ von Barthes verstanden als »die Art und Weise, wie die Stimme im Körper sitzt – oder wie der Körper in der Stimme sitzt.«<sup>54</sup> Zum ›Phänogesang‹ nun zählt all das, was zur Struktur, zum Kode, zum Regelwerk von Sprache gehört, also der Kommunikation, dem Ausdruck und der Darstellung dient. Phänogesang und Genogesang verhalten sich dann zueinander wie Gesetz und Spiel, wie Sinn und Sinnlichkeit, wie Signifikat und Signifikanz, wie ein Kode und seine Subversion. Wechselverhältnis und Widerstreit dieser Dimensionen der Stimme zeigen sich an der Sprachäußerung in der Differenz von »Artikulation« und »Aussprache«.<sup>55</sup> Während die ›Artikulation‹ sich um die Klarheit des Sinns im Ausdruck bemüht, damit eine nahezu gleiche Lautstärke je-

dem Konsonanten zubilligen will und zugleich sich zu einer »Kunst der Ausdruckswirkung«<sup>56</sup> kodifizieren lässt, ist die ›Aussprache‹ stets singular, unabhängig fixierter Codes, sie verschleift Konsonanten, verleiht jeder Silbe eine einmalige Signifikanz in genau dieser einen partikulären Äußerung.<sup>57</sup> Es ist diese Dimension des ›Genogesangs‹, welche bewirkt, dass für Barthes die Stimmenwahrnehmung immer ein affektiver Prozess ist:<sup>58</sup> Eine Stimme zu hören, zieht unweigerlich eine Bewertung (évaluation) nach sich; so birgt für Barthes die Stimmlichkeit immer auch ein erotisches Element.

Der Mediävist Paul Zumthor hat das Spannungsverhältnis, um das Kristevas wie Barthes' Überlegungen kreisen, als Heterogenität von Stimme und Sprache gefasst und für die Oralitätsforschung fruchtbar gemacht. Antike wie mittelalterliche Epen sind für Zumthor nur performativ verstehbar, also als Aufführungsanweisungen. Diese ›Texte‹ wurden wirksam allein als konkrete ›Werke‹ (œuvre), die ein Sänger vor einem Publikum zur Aufführung brachte. Anders als im zu lesenden Text verbinden sich beim gehörten Werk Sprache, Klänge, optische Elemente.<sup>59</sup> Wenn der Text charakterisierbar ist durch seinen Sinngehalt, so fällt der Stimme in der Aufführungssituation die Aufgabe zu, »soziokorporelle« Aspekte zur Geltung zu bringen:

Der Einsatz der Stimme bewirkt fast immer eine Lockerung der linguistischen Zwänge. In ihm schlagen Spuren eines ›wilden Wissens‹ durch, die der Sprachfähigkeit selbst in Situationen konkreter Komplexität und intensiver persönlicher Beziehungen entströmen. Der mündlich produzierte Text stemmt sich in dem Maße, in dem er sich einer körpertragenden Stimme verdankt, mehr als der geschriebene Text gegen jede Wahrnehmung, die ihn aus seiner sozialen Funktion, seinem Ort in einer realen Gesellschaft, aus der Tradition, auf die er sich vielleicht beruft und aus den Umständen, unter denen er gehört wird, lösen möchte.<sup>60</sup>

Auch Zumthor, wie schon Barthes, entzieht der Stimme jene Neutralität, die ihr als Medium der Sprache allzu gerne verliehen wird. Stimme und Sprache gehen in der konkreten Redeäußerung ein Spannungsverhältnis ein:

Die Rede ist nicht einfach Vollstreckerin des Sprachsystems. Sie bestätigt nicht nur nicht vollständig dessen Vorgaben, sondern handelt ihm oft, in ihrer ganzen Körperlichkeit ... zuwider ... die Rede gibt sich als Erzählung, sie wird aber gleichzeitig im Klang der Stimme und der Bewegung des Körpers, die ihr Ausdruck verleihen, zum Kommentar des Erzählten.

Erzählung und Deutung – sie leben zusammen im Werk und spielen doch ihr eigenes, ihr eigenständiges Spiel.<sup>61</sup>

Wir beenden hier unsere Durchquerung der Überlegungen von Kristeva, Barthes und Zumthor. Unser Anliegen war es, das ›Semiotische‹ (auch: ›Genotext‹) und das ›Symbolische‹ (auch: ›Phänotext‹) bei Kristeva, ›Genogesang‹ (auch: ›Aus-sprache‹) und ›Phänogesang‹ (auch: ›Artikulation‹) bei Barthes und schließlich ›Werk‹ (auch: ›Rede‹) und ›Text‹ (auch: ›Sprachsystem‹) bei Zumthor als begriffliche Schemata zu verstehen, die eine Ähnlichkeit aufweisen mit Nietzsches Unterscheidung zwischen ›Dionysischem‹ (auch: ›Tonsprache‹) und ›Apollinischem‹ (auch: ›Begriffssprache‹). Doch nicht um die begriffsgeschichtlich und ideengeschichtlich aufschlussreiche Wanderung und Evolution einer Unterscheidung ist es uns hier zu tun. Unsere Vermutung ist vielmehr, dass sich die Differenz, um die es im binären Schema dieser Begriffe geht, auf das Verhältnis von Stimme und Sprache übertragen lässt. Dieser Vermutung nachzugehen, bleibt eine Forschungsaufgabe.

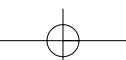
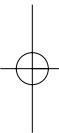
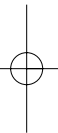
- 1 Roland Barthes: Die Musik, die Stimme, die Sprache, in: ders.: Der entgegenkommende und der stumpfe Sinn. Kritische Essays III, Frankfurt/M. 1990, S. 279–285 (hier: S. 280).
- 2 Die bemerkenswerteste Ausnahme ist Nietzsche. Wir kommen darauf zurück. In der zeitgenössischen Philosophie sind es vor allem die Phänomenologen, welche die Stimmlichkeit reflektierten, so Ihde in den USA (Don Ihde: Some Auditory Phenomena, in: Philosophy Today 10 (1966), S. 227–235) und Waldenfels in Deutschland (Bernhard Waldenfels: Hearing oneself speak. Derrida's Recording of the Phenomenological Voice, in: Leonard Lawlor (Hg.): Derrida's Interpretation of Husserl, Memphis 1994, S. 65–77; ders.: Antwortregister, Frankfurt/M. 1994; ders.: Vielstimmigkeit der Rede, Frankfurt/M. 1999). In einer nicht-phänomenologischen Perspektive in Frankreich: Barthes: Die Musik, die Stimme, die Sprache (Anm. 1) sowie in Deutschland: Dieter Mersch: Jenseits von Schrift. Die Performativität der Stimme, in: Dialektik 2 (2000), 79–92.
- 3 Exemplarisch: Hellmut Geißner: Sprechwissenschaft. Theorie der mündlichen Kommunikation, Königstein 1981; Jeff Pittam: Voice in Social Interaction. An Interdisciplinary Approach, London 1994.
- 4 Ferdinand de Saussure: Grundfragen der allgemeinen Sprachwissenschaft, Berlin 1967, S. 141 f.: »Übrigens ist es unmöglich, daß der Laut an sich, der nur ein materielles Element ist, der Sprache angehören könnte. Er ist für sie nur etwas Sekundäres, ein Stoff, mit dem sie umgeht. Das bezeichnende Element in der Sprache ist seinem Wesen nach keineswegs lautlich, es ist unkörperlich, es ist gebildet nicht durch seine stoffliche Substanz, sondern einzig durch die Verschiedenheiten, welche sein Lautbild von allen anderen trennen.«
- 5 Über die externe Sprache, also die empirisch beobachtbaren Sprachäußerungen führt Chomsky aus: »The E-language that was the object of study in most traditional or structuralist grammar or behavioral psychology is now regarded as an epiphenomenon at best.« Noam Chomsky: Knowledge of Language: It's Nature, Origin, and Use, Oxford/New York 1986, S. 25.
- 6 Aristoteles: De anima, zit. nach Mersch: Jenseits von Schrift (Anm. 2), S. 79, der dieses Zitat zum Motto seiner philosophischen Reflexion der Stimme macht.
- 7 Hans Lullies: Stimme und Sprache, in: E. Dunker u. a. (Hg.): Hören, Stimme, Gleichgewicht, München 1972, S. 215–257 (hier: S. 216 ff.).
- 8 Ob mit Männer- oder Frauenstimme, ob ›im Brustton‹ oder ›mit Kopfstimme‹, ob eher schrill oder sonor gesprochen wird, variiert mit geschlechtlichen und organischen Dispositionen ebenso wie mit dem aktuellen Muskeltonus, mit Druck und Geschwindigkeit des geformten Luftstromes.



- 9 Waldenfels: Antwortregister (Anm. 2), S. 490.
- 10 Darauf hat Kedl mit Nachdruck verwiesen: »I hear a train« and »I hear a sound« are not on a par« (S. 38) und »I cannot think of what I hear as being someplace other than where the thing is which makes the sound. I locate sounds by locating what it is which makes them« (G. Kent Kedl: Language. Sounds we use to Communicate, in: Philosophical Investigations 3 (1981), S. 26–43 (hier: S. 40).
- 11 Dazu Waldenfels: Antwortregister (Anm. 2), S. 489.
- 12 Zur Ereignishaftigkeit auch: Mersch: Jenseits von Schrift (Anm. 2), S. 88 ff.; Doris Kolesch: »Listen to the radio«. Artauds Radio-Stimme(n), in: Günter Ahrends u. a. (Hg.): Forum Modernes Theater, Bd. 14/2, Tübingen 1999, S. 115–143, (hier: S. 116 ff.).
- 13 Waldenfels: Hearing oneself speak (Anm. 2), S. 495.
- 14 Kolesch: Listen to the radio (Anm. 12), S. 119.
- 15 Wolfgang Senf: Anthropologische Gesichtspunkte der Stimme, in: Sprache-Stimme-Gehör 13 (1989), S. 19–25 (hier: S. 22).
- 16 Felix Trojan: Die Generatoren des stimmlichen Ausdrucks, in: Klaus R. Scherer (Hg.): Vokale Kommunikation, Weinheim/Basel 1982, S. 59–77.
- 17 Jean-Pierre Vernant: Mythos und Gesellschaft im alten Griechenland, Frankfurt/M 1987, S. 192; Mersch: Jenseits von Schrift (Anm. 2), S. 80 ff.
- 18 So sind Zeigen, Zeugen und Sagen im indogermanischen Sprachgebiet etymologisch verwandt.
- 19 Dazu: Sybille Krämer: Sprache, Sprechakt, Kommunikation. Sprachtheoretische Positionen im 20. Jahrhundert, Frankfurt/M. 2001.
- 20 Dieser Begriff ist von Apel (Karl O. Apel: Die Logosauszeichnung der menschlichen Sprache, in: Hans G. Bosshardt (Hg.): Perspektiven auf Sprache, Berlin/New York 1986, S. 45–87).
- 21 Mehr dazu: Krämer: Sprache, Sprechakt, Kommunikation (Anm. 19), S. 95 ff.
- 22 Ausführlicher: ebd., S. 103.
- 23 Jacques A. Miller: Jacques Lacan et la voix, in: Ivan Fonagy u. a. (Hg.): La Voix, Paris 1989, S. 180.
- 24 Dieter Mersch's Rehabilitation und Rekonstruktion der Stimmlichkeit zielt – im Anschluss an Roland Barthes – auf diese Überwindung der Idee vom Sprechen als bloßem Zeichengebrauch.
- 25 Mersch: Jenseits von Schrift (Anm. 2), S. 86.
- 26 Egon Aderhold: Künstlerisches Sprechen – Gekünsteltes Sprechen? Ein Beitrag zum Thema »Freisprechen und Nachsprechen«, in: Horst Gundermann (Hg.): Die Ausdruckswelt der Stimme. 1. Stuttgarter Stimmtage, Akademie für gesprochenes Wort, Heidelberg 1998, S. 51–62.
- 27 Dazu auch: Hede Helfrich: Satzmelodie und Sprachwahrnehmung. Psycholinguistische Untersuchungen zur Grundfrequenz, Berlin 1985.
- 28 Thea M. Mertz: Persona – per sonare oder meine Stimme – meine Stimmung, in: Siegrun Lemke/Susanne Thiel (Hg.): Sprechen. Reden. Mitteilen. Prozesse allgemeiner und spezifischer Sprechkultur, München/Basel 1996, S. 128–136. Lust- und Unlustgefühle teilen sich in der Stimme mit, insofern Rachenweite bzw. Muskelweite und Rachenenge bzw. Muskelanspannung unterschiedliche Stimmungen signalisieren. Dazu: Trojan: Die Generatoren des stimmlichen Ausdrucks (Anm. 16). Normalerweise realisiert sich eine Stimmlage als ein Mischklang zwischen Kopftönen und Brusttönen. Doch in Zuständen der Erregung kommt es zum Ausschlag jeweils eines der beiden Register: So markiert der Kopftönen eine Haltung der Defensivität, im Brusttönen jedoch artikulieren sich »Impo- niergehebe« und Aggressivität. Dazu: Hellmut Geißner: Über den Brustton der Überzeugung – zur Semantik des Imponiergehabes, in: Gundermann: Die Ausdruckswelt der Stimme (Anm. 26), S. 102–108.
- 29 Zu dieser Unterscheidung: Albrecht Riethmüller: »Stoff der Musik ist Klang und Körperbewegung«, in: Hans U. Gumbrecht/Karl L. Pfeiffer (Hg.): Materialität der Kommunikation, Frankfurt/M. 1988, S. 51–62 (hier: S. 56).
- 30 Zur musiké in der altgriechischen Bedeutung: Thrasybulos G. Georgiades: Nennen und Erklängen. Die Zeit als Logos, Göttingen 1985, S. 188 f.; ders.: Musik und Rhythmus bei den Griechen. Zum Ursprung der abendländischen Musik, Hamburg 1958, S. 41–48; Hermann Koller: Musik und Dichtung im alten Griechenland, Bern 1963, S. 5–16; Joseph Müller-Blattau: Das Verhältnis von Wort und Ton in der Geschichte der Musik. Grundzüge und Probleme, Stuttgart 1952, S. 8 ff.; schließlich Corina Caduff: Vom »Urgrund« zum Supplement. Musik in den Sprachtheorien von Rousseau, Nietzsche und Kristeva, in: Musik & Ästhetik 1 (1997) S. 37–55.
- 31 Rudolf Fietz: Medienphilosophie. Musik, Sprache und Schrift bei Friedrich Nietzsche, Würzburg 1992.

- 32 Friedrich Nietzsche: Die Unschuld des Werdens. Der Nachlass, ausgewählt und geordnet von A. Baeumler, Bd. I, Stuttgart 1956, 190, Fragment 508.
- 33 Solche Spuren finden sich im Zusammenhang der Entstehungsgeschichte der *Geburt der Tragödie* (1872), im vierten Teil der *Dionysischen Weltanschauung* (1870) und in einem Fragment aus dem Frühjahr 1871. Dieses Fragment erschien 1973 als Teil eines von Jacob Kraus herausgegebenen Gesamtextes ›Über Wort und Musik‹ (1871, Bruchstück). Verglichen mit der Kritischen Studienausgabe (KSA) handelt es sich bei dieser Fragment-Edition um den Zusammenschchnitt zweier Fragmente 7 und 12 aus dem Band 7 der KSA. Dazu: Caduff: Vom ›Urgrund‹ zum Supplement (Anm. 30), S. 45.
- 34 Nietzsche, Friedrich: Sämtliche Werke. Kritische Studienausgabe (KSA), hg. v. Giorgio Colli und Mazzino Montinari, Bd. 1, Berlin/München 1980, S. 572.
- 35 Ebd., Bd. 1, S. 572.
- 36 Ebd., Bd. 7, S. 362.
- 37 Ebd., Bd. 1, S. 575 f.
- 38 Ebd., Bd. 7, S. 379.
- 39 Ebd., Bd. 1, S. 572.
- 40 Ebd., S. 574.
- 41 Ebd., Bd. 7, S. 575.
- 42 Caduff: Vom ›Urgrund‹ zum Supplement (Anm. 30), S. 48.
- 43 Nietzsche: KSA (Anm. 34), Bd. 7, S. 360; dazu auch: »so gewiß nämlich der einzelne Ton, dem Bild gegenüber bereits dionysisch und das einzelne Bild, samt dem Begriff und das Wort, der Musik gegenüber apollinisch ist«.
- 44 Dazu: Alice Lagaay: *Metaphysics of Performance*, Berlin 2000, S. 51 ff.
- 45 Nietzsche zitiert nach: Fietz: *Medienphilosophie* (Anm. 31), S. 69.
- 46 Julia Kristeva: *Die Revolution der poetischen Sprache*, Frankfurt/M. 1978, S. 37.
- 47 Ebd., S. 35.
- 48 Ebd., S. 78.
- 49 Ebd., S. 78.
- 50 Ebd., S. 79.
- 51 Ebd., S. 94 ff.
- 52 Es geht hier insbesondere um Texte, die allesamt in deutscher Version zugänglich sind unter dem Titel *Der entgegenkommende und der stumpfe Sinn* (Anm. 1): Vor allem: (1) Die Rauheit der Stimme, S. 269–278; (2) Die Musik, die Stimme, die Sprache, S. 279–285.
- 53 Barthes: *Die Musik, die Stimme, die Sprache* (Anm. 1), S. 272.
- 54 Ebd., S. 284.
- 55 Ebd., S. 283.
- 56 Ebd.
- 57 »Die Artikulation, sagte er [Barthes bezieht sich hier auf den Sänger Panzera], ist die Vortäuschung und der Feind der Aussprache; man muß *aussprechen*, keineswegs *artikulieren* [...], denn die Artikulation negiert das *Legato*; sie will jedem Konsonanten die gleiche Lautstärke verleihen, wo doch in einem musikalischen Text kein Konsonant mit dem anderen identisch ist: Anstatt einem an sich und ein für allemal vorliegenden olympischen Coder der Phoneme zu entspringen, muß jede Silbe in die allgemeine Bedeutung des Satzes eingefaßt werden.« Ebd.
- 58 Kolesch: *Listen to the radio* (Anm. 12), S. 119.
- 59 Paul Zumthor: *Körper und Performanz*, in: Gumbrecht: *Materialität der Kommunikation* (Anm. 29), S. 703–713 (hier: S. 705).
- 60 Zumthor: *Körper und Performanz* (Anm. 59), S. 707.
- 61 Ebd., S. 709.

## II. STIMM-WECHSEL



**Heike Behrend**

## **GEISTERSTIMMEN IN AFRIKA: DIE STIMME ALS MEDIUM DER FREMDPRÄSENZ**

Als Emin Pascha, alias Eduard Schnitzer, im Jahre 1888 Westuganda besuchte, begegnete ihm eine Gruppe von Frauen:

Clothed in bark cloths, yellowish brown or dyed black [...] they also not infrequently wear skins of goats or sheep, occasionally cheetah or otter skins, and adorn or disfigure their heads with objects of every conceivable description. These ladies are certainly not beautiful, and they would hardly be eligible for vestal virgins, but they are feared, and therefore venture to take their liberties. As is always the case where professional interests are concerned, they vie with one another in eccentricities. One at Rionga's court grunted every minute; at Anfinas, one of them spoke in the highest falsetto, while another sat down beside one of the company and wanted her shoulders rubbed and her head bent.<sup>1</sup>

Emin Pascha entging, dass diese Frauen nicht nur Priesterinnen, sondern vor allem auch Geistmedien waren. Das exzentrische Grunzen und hohe Falsett, das er zu hören bekam, stammte von fremden Stimmen, den Stimmen von Geistern, die aus ihren Medien sprachen.

### **1. SELBSTPRÄSENZ UND FREMDPRÄSENZ**

In der westlichen philosophischen Tradition galt die Stimme vor allem als ein Medium der Selbstpräsenz. Die Stimme und damit das Sprechen waren sowohl Akt der Selbstvergewisserung eines Subjekts als auch dessen Entzweiung, ein sich Fremdwerden beim Hören der eigenen Stimme. Das Subjekt hob die Entzweiung jedoch wieder auf, indem es sich die entäußerte, sprechende Stimme im Prozess des Sich-Selbst-Zuhörens wieder aneignete.

In Opposition, aber auch als Komplement zu diesem Modell der Stimme als Medium der Selbstpräsenz möchte ich im Folgenden ›pagane<sup>2</sup> und christliche Praktiken der Geistbesessenheit in Afrika vorstellen, die Stimmen vor allem als Medium der Fremdpräsenz einsetzen. Geistbesessenheit, eine Körpertechnik, eingebunden in ein Ritual, lässt sich auch als eine Form indirekter Kommunikation beschreiben, weil im Zustand der Geistbesessenheit etwas gesagt wird, für

das der Sprecher nicht verantwortlich gemacht werden kann, denn es ist – zumindest aus lokaler Perspektive – der Geist, der spricht und nicht das Medium. Geistbesessenheit ist ein Diskurs des Nicht-Ich, der das Paradox vor Augen und Ohren führt, dass eine Person gleichzeitig ist und nicht ist, was sie vorgibt zu sein. Sie ist eine Praxis zu sprechen, ohne das Subjekt der Rede zu sein, und zu wissen, ohne Subjekt des Wissens zu sein. Die eigene Handlungs- und Redemacht wird aufgegeben, um eine fremde zu erlangen. Im Zustand der Geistbesessenheit erleidet der Besessene Sprache.

Aus der Perspektive unserer eigenen Kultur wirft die Praxis der Geistbesessenheit vor allem zwei Fragen auf: Zum einen die Frage nach dem Begriff der sozialen Person,<sup>3</sup> der die Instabilität und Dissoziation von Körper und Stimme erlaubt, und zum anderen die Frage nach ›agency‹,<sup>4</sup> nach der Beziehung zwischen intentionalem und erzwungenem Handeln, nach aktivem Handeln und dem Erleiden desselben.

Bevor ich verschiedene Beispiele christlicher Besessenheit innerhalb einer Laienorganisation der katholischen Kirche gebe, stelle ich zunächst eine Kosmologie der Kraft vor, die für das Verständnis von ›paganer‹ wie christlicher Geistbesessenheit unabdingbar ist. Dann werde ich vorkoloniale Praktiken der Geistbesessenheit rekonstruieren, um danach die hybriden Verbindungen von Fremd- und Selbstpräsenz in indigenisierten christlichen Besessenheitspraktiken zu beschreiben.

## 2. EINE ELEKTRISCHE KOSMOLOGIE DER KRAFT

In Westuganda – aber auch in anderen Regionen Afrikas – gibt es eine Vorstellung von Kraft oder Macht, eine Art pneumatisches Medium, das zirkuliert und das von Dingen, Tieren, Menschen und Worten Besitz ergreifen und sie dann ermächtigen kann. Alles kann potentiell von dieser Kraft erfasst, bewegt und somit verwandelt werden. Was von einer Person oder einem Ding vergeistigt wird, ist nicht der konkrete Gegenstand oder die historische Person, sondern ihr Generisches, Exemplarisches, ihr Wesen, ihr Inbegriff, eben ihr ›Geist‹.

Wie Elektrizität – ein Vergleich, den meine ugandischen Gesprächspartner immer wieder ins Spiel brachten – ist die Kraft ubiquitär und unsichtbar, aber verbindet, wenn aktiviert, durch nicht sichtbare Kanäle das vorher getrennte. Wird eine Person, ein Tier oder Ding von ihr ergriffen, dann wird sie oder es aufgeladen, aktiviert wie eine Batterie und gewinnt ›agency‹, Handlungsmacht. Personen können sich ausdehnen oder schrumpfen, je nachdem, ob und wie sie an der

Kraft partizipieren und mit Dingen und anderen Personen verbunden sind. Sie sind teilbar, ihre Grenzen eher durchlässig, und ein ständiger Austausch mit dem Außen und anderen Personen und Dingen findet statt.

Solange eine Person lebt, befindet sich ihre ›Lebenskraft‹, in Westuganda *mwoyo* oder ›Seele‹ genannt, in ihrer Brust oder im Zwerchfell. Als Teil der Person ist sie jedoch beweglich und kann sich im Zustand des Schlafes entfernen, und was sie erlebt, träumt der Schlafende. Im Zustand der Geistbesessenheit wird sie von einem Geist verdrängt und durch ihn ersetzt. Und beim Tod verlässt sie den Körper endgültig und verwandelt sich in einen Ahnengeist.<sup>5</sup>

Geister sind, so könnte man sagen, die mehr oder weniger personifizierte Form dieser Kraft oder Macht.<sup>6</sup> Sie war – zumindest in vorkolonialer Zeit – moralisch ambivalent: Je nachdem, wer sie benutzte, ob ein Heiler oder ein Zauberer, konnte sie zum Heilen wie auch zum Schaden eingesetzt werden. Sie kann akkumuliert, aber auch entzogen werden und entsprechend Störungen hervorrufen.

In dieser Kosmologie ist alles mehr oder weniger im Fluss. »Man kann nie sicher sein, dass etwas ist, was es zu sein scheint«, erklärte mir ein ehemaliger Heiler und Wahrsager, der zum Katholizismus konvertiert war. Sichtbares und Unsichtbares interferieren, und Sichtbarkeit ist nicht unbedingt mit Evidenz gekoppelt. Sichtbare Formen erscheinen eher als angehaltene Metamorphose,<sup>7</sup> und liminale Zustände wie Träume und Besessenheit ermöglichen, an der Verwandlung von Formen und Dimensionen zu partizipieren.<sup>8</sup>

Trotz Kolonisierung und Missionierung blieb diese Kosmologie der Kraft bis heute erhalten. Mit der Einführung von Elektrizität in den großen Städten während der Kolonialzeit fand sie zusätzliche Bestätigung und wurde, wie bereits erwähnt, vor dem Hintergrund des Elektrizitätsdiskurses reformuliert. Die äußeren Mächte wurden nicht abgeschafft, sondern beibehalten und sogar ausdifferenziert. Eine Naturalisierung und Medikalisierung des Körpers, die Schließung seiner Grenzen nach außen, fand nicht statt.<sup>9</sup>

### 3. INITIATION

Geistbesessenheit ist nur eine spezielle Form des Außer-Sich-Seins. Sie bildet zusammen mit Ergriffenheit, Inspiration, Trance und Traum ein Kontinuum verschiedener Zustände, die einen je unterschiedlich starken Bruch mit dem Alltagsleben vollziehen und oft zur Generierung eines besonderen Wissens dienen, das sich den Regeln normativer Rede entzieht. Im Zustand der Besessenheit oder Ergriffenheit manifestieren sich äußere Mächte, Geister oder Gottheiten, deren Prä-

senz die Frage nach der Bedeutung übersteigt. Sie werden durch Mimesis angeeignet, genauer, sie erzwingen die Nachahmung und damit die mehr oder weniger radikale Transformation der Person des Mediums.<sup>10</sup>

In Westuganda wurden und werden zum Teil noch immer – allerdings zur Zeit stark marginalisiert – Ahnengeister, die Geister von Kulturheroen und Fremdgeister in Kulturen verehrt und vergegenwärtigt. Sie alle verkörpern sich in Medien, die in einer langwierigen Initiation lernen, die Geister zu domestizieren und zum Heilen, Schaden oder zum Zweck der Divination einzusetzen.

Geistbesessenheit wird immer durch Leiden, eine Krankheit oder Krise, manchmal auch ein Trauma, ausgelöst. Heilung erlangt der oder die Besessene entweder durch einen Exorzismus oder durch die Anerkennung des Geistes und die Initiation in seinen Kult. Während der Initiation lernt dann der Initiand in einem komplexen Spiel von Unterwerfung und Ermächtigung, die äußere Macht, den Geist, der ihn ergriff, als eine (andere) soziale Person zu artikulieren.

So tauchte zum Beispiel im Zuge der christlichen Missionierung in Westuganda ein neuer Geist auf, der sich »Ich bete täglich zu Gott nannte.«<sup>11</sup> Sein Medium war offensichtlich vom ›Geist‹ des Christentums so stark beeindruckt worden, dass es sich gezwungen sah, im Zustand der Besessenheit christliche Rituale mimetisch zur Darstellung zu bringen. Doch war die Identität des neuen Geistes am Anfang noch unsicher. Erst nach und nach gab er sich in Träumen und Visionen zu erkennen und stellte seine Forderungen nach einer bestimmten Kleidung (eines Pastors), Parfüm, besonderen Nahrungsmitteln und anderen Paraphernalia, in denen er sich als soziale Person zunehmend materialisierte und sinnliche Präsenz – Hörbarkeit, Geruch und Sichtbarkeit – erlangte.

In einem zweiten Schritt lernte der Initiand dann, den Geist zu dissoziieren, also als Veräußerung der Ursache wieder abzutrennen und ihn in einen Schrein zu bannen. Heilung heißt also nicht das Verschwinden oder die Vertreibung des Geistes, sondern seine zyklische Adressierung, Aneignung und Bannung.

#### 4. DIE GEWALT DER STIMME

In Westuganda ist Geistbesessenheit eingebunden in ein multimediales Ritual, das auditive, visuelle und olfaktorische Medien – Gesang, Sprache, Trommeln, Rasseln, Parfüm und Tanz – auf vielfältige Weise verbindet.

Zu Beginn des Rituals – um den Geist zum Erscheinen zu bewegen – kommt der menschlichen Stimme eine entscheidende Bedeutung zu: Lieder, die dem Geist speziell gewidmet sind und ihn anlocken sollen, werden gesungen, wäh-



rend gleichzeitig das rhythmische Schütteln der Rasseln sowie das Schlagen der Trommeln den Zustand der Dissoziation beim Medium auslösen.<sup>12</sup> Die Musikinstrumente, insbesondere die Trommel, sind selbst oft dem Geist geweiht; sie erscheinen manchmal als seine Stimme, manchmal als sein Wohnort.<sup>13</sup>

Zusammen schaffen polyphoner Gesang, Trommeln und Rasseln einen Klangraum, der die Grenzen zwischen dieser Welt und der der Geister zu überbrücken sucht. Die Geister werden in diesen Klangraum einbezogen und können sich der Attraktion der Stimmen und Töne kaum entziehen. Stimmen und Tönen kommt hier also eindeutig Handlungsmacht, ›agency‹, zu. Sie drängen die Geister zu erscheinen.

Auch nachdem der Geist sich verkörpert hat, kontrollieren vor allem Trommel und Rasseln weiterhin seinen Tanz; sie feuern ihn an, beruhigen oder entlassen ihn aus der Besessenheit. Er ist ihnen ausgeliefert, sie sind seine Meister. Erst wenn Trommeln, Rasseln und Gesang verstummen, erhebt der Geist seine Stimme zur Divination.

## 5. DISKONTINUITÄT DER SOZIALEN PERSON

Die pagane Besessenheit markiert die Diskontinuität, den Bruch, zwischen alltäglicher und besessener Person auf radikale Weise. Die Transformation des Mediums in den Geist wird auf dramatische Weise in Szene gesetzt, denn der Prozess der Depersonalisierung, die Auslöschung, die die Person des Mediums erfährt, ist gewalttätig und schmerzhaft. So beginnt es, bevor der Geist ›in seinen Kopf steigt‹, bitterlich zu weinen, fällt dann zu Boden und erleidet einen ›kleinen Tod‹.<sup>14</sup>

Tatsächlich erfährt das Medium eine doppelte Unterwerfung: zum einen unter die Musik, die es bewegt und zum Tanzen zwingt, zum anderen unter den Geist, der aus ihm spricht und handelt.<sup>15</sup>

Auf den ›kleinen Tod‹ folgt die ›Auferstehung‹, die Verwandlung des Mediums in den Geist: Körperhaltung, Gestik, Mimik, Sprache und Stimme des Mediums verändern sich grundlegend und führen vor Augen und Ohren, dass nun ein Anderer, der Geist, im Medium präsent ist.

Auch das Ende der Besessenheit, das Verschwinden des Geistes und die Rückkehr zur alten Person, wird auf dramatische Weise in Szene gesetzt. Wieder wird der Bruch, die Diskontinuität, die das Medium erleidet, betont: Es fällt erneut zu Boden und liegt für einige Zeit bewegungslos, bis es langsam zu sich kommt, gähnt und sich die Augen reibt, als sei es gerade aus einem tiefen Schlaf erwacht.

Zur Markierung des radikalen Bruchs mit der Alltagsperson im Zustand der Besessenheit gehört auch die völlige Amnesie des Mediums. Es kann sich nach der Besessenheit an nichts erinnern, und Zuschauer oder Kultführer müssen ihm erzählen, was es als Geist tat und sagte. Diese Amnesie wird häufig auch noch theatralisch in Szene gesetzt, so, wenn das Medium als Geist Opfertgaben, z. B. eine gebratene Ziege in großen Mengen zu sich nimmt, nach der Besessenheit aber ostentativ über großen Hunger klagt und nach einem guten Mal verlangt.

Die radikale Selbstvergessenheit des Mediums tilgt natürlich auch den Akt der Selbstvergewisserung im Sich-Selbst-Zuhören. Tatsächlich ist die Fremdpräsenz des Geistes im Medium so total, dass es auch die fremde Stimme, die aus ihm spricht, nicht hört, zumindest nicht erinnern kann. Die eigene Person des Mediums ist vollkommen ausgelöscht; allein sein Körper dient als Gefäß des Geistes.

In der lokalen Theorie der Geistbesessenheit ist die Theatralisierung der Selbstvergessenheit des Mediums von zentraler Bedeutung, um trotz der Verkörperung des Geistes im Medium, also einer zeitlich begrenzten Identität der beiden, doch ihre Trennung, ihre Nicht-Identität, zu behaupten. Sobald Geister bzw. ihre Medien zu offensichtlich menschliche Interessen – also z. B. die des Mediums – verfolgen, verlieren sie den Anspruch auf Glaubwürdigkeit. Deshalb ist die Rede der Geister meist vieldeutig und von inneren Widersprüchen gekennzeichnet. Denn nur durch (scheinbare) Selbstlosigkeit können sie sich erfolgreich legitimieren und ihre Botschaft Wahrheit, Glaubwürdigkeit und Autorität beanspruchen.<sup>16</sup>

## 6. GEISTERSTIMMEN

In der lokalen Theorie wird die Stimme mit dem Atem und mit Nahrung in Verbindung gebracht. Luft und Nahrung werden durch den Mund einverleibt, und im Körper in Sprache oder Töne umgewandelt, die je nach Kraft oder Macht einer Person mehr oder weniger Wirkung entfaltet. Die Worte einer Person sind eine Substanz, die von Mund zu Mund weitergegeben wird.<sup>17</sup> Aus dem Mund einer Person entlassen, dringen sie in fremde Körper ein. Ihnen eignet eine Materialität, die es erlaubt, sie wie Gaben wechselseitig auszutauschen.

Wie mir verschiedene Gesprächspartner in Westuganda erklärten, besitzen Töne, mehr als der Blick, die Fähigkeit, das Herz bzw. den Magen, also das emotionale Zentrum einer Person (auch der Geister) zu bewegen. So wie Worte, die im Ärger ausgesprochen werden, die adressierte Person krank machen können, so können umgekehrt wohlmeinende, segnende Worte zur Heilung beitragen.

Um die Diskontinuität von Alltagsperson und Geist im Zustand der Beses-

senheit auf sprachlicher und stimmlicher Ebene deutlich zu machen, sprechen Geister in Westuganda ein eigenes esoterisches Geistervokabular, das während der Initiation erlernt wird. Grammatik und Syntax der lokalen Sprache bleiben erhalten, aber bestimmte Dinge und Handlungen werden mit Fremdwörtern bezeichnet, manche aus dem Swahili oder aus dem Luganda.<sup>18</sup> Diese esoterische Sprache ist standardisiert, semantisch nicht entleert, und Leute, auch Nicht-Initiierte, die häufig an Besessenheitsritualen teilnehmen, können sie verstehen. Außerdem markieren Geister, wie schon Emin Pascha 1888 feststellte, ihre Präsenz stimmlich, indem sie im Falsett sprechen und in ihre Rede immer wieder ein kleines Grunzen oder Stöhnen einfügen, zum Beispiel ein exzentrisches »nunu!«<sup>19</sup>

## 7. STIMMENKLAU

Als John Beattie, ein englischer Ethnologe, in den 1950er Jahren in Westuganda forschte, begegnete ihm zu einer Zeit, als das Telefon in den größeren Städten in Uganda eingeführt wurde, eine neue Praxis der Divination.<sup>20</sup> Mit Hilfe des Horns eines Büffels, einer Antilope oder Ziege sowie einer bestimmten Medizin konnten Heiler und Wahrsager eine eher unpersönliche Kraft generieren und aussenden, um die Stimme einer entfernten Person zu stehlen und zu befragen. Die gestohlene Stimme sprach dann durch das Horn und beantwortete die Fragen des Heilers. (»Hast Du X verhext?« – »Ja, ich habe X verhext« oder »ich habe es nicht getan«). Nach der Befragung kehrte die Stimme zu ihrem Besitzer zurück, der meist schlafend, manchmal vom Stimmenklau träumend, am nächsten Morgen geschwächt erwachte. (Wenn man im Traum redet, dann deutet das auf Stimmenklau hin, erzählte mir eine Frau, der die Stimme gestohlen worden war.) Auch Zauberer konnten Stimmen rauben: Sie sperrten die gestohlene Stimme in einen Topf und zwangen sie zu sprechen.

Die Praxis des Stimmenklus, offensichtlich auch eine lokale Aneignung des Telefons,<sup>21</sup> war Teil einer Geständnistechologie, die der Generierung von in gewisser Weise subjektlosem Wissen diente. Was der Heiler nicht selbst sagen konnte, dass nämlich diese oder jene Person eine Hexe sei, bestätigte bzw. negierte die geraubte Stimme. Die Trennung von Körper und Stimme ermöglichte also eine entpersönlichte Geständnistechologie ohne Subjektivierungseffekt.

## 8. ZUR GESCHICHTE DER GEISTBESESSENHEIT IN WESTUGANDA

Während Geistmedien in Ostafrika in vorkolonialer Zeit einen hohen Status genossen – sie bildeten in den interlakustrinen Königstümern eine Art Gegenmacht zur Macht des Königs – wurden sie während der Kolonialzeit und insbesondere durch die christliche Missionierung zum Inbegriff des Primitiven und Vor-Modernen. Vor dem Hintergrund der Einführung der Schrift durch christliche Missionare und den kolonialen Staat, der »Ewigkeitsmaschine«, wie McLuhan sie nennt, die zum Garanten von Geschichtlichkeit, Hochreligion und Moderne wurde, standen Geistmedien bald für das Andere der Schrift, nämlich für Oralität und Ritualität. Sie wurden als Scharlatane und Betrüger verfolgt und konnten oft nur im Untergrund arbeiten.

Die christlichen – anglikanischen und katholischen – Missionare, die Ende des 19. Jahrhunderts aus England und Frankreich nach Westuganda kamen, sahen in der lokalen Religion vor allem Teufelsanbetung und versuchten, diese Sicht auch den Bewohnern beizubringen. Afrikanische Gottheiten, Geister und Hexen erschienen ihnen als Agenten des Bösen, die die Etablierung des christlichen Ideals eines selbstverantwortlichen Individuums hintertrieben. Obwohl gegen Ende des 19. Jahrhunderts in Europa der Spiritismus – auch innerhalb der christlichen Kirchen – eine Blüte erlebte und damit auch Praktiken der Geistbesessenheit eine Renaissance erfuhren, griffen die christlichen Missionare in Afrika nicht auf diese zurück, sondern diabolisierten stattdessen die afrikanischen Geister und konstruierten sie – trotz vieler Gemeinsamkeiten – als das Andere des Christentums.

Während in vorkolonialer Zeit Geister in Westuganda moralisch ambivalent waren, sie sowohl zum Guten wie zum Bösen, zum Heilen wie zum Schaden und Töten benutzt werden konnten, spalteten die christlichen Missionare diese Ambivalenz auf in einen Dualismus, in dem der christliche Gott für das Gute stand und all die in sich ambivalenten afrikanischen Geister und Gottheiten nun eindeutig mit dem Bösen identifiziert wurden. Dadurch trugen sie zur Produktion und Vermehrung böser oder satanischer Geister bei; sie erzeugten, was sie bekämpften. Denn anders als in Europa wurde die Realität von Geistern, Teufeln und Dämonen zumindest von der Elite nicht negiert, sondern gerade in der Gesellschaft anerkannt und reproduziert.

Erst in postkolonialer Zeit gelang es Geistmedien und »traditionellen« Heilern in Uganda, wieder Macht zu gewinnen und mit den postkolonialen Regierungen zahlreiche Allianzen einzugehen. Während der postkoloniale Staat sich zum Hüter von »Traditionen« und damit auch den »traditionellen Religionen« erklärte, haben dagegen christliche Kirchen in den letzten Jahren verstärkt allem

›Traditionellen‹, das sie mit dem Paganen gleichsetzten, den Kampf angesagt. Dieser Prozess steigerte sich, als zu Beginn der 1990er Jahre auch Westuganda von einer globalen Welle christlichen Fundamentalismus erfasst wurde. Der Dualismus, den die europäischen Missionare eingeführt hatten, wurde von afrikanischen Christen übernommen und nun radikalisiert. In dem Versuch, eine christliche Moderne in Afrika zu etablieren, aktualisierten sie gerade in ihrer Opposition eine Vielzahl okkultur Mächte und trugen wesentlich zu deren Ausdifferenzierung bei.

Gleichzeitig fand aber auch eine verstärkte Popularisierung des Christentums statt. Weniger der Klerus als vielmehr Laien trieben die christliche Erneuerung voran. ›Gegen den toten Buchstaben‹ versuchten pfingstlerische und charismatische Christen in Westuganda, wieder ›den lebendigen Geist‹ einzuführen und griffen dabei zunehmend auch auf ekstatische Praktiken wie Zungenreden, Visionen und Auditionen sowie die Heilung durch Exorzismus zurück. Diese wiederum gingen hybride Verbindungen mit den lokalen Traditionen der Besessenheit ein. Im Zuge dieser ›Vergeistigung‹ löste sich der heilige Geist aus der Trinität und gewann zentrale Bedeutung. Und obwohl gerade die charismatischen und pfingstlerischen Kirchen und Bewegungen Geistbesessenheit als pagane und satanische Praxis radikal ablehnten, kehrte das Verdrängte manchmal sehr un-zweideutig im christlichen Kontext wieder. So wurden zum Beispiel innerhalb der katholischen Kirche in Westuganda Märtyrer ›vergeistigt‹ und in christliche Besessenheitskulte integriert, die stark den vorkolonialen Cwezi-Kulten ähneln.<sup>22</sup>

## 9. KREUZZUG

Im Jahr 1998 begann ich, in Westuganda über eine katholische Laienorganisation, die Gilde der Ugandischen Märtyrer (GUM),<sup>23</sup> ethnografisch zu arbeiten. Seit Beginn der 1990er Jahre hatten auch in ihr charismatische Fundamentalisten an Einfluss gewonnen, die gleichzeitig christliche Praktiken verstärkt lokalisierten und indigenisierten und somit eine Version des Katholizismus schufen, die Teilen des orthodoxen Klerus nicht wenig zu schaffen machte. Als dann 1996 die GUM unter der Führung von Lawrence Kassaija, einem ehemaligen Boxer, anfang, mit Hilfe des heiligen Geistes Hexen und andere Agenten Satans zu jagen, konnte sie zahlreiche, an unabhängige Kirchen verlorene ›Seelen‹ wieder in die katholische Kirche zurückholen und erheblich an politischer Macht gewinnen.

Die von ihr organisierten Hexenjagden – Kreuzzüge genannt – knüpften so-

wohl an vorkoloniale Praktiken der Hexenfindung als auch an christlichen Techniken der Inquisition an. Die ersten ›Kreuzzüge‹ der GUM verliefen, wie mir erzählt wurde, äußerst gewalttätig. So identifizierten 1996 die GUM-Mitglieder in einem Dorf eine ältere Frau als notorische Hexe, die sich sofort in eine schwarze Katze verwandelte. Als sie verbrannt werden sollte, wechselte sie wiederum ihre Gestalt, wurde menschlich und musste mit schweren Brandwunden ins Krankenhaus gebracht werden. Sie verklagte die GUM. Darauf kamen einige der an der Hexenjagd beteiligten, unter ihnen auch Lawrence Kassaija, ins Gefängnis. Nach diesem Zwischenfall verbot die katholische Kirche den Anführern erst einmal weitere Kreuzzüge. Zwei Jahre lang lernten die GUM-Mitglieder in Workshops Bibelfestigkeit und den gewaltfreien Umgang mit Hexen. Erst danach durften sie wieder zu neuen Kreuzzügen aufbrechen.

Am 24. und 25. August 2002 führte die GUM einen vom Bischof genehmigten Kreuzzug in Kyamiaga im Buhesi-subcounty durch,<sup>24</sup> in dessen Verlauf es zu verschiedenen Formen von christlicher und satanischer Geistbesessenheit kam, die ich im Folgenden vorstellen möchte.

Der Kreuzzug begann in einer Kirche, in der sich die Mitglieder in der Nacht versammelten. Durch Fasten, Gebete, Gesang und Tanz ›konzentrierten‹ und steigerten sie die Kraft des heiligen Geistes, der ihre Körper wie eine elektrische Batterie ›füllte‹ und ›auflud‹. Wie bei der paganen Geistbesessenheit kam dabei den akustischen Medien eine besondere Bedeutung zu. Christliche Lieder, Gebete und Trommeln schufen auch hier einen Klangraum, der die Grenzen zwischen dieser Welt und der Gottes sowie des heiligen Geistes überbrückte. Es entstand eine Klanggemeinschaft, die weniger auf Gleichklang als vielmehr auf einer komplexen, oft dialogischen Polyphonie basierte. Offensichtlich fand außerdem noch ein Rückkopplungsprozess statt, ein Effekt, der durch die Rückwirkung der Akustik des Raumes, seiner Resonanz, auf die singenden und tanzenden Personen ausgelöst wurde und wohl auch Trance-fördernd wirkte.<sup>25</sup> Bestimmte Lieder, insbesondere diejenigen, die den heiligen Geist besangen, riefen denn auch bei einigen Mitgliedern eine Dissoziation hervor, die zur Glossolie führte.

Mit ›Waffen‹ ausgerüstet, neben Gebet und Gesang auch der Bibel und mit Weihwasser gefüllten Flaschen, bildeten sie am nächsten Morgen Gruppen von 20 bis 30 Frauen und Männern, die dann singend von Haus zu Haus zogen. Gelangten sie zu einem Haus, in dem ›satanische Dinge‹ versteckt waren, ›entlud‹ sich die Kraft des heiligen Geistes bei einigen der Mitglieder, die dann je nach Stärke des Satanischen nur leicht zitterten oder zu Boden fielen, schrienen und sich auf der Erde wälzten. Ihr mit dem heiligen Geist gefüllter Körper wurde so zu einem Indikator des Satanischen.

Zu Beginn der ›healing session‹, die am nächsten Tag stattfand, fassten sich die Mitglieder mit der Gabe zu heilen bei den Händen und bildeten einen Kreis, um die Kraft wie einen Stromkreis kurzzuschließen und so ihre Intensität zu erhöhen. Sie wurden zu ›instruments of god‹, die durch die Berührung – meist Handauflegen – von Kranken, von satanisch Besessenen, die bösen Geister zwingen, sich zu erkennen zu geben. Damit schlossen sie einerseits an die lange christliche Tradition einer Theologie der Elektrizität an, die die göttliche Kraft als Zeichen einer verborgenen Gegenwart Gottes in der Welt versteht.<sup>26</sup> Andererseits setzten sie die Kraft des heiligen Geistes aber auch wie beim paganen Stimmenklauf ein, um die satanischen Geister zu einem Geständnis zu zwingen und sie nach einem meist dramatischen Kampf zu exorzieren.

#### 10. SATANISCHE GEISTERSTIMMEN

Die satanischen Geisterstimmen wiesen eine starke Kontinuität zu Formen und Erscheinungsweisen vorkolonialer, ›paganer‹ Geister auf. So ertönten aus einem Mann, der von einem Rivalen verhext worden war, zuerst Stimmen von Tiergeistern, von Hund, Katze, Ziege, Schwein und Affe. Diese Tiere gelten als unrein; Affe und Katze sind außerdem noch Tiere der Wildnis, eines Anderen der eigenen Kultur, das auch mit Hexerei assoziiert wird. Gleichzeitig evozierten Schweinegrunzen und Hundegeheul auch christliche Klangbilder der Hölle, »Höllengläuberei«.<sup>27</sup>

Nachdem die Tiergeister durch den Einsatz der ›Waffen‹ – fromme Lieder, Gebete und Weihwasser – exorziert worden waren, meldeten sich menschliche Stimmen in dem Mann zu Wort: Zuerst sein eigener Geist – etwas, was mir bisher bei meinen Studien zur Geistbesessenheit noch nicht untergekommen ist – und danach die Stimme seines Feindes, des Mannes, der ihn verhext hatte. Während die Stimme des eigenen Geistes seine Unschuld beteuerte und die eigene Moralität betonte (»May I tell you why they did this to me? I preach well, I speak with Christians very well. They want to send me away. I do not love girls. There are some beautiful girls but I do not make love to them!«), attackierte ihn die Stimme des Gegners<sup>28</sup> (»You are only proud. You cannot marry those beautiful girls. We were catechists and we used to marry those girls. I married Kabasweka«) und gestand ein, mit Mädchen, die ihm anvertraut worden waren, geschlechtliche Beziehungen eingegangen zu sein. Obwohl die Stimme des Gegners nicht eigentlich bekannte, ihn verhext zu haben, war aber für alle Beteiligten, mit denen ich sprach, seine Rede ein Eingeständnis auch dieser Schuld.

Es fand hier also mit Hilfe des heiligen Geistes eine Divination statt; er

zwang, wie beim Stimmenklau, jedoch nicht nur den Verursacher des Leidens aus dem Körper des Leidenden zu sprechen und seine Schuld einzugestehen, sondern auch das Opfer selbst. Der Körper des Besessenen wurde zum Ort einer Mehrstimmigkeit, eines ›social drama‹. Die Reden der beiden machten also auf indirekte Weise, eben durch Geisterstimmen, einen privaten Konflikt öffentlich. Tatsächlich ist es höchst gefährlich, jemanden der Hexerei anzuklagen. Hexerei ist ein Kampf auf Leben und Tod. Es besteht immer die Gefahr, dass der Hexer sich als der Stärkere erweist und sich am Ankläger rächt. Die angeklagte Person kann aber auch unschuldig sein, und dann richtet sich die zerstörerische Kraft gegen den Ankläger. Sprechen dagegen Geister die Anklage aus, dann liegt die Verantwortung für das Gesagte bei ihnen, und der Kranke und Leidende kann nicht zur Rechenschaft gezogen werden, wird also entlastet.

Während im paganen Kontext den Stimmen der Geister nur eine relative Wahrheit zugesprochen wird, solange sie überzeugend die Interesselosigkeit desjenigen, der spricht, also die Trennung von Geist und Medium vorführen, wird nun im christlichen Kontext eine absolute Wahrheit behauptet, die sich auf die hegemoniale Position des heiligen Geistes als eine Art Supermacht beruft. Kann der Rede paganer Geister immer widersprochen werden, so garantiert nun die Präsenz des einen heiligen Geistes die Wahrheit des Gesagten. Es handelt sich hier also offensichtlich um einen Wahrheitsbegriff, der sich auf die Autorität und den Status der Quelle des Gesagten verlässt und weniger das Gesagte einer Prüfung unterzieht.<sup>29</sup>

Im Gegensatz zur paganen Geistbesessenheit wird im christlichen Kontext die Diskontinuität zwischen Alltagspersona und der besessenen zwar nicht aufgehoben, aber doch stark reduziert. Die Präsenz satanischer Geister bleibt unzweideutig, aber die Stimmen der Geister (außer bei Tiergeistern) unterscheiden sich nicht wesentlich von der Alltagsstimme des Mediums. Das esoterische Geistervokabular sowie das exzentrische Gurren und Falsett sind verschwunden. Obwohl die satanischen Geister höchst obszöne Reden führen und auf das abscheulichste Fluchen (so dass meine Mitarbeiterinnen sich anfangs weigerten, mir das Gesagte zu übersetzen), hat offensichtlich eine ›Stimmenglättung‹, eine stimmliche Entdifferenzierung und Disziplinierung, stattgefunden, die ich als Versuch deuten möchte, im christlichen Kontext die Kontinuität der sozialen Person zu erhalten. Trotz Fremdpräsenz garantiert die Stetigkeit und Fortdauer der Stimme die Selbstpräsenz der Person des Mediums. Der heilige Geist als Autor und Ausleger wird zum Garanten der Kontinuität der christlichen Persona. Dem entspricht die Möglichkeit, nun auch vom eigenen Geist – und nicht nur von einem fremden – ergriffen zu werden.



Während die pagane Geistbesessenheit die Fremdpräsenz eines Geistes in einen Kult integriert und also eine oft lebenslange Beziehung zwischen Geist und Medium etabliert, werden satanische Geister im christlichen Kontext nur aufgerufen, um vertrieben zu werden. Ihre Präsenz wird vorgeführt, um ihre Austreibung und damit die Mächtigkeit und den Sieg des heiligen Geistes um so eindringlicher zu bezeugen. Aus christlicher Perspektive ist die Fremdpräsenz zeitlich begrenzt, Ausdruck einer temporären Krise oder Krankheit, die behoben wird, um die verlorene Einheit der christlichen Person wieder herzustellen.

Aber die christliche Vorstellung von der Person birgt die Gefahr, dass eine solche Krise der Präsenz sich wiederholt und chronisch wird, weil ihr die zyklische Form der Aneignung und Bannung nicht mehr zur Verfügung steht.

## 11. DIE STIMMEN DES HEILIGEN GEISTES

Die unvergleichliche Macht des christlichen Gottes sowie des heiligen Geistes verbietet Mimesis. Man kann diesen Gott und Geist nicht imitieren, also auch im Zustand der Besessenheit nicht darstellen.<sup>30</sup>

Entsprechend zurückgenommen sind denn auch die Formen der Verkörperung und die Stimmen des heiligen Geistes. So ist die Glossolalie, die auch innerhalb der GUM praktiziert wird, im Gegensatz zur Sprache paganer und satanischer Geister nicht übersetzbar. Es geht hier offensichtlich weniger um Kommunikation, als vielmehr darum, die Präsenz des heiligen Geistes vorzuführen. Es handelt sich um eine phatische Form der Kommunikation, um eine Übertragung, deren Botschaft die Tatsache der Übertragung ist.

Während des Kreuzzugs ereigneten sich jedoch noch andere Formen der Vergegenwärtigung des heiligen Geistes, die zeigen, wie stark er lokalisiert wurde und welchen sehr direkten Anschluss er an pagane und satanische Formen der Besessenheit gefunden hat. Als die GUM-Mitglieder in einem Gehöft einen »satanischen« Baum entdeckten, machte sich der heilige Geist im Körper einer Frau bemerkbar, die zu zittern anfang, die Arme ausbreitete als ob gekreuzigt und dabei tierische Laute ausstieß, die gefährlich in die Nähe satanischer Tiergeister wiesen. Tatsächlich waren die vor-sprachlichen Laute, die sie von sich gab, die Laute einer Taube, des Vogels, der aufs Engste mit der Ikonografie des heiligen Geistes verbunden ist, gleichzeitig aber auch an die pagane Besessenheit durch Tiergeister anschloss.

Die Besessenheit durch den heiligen Geist erfolgte hier als eine nur partielle Transformation oder besser, eine Form der Koexistenz von Selbst- und Fremdprä-

senz, die die Kontinuität und Einheit der Person nicht mehr grundsätzlich gefährdet. Dem entspricht auch, dass die totale Amnesie, die zur paganen sowie satanischen Besessenheit gehört, beim heiligen Geist durch Formen einer nur partiellen Selbstvergessenheit abgelöst wird. Die vom heiligen Geist Ergriffenen, so erklärten mir die GUM-Mitglieder, bleiben zu 50% bei Bewusstsein. So wie man vom Halbschlaf spricht, könnte man hier von einer Halbpräsenz oder Halbabsenz der vom heiligen Geist ergriffenen sprechen. Sind sie vom heiligen Geist erfüllt, dann können sie zwar nicht mehr aus eigener Kraft sprechen, aber sie nehmen sehr wohl wahr, was sich in ihrem Umkreis ereignet und können sich später auch daran erinnern. Somit ermöglicht die Besessenheit durch den heiligen Geist ein Reflexivwerden der Kommunikation, eine teilweise Selbstvergewisserung. Sie nähert sich damit an das christliche Ideal eines selbstverantwortlichen Subjekts an.

Während in Europa zumindest teilweise die Entkörperlichung der Stimme, ihre Abkopplung vom Körper, zugunsten einer inneren Anschauung gelang, konnte sich der christliche, unkörperliche Begriff des Verstehens innerhalb der GUM nicht durchsetzen.<sup>31</sup> Obwohl die soziale Person mit den vielen fremden Stimmen zwar vereinheitlicht und die radikale Diskontinuität im Zustand der paganen Geistbesessenheit zugunsten einer nur ›halben‹ Fremdpräsenz im christlichen Kontext abgeschwächt wurde, blieb doch der afrikanische Körper – nicht nur der pagane, sondern auch der christliche – weiterhin offen für die Invasion äußerer Mächte. Freilich gewann die christliche Person an Handlungsmacht und Reflexivität, doch ihre Autonomie wird weiterhin durch äußere Mächte eingeschränkt und bedroht. Vor allem aber verlor sie weitgehend ihre Wandlungsfähigkeit, die Möglichkeit, im Ritual ein radikal Anderer zu werden und als der Andere zu sprechen.

Die Forschung in Westuganda wurde finanziert von der VW-Stiftung, dem Ministerium für Wissenschaft und Forschung in NRW sowie dem SFB/FK 427 »Medien und kulturelle Kommunikation«, wofür ich an dieser Stelle herzlich danken möchte. Danken möchte ich auch Cornelia Epping-Jäger, Erika Linz und Erhard Schüttpez für Hinweise, Kritik und freundliche Unterstützung.

- 1 Emin Pasha: *Emin Pasha in Central Africa*, London 1988, S. 285.
- 2 Die Bezeichnung dieser Besessenheitskulte als ›pagan‹ ist nicht sehr glücklich, weil sie so in negativer Weise aus christlicher Perspektive definiert werden. Die Alternativen ›traditionell‹ oder ›vorkolonial‹ wären aber noch problematischer, weil damit die Modernität dieser Kulte unterschlagen würde.
- 3 Zum Begriff der sozialen Person vgl. z. B. Marcel Mauss: *Eine Kategorie des menschlichen Geistes. Der Begriff der Person und des ›Ich‹* [1938], in: ders.: *Soziologie und Anthropologie*. Bd. II, Berlin 1978, S. 223–252; Germaine Dieterlen (Hg.): *La Notion de la Personne en Afrique Noire*, Paris 1981 und den Artikel von Fritz Kramer: *Identität*, in: Bernhard Streck (Hg.): *Wörterbuch der Ethnologie*, Köln 1987.
- 4 Der nur schwer ins Deutsche zu übersetzende Begriff ›agency‹ ist der Versuch, Handeln zu theoretisieren. Gegen den semiotischen Ansatz liegt hier die Betonung auf dem Tun, dem Handeln, der In-

- tention. Der Begriff geht u. a. zurück auf »Die Gabe« von Marcel Mauss (Marcel Mauss: Die Gabe, Frankfurt/M. 2001), die Gabe als Teil der sozialen Person, die »agency« hat, also Subjektstatus erlangen oder Objektstatus erleiden kann. Man könnte ihn aber auch auf Edward Byrnett Tylors Animismustheorie (Primitive Culture, London 1871) zurückführen. In neuerer Zeit haben Godfrey Lienhardt (Divinity and Experience. The Religion of the Dinka, Oxford 1961) und Fritz Kramer (Der Rote Fez. Über Besessenheit und Kunst in Afrika, Frankfurt/M. 1987) mit der Theorie der »Passiones«, Marilyn Strathern (The Gender of the Gift: Problems with Women and Problems with Society in Melanesia, Berkeley 1987) und Alfred Gell (Art and Agency. An Anthropological Theory, Oxford 1998) den Begriff weiter entwickelt.
- 5 John Beattie: The Ghost Cult in Bunyoro, in: Ethnology 3 (1964), S. 127–151 (hier: S. 127).
  - 6 John Beattie: Spirit Mediumship in Bunyoro, in: John Beattie/John Middleton (Hg): Spirit Mediumship and Society in Africa, London 1969, S. 159–170 (hier: S. 170).
  - 7 Barbara Keifenheim: Zur Bedeutung Drogen-induzierter Wahrnehmungsveränderungen bei den Kashinawa-Indianern Ost-Perus, in: Anthropos 94 (1999), S. 501–514 (hier: S. 501). Ich möchte hier anmerken, dass diese Konzeptionen in Südamerika entwickelt wurden, aber auch für einige Regionen in Afrika Gültigkeit beanspruchen können. Dies um so mehr, als z. B. neue Medientechnologie wie Video in Ghana und Nigeria mit Hilfe von Special Effects verstärkt erlaubt, die vorher unsichtbaren Metamorphosen von Mensch, Tier und Geist nun zu visualisieren.
  - 8 Ebd., S. 502.
  - 9 Luise White: The Needle and the State. Immunization and Inoculation in Africa, or, the Practice of Un-national Sovereignty, 1997 [unveröffentlichtes Manuskript].
  - 10 Fritz Kramer: Der rote Fez. Über Besessenheit in Afrika, Frankfurt/M. 1987.
  - 11 John Beattie: Group Aspects of the Nyoro Spirit Mediumship Cult, in: Rhodes-Livingston Journal 30 (1961), S. 11–38 (hier: S. 29).
  - 12 John Beattie: Initiation into the Cwezi Spirit Possession Cult in Bunyoro, in: African Studies 16/3 (1957), S. 150–161 (hier: S. 153).
  - 13 Gilbert Rouget: La musique et la transe [1980], Paris 1990, S. 221.
  - 14 Beattie: The Ghost Cult in Bunyoro (Anm. 5), S. 132.
  - 15 Rouget: La musique et la transe (Anm. 13), S. 215.
  - 16 Heike Behrend: Alice und die Geister. Krieg im Norden Ugandas, München 1993, S. 157 ff.
  - 17 Rosalind Morris: In the Place of Origins. Modernity and its Mediums in Northern Thailand, Durham/London 2000, S. 300.
  - 18 Beattie: Initiation into the Cwezi Spirit Possession Cult (Anm. 12), S. 157.
  - 19 John Beattie: Consulting a Diviner. A Text, in: Ethnology 5 (1966), S. 202–217 (hier: S. 208).
  - 20 John Beattie: Divination in Bunyoro, in: Sociologus 14/1 (1964), S. 44–62 (hier: S. 56 f.).
  - 21 Heike Behrend: »Call and Kill«. Zur Verzauberung und Entzauberung westlicher technischer Medien in Afrika, in: Erhard Schüttpeitz/Albert Kümmel (Hg): Signale der Störung, München 2003, S. 287–300.
  - 22 Beattie: Initiation into the Cwezi Spirit Possession Cult (Anm. 12).
  - 23 Für weitere Informationen über die GUM siehe Heike Behrend: Man ist, was man isst. Variationen der Eucharistie. Katholiken und Kannibalen in Westuganda, in: Kursbuch 143 (2001), S. 167–183.
  - 24 Ein Videofilm mit dem Titel »Satan Crucified« über diesen Kreuzzug ist in Arbeit.
  - 25 Sabine Rittner: Singen und Trance – die Stimme als Medium zur Induktion veränderter Bewusstseinszustände, in: Horst Gundermann (Hg): Die Ausdruckswelt der Stimme, Stuttgart 1998, S. 317–325 (hier: S. 319).
  - 26 Ernst Benz: Theologie der Elektrizität, Wiesbaden 1970.
  - 27 Karl-Heinz Göttert: Geschichte der Stimme, München 1998, S. 118.
  - 28 Soweit ich in Erfahrung bringen konnte, arbeiteten beide Männer als Katechisten; da aber der eine die ihm anvertrauten Mädchen und Frauen verführt hatte, war er entlassen worden und der andere hatte seine Stelle übernommen. Darauf rächte sich der Entlassene und verhexte seinen Nachfolger.
  - 29 Zum Beispiel: Jack Goody/Ian Watt: Konsequenzen der Literalität, in: Jack Goody (Hg): Literalität in traditionellen Gesellschaften, Frankfurt/M. 1981, S. 45–104.
  - 30 Rouget: La musique et la transe (Anm. 13), S. 82.
  - 31 Göttert: Geschichte der Stimme (Anm. 27), S. 35.

**Cornelia Epping-Jäger**

**»EINE EINZIGE JUBELNDE STIMME«.**

**ZUR ETABLIERUNG DES DISPOSITIVS LAUT/SPRECHER IN DER POLITISCHEN  
KOMMUNIKATION DES NATIONALSOZIALISMUS**

Gegenstand des folgenden Beitrages ist die These, dass in Deutschland während der Jahre 1932–1936 ein von der NSDAP konzipierter medialer Großversuch stattfand, in dessen Verlauf sich aus den vernetzten Dynamiken von technologischer Forschung und politischem System das Dispositiv Laut/Sprecher<sup>1</sup> als eine Diskursapparatur etabliert, die zentriert ist um die massenkommunikative Distribution der ›Stimme des Führers‹ und ihrer personalen Substitute. Die Fortschritte, die die technologische Forschung sowohl auf dem Gebiet der Massenbeschallung auf freiem Feld als auch durch das Kurzschließen der Massenbeschallung mit dem Verbreitungsmedium Rundfunk erzielte, erlaubten dem NS-System eine mediale Konsolidierung als Hörergemeinschaft: Das aus dem Großexperiment hervorgehende Dispositiv ›Laut/Sprecher‹ organisiert die Verinnerlichung der Macht als akustische Erlebnisform der ›Volksgemeinschaft‹, die

**Abb. 1**  
1932



sich in der synchronen Ubiquität der ›Führerstimme‹ und ihrer Gemeinschaftsresonanz medial konstituiert.

Die zentrale Phase dieses Experiments lässt sich zeitlich situieren zwischen der Wahlkampagne »Hitler über Deutschland« des Jahres 1932, der auf dem Tempelhofer Feld abgehaltenen »Maifeier« des Jahres 1933 und den deutschlandweit übertragenen »Wahlreden« Adolf Hitlers im März 1936 aus der Krupp-Halle in Essen und der Schlageter-Halle in Königsberg. Zwischen diesen massenkommunikativen Großereignissen liegen vier Jahre der Ausdifferenzierung des medialen Dispositivs, das ein wesentliches – vielleicht konstitutives – Moment der Kommunikationskultur des NS-Systems darstellte.

### 1. DIE NSDAP ALS REDNERPARTEI: DIE HERRSCHAFT DER STIMME

Nach ihrem in Anschluss an den Marsch auf die Feldherrenhalle 1924 ausgesprochenen Verbot wurde die NSDAP im Februar 1925 neu gegründet, und sie konstituierte sich explizit als ›Rednerpartei‹.<sup>2</sup> Im selben Jahr berichtete eine Technikzeitschrift erstmals über den Einsatz von Großlautsprechern bei einer amerikanischen Wahlveranstaltung.<sup>3</sup> Deutsche Firmen, die zu diesem Zeitpunkt bereits eigene elektro-akustische Technologien zur ›Massenbeschallung‹ erprobten, griffen die amerikanische Erfindung auf und entwickelten sie weiter.<sup>4</sup> In Kirchen<sup>5</sup> und Domen,<sup>6</sup> aber auch in großen Fest- und Industriehallen<sup>7</sup> und auf Außengeländen<sup>8</sup> testete man in Versuchsserien aus, wie – so formuliert Hans Gerdien, der Telefunken-Chefkonstrukteur, bereits 1926 den für den Einsatz des Laut/Sprechers stabilen Topos – »die riesigen Dimensionen eines Saales vollkommen durch eine einzige menschliche Stimme zu beherrschen«<sup>9</sup> seien. Zu Beginn der 30er Jahre sind der Sportpalast,<sup>10</sup> der Reichstag,<sup>11</sup> Zirkus Busch und Krone und ein Großteil jener Orte, an denen Nationalsozialisten ihre Rede-Veranstaltungen abhalten, mit den entsprechenden Übertragungsanlagen ausgestattet.<sup>12</sup> In dieser Phase scheint es zu einer Zusammenarbeit zwischen Technikfirmen und NSDAP gekommen zu sein, die als Rednerpartei früh die medialen Chancen der Lautsprecher-Technologie erkannte.<sup>13</sup> Vom Hauptredner der Partei, Adolf Hitler, ist bekannt, dass er die akustischen Besonderheiten der Auftrittsorte vor seinen Reden austestete, um sie in seinen Reden operational zu nutzen. »Heute« – so notierte Joseph Goebbels etwa 1928 in sein Tagebuch – »probt der Chef im Sportpalast die Lautsprecher aus.«<sup>14</sup>

Neben fest installierten Anlagen griff die NSDAP intensiver als andere Parteien der Weimarer Republik auf transportable Lautsprecher-Anlagen zurück, die

eine rasche und flexible technische Ausdehnung von Beschallungsräumen erlauben. Während SPD und KP stärker mit Megaphonen – aus der Masse heraus also – agierten, setzte die NSDAP den Lautsprecher und die durch ihn elektro-akustisch gewandelte Stimme des Redners als unidirektionales, akustisches Herrschaftsinstrument ein, dessen unter Umständen noch mangelnde ›Durchschlagskraft‹ durch die körperliche Überzeugungsarbeit eigens gegründeter SA-Verbände kompensiert wurde.<sup>15</sup>

Die Verbreitung und technische Fortentwicklung der Lautsprechertechnologie führte im Horizont ihrer immer intensiveren politischen Nutzung insbesondere durch die NSDAP im Jahr 1932 zu einem absolut neuen medialen Szenario des Laut/Sprecher-Einsatzes: zu der Wahlkampagne »Hitler über Deutschland«. <sup>16</sup> Vom zweiten Wahlgang der Reichspräsidentenwahl an, über die Landtagswahlkämpfe bis zur Reichstagswahl im Oktober absolvierte Hitler insgesamt vier Deutschlandflüge und sprach auf 200 Kundgebungen.<sup>17</sup>

Die als Nicht-Regierungspartei von der Nutzung des Rundfunks ausgeschlossene NSDAP kompensierte diesen Ausschluss also durch ein alternatives massenkommunikatives Programm und setzte zum medialen Angriff an. Joseph Goebbels notiert:

Eine entscheidende Neuerung: der Führer wird seinen nächsten Wahlkampf mit dem Flugzeug bestreiten und täglich drei- bis viermal nach Möglichkeit auf offenen Plätzen oder in Sportarenen sprechen. Dann kann er trotz der Kürze der noch zur Verfügung stehenden Zeit etwa ein und eine halbe Million Menschen erfassen.<sup>18</sup>

Eine gecharterte D 1720 fliegt den ›Führer‹ in die verschiedensten Teile der Republik. Am 6. April 1932 etwa werden die Stationen Königsberg-Berlin-Würzburg-Nürnberg-Regensburg angesteuert; am 7. April 1932 folgen Nürnberg-Frankfurt-Darmstadt-Ludwigshafen; am 18. April geht es von München über Beuthen und Görlitz nach Breslau. Folgt man den Zahlen des ›Führerbegleiters‹, dann sprach Adolf Hitler vor 250 000 Menschen in Schlesien, vor 80 000 in Dresden, 150 000 hörten ihn in Königsberg, 200 000 in Berlin. Insgesamt errechnete man eine Präsenzöffentlichkeit von 10 Millionen Menschen.<sup>19</sup>

Hier ist nicht der Ort, die Medienkampagne in allen Einzelheiten nachzuzeichnen, daher seien nur die wichtigsten Konstituenten der medialen Inszenierung erwähnt. Zentral ist, dass die Wahlkämpfe eine akustische und zeitliche Neustrukturierung des öffentlichen Raumes vorstellten: Schon Tage vor den eigentlichen Wahlkampfenden bereiten SA- und Propagandatrupps mit Sprech-

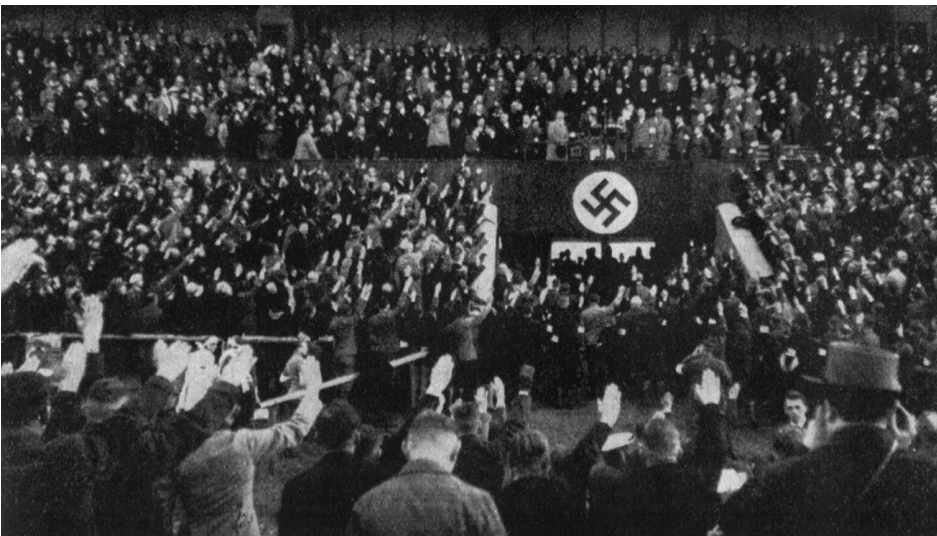


Abb. 2

»Eine Viertelmillion Schlesier stand im Banne des Führers«

chören sowie parteieigenen Lautsprecherlastwagen den Advent des ›Führers‹ vor. Herkömmliche Wahlveranstaltungen beschränkten sich auf kurz an- und wieder abreisende Redner, die entweder inmitten der Menge oder von markierten Redepositionen aus sprachen. Die NSDAP dagegen legt zunächst einen mehrtägigen Stimm- und Klangteppich über die für die Kampagne ausgewählten Orte und inszeniert einen bis dahin noch nicht erfahrenen, zeitlich und topologisch extrem

ausgeweiteten Hörraum, der so als Resonanzraum für den erwarteten Redner vorbereitet werden soll. Die Zuhörer werden straff organisiert und auf das akustische Zentrum hin ausgerichtet: Gemietete Lastwagen, Sonderzüge, Bus- und Radkolonnen sowie Sternmärsche führen sie aus allen, zum Teil erheblich auseinander liegenden Regionen zusammen. »Unabsehbare Massen« – so berichtet der Kommentator Berchtold etwa aus Görlitz – »stehen in den Straßen, auf den Dächern, hängen in den Ästen der Bäume.«<sup>20</sup> Die gezielt mit dem Prinzip ›Verspätung‹ kalkulierenden Reden Adolf Hitlers werden – vor allem wenn es um große Auditorien geht – in die Abend- und Nachstunden gelegt und mittels Lautsprechern nicht nur in die Hallen und Zelte, sondern auch in den öffentlichen Raum übertragen:

Die Frankenhalle war seit Tagen ausverkauft, so mussten 14 Lautsprecher die Rede des Führers zu den Zehntausenden tragen, die sich im Freien zusammenballten. [...] Satz um Satz dröhnte in die kalte Nacht, senkte sich in die Herzen.<sup>21</sup>

Und:

Eine einzige Masse im Dunkel der Nacht, ein einiges Volk, Deutsche, nichts als Deutsche. Im hellen Lichte über ihnen steht der Führer und durch die Stille, die jetzt über allen feierlich liegt, erhebt sich seine klare Stimme zum Appell an seine Schlesier. Er spricht bis das Deutschlandlied hervorquillt aus übervollem Herzen und einem Orkan gleich braust über Görlitz ...<sup>22</sup>

Breslau. Hitler spricht auf der offenen Radrennbahn, »Zehntausende hören ihn und auch noch so viele in den Straßen; wo gestern der Ministerpräsident Braun noch vom Manuskript ablas, predigt Adolf Hitler heute die nationalsozialistische Volksgemeinschaft.«<sup>23</sup>

Die Inszenierung operierte also angesichts der jeweils gewaltigen und organisierten Nachfrage mit dem Prinzip der Knappheit der Ressource ›Führerstimme‹, einer Knappheit, die jedoch durch die Ausdehnung des Übertragungsraumes technologisch kompensiert wurde. Die bei den generalstabsmäßig organisierten Planungen gemachten Erfahrungen – insbesondere auch mit auftretenden Störungen – werden systematisch ausgewertet und in einem »Organisationsgedächtnis« archiviert, das Musterchoreographien für spätere Massenveranstaltungen bereithält.<sup>24</sup>



Als erstes Resümee lässt sich festhalten: Die Rundflüge zielten auf eine ubiquitäre Verbreitung der Stimme des ›Führers der Bewegung‹. Und obgleich das Konzept sicher zunächst nur auf eine mediale Substitution des nicht zur Verfügung stehenden Rundfunks und seiner Verbreitungsqualität zielte, zeigten sich andererseits hier bereits die *Konturen eines neuen medialen Dispositivs*, das sich in den folgenden Jahren ausdifferenzieren sollte, wobei es nachhaltig durch eine nie restlos ausgeräumte Distanz zum Rundfunk geprägt blieb. Das Radio nämlich war für die Zwecke der NSDAP, die ihrem Selbstverständnis nach ja eine sozialrevolutionäre Massenbewegung darstellte, schon deshalb ein problematisches ›Massenmedium‹, weil seine differenzierten Adressenordnungen für das Entdifferenzierungsbedürfnis der Partei, d. h. für das Ziel einer medialen Installation von Gemeinschaftserlebnissen, wenig geeignet erschien.

Kann man sich der sichtbaren Propaganda zur Not entziehen, so ist dies bei der hörbaren Propaganda schon schwieriger. Eine Ausnahme bildet das Radio, das einem gestattet, Ministerreden abzuschalten. Aber schon der Lautsprecherwagen verkündet seine Botschaft, ohne dass wir das Ohr dagegen verschließen können.<sup>25</sup>

Der NSDAP ging es als ›Massenbewegung‹ zunächst nicht um die anonym bleibenden, kollektiv nur schwer zu adressierenden und in ihrem Rezeptionsverhalten kaum zu kontrollierenden Rundfunkhörer, sondern vielmehr um die im öffentlichen Raum versammelten massenhaft je Einzelnen, die gleichsam einer Logik direkter Interaktion verfügbar gemacht werden sollten.

Die durch den Lautsprecher organisierte Ausrichtung der Massenkommunikation auf die Kontrolle und Resonanz von Auditorien reflektiert die Zentralstellung der ›Führerstimme‹ und ihres kollektiven Resonanzraumes: Sie erlaubte, im Gegensatz zu der Standardnutzung des Rundfunks, sowohl die mediale Konstruktion des Gemeinschaftserlebnisses als auch seine Fokussierung auf die ›bewegende‹ Stimme des ›Führers der Bewegung‹. Zudem ließ sich so, über die Simulation einer »Kommunikation im Raume wechselseitiger Wahrnehmung«,<sup>26</sup> die durch die Lautsprecher-Technologie körperdistanzierte Stimme als ›embodied voice‹ inszenieren.

Musste die NSDAP in ihrer Frühphase noch in Arbeiterkleidung getarnte Parteigenossen Schlägereien anzetteln lassen, damit die Zeitungen, wenn schon nicht über den Redner Hitler, doch über die um ihn herum stattfindenden Schlägereien berichten konnten, so war dies in den Zeiten des spektakulären Lautsprecherwahlkampfes nicht mehr nötig, in dem die Auditorien offenbar die

neuen Stimm-Performances als heute kaum noch nachvollziehbare akustische Faszination erlebten.

Das Wissen um die Innovativität des sich herausbildenden neuen Dispositivs – das ist auffällig am Agieren der NSDAP – geht zu Beginn der dreißiger Jahre als operationalisierbare Ressource in die mediale Strategie der Partei ein, die in einem ausdifferenzierten *Medienverbund* – mit Sonder-Publikationen, mit auf Schallplatten aufgenommenen Ausschnitten aus Massenreden und entsprechenden Filmaufnahmen<sup>27</sup> – die Lautsprecher-Wahlkämpfe transkribiert, um sie so dem einzelnen Hörer oder Versammlungsteilnehmer als *Gesamtinszenierung* erneut zugänglich zu machen: Auf diese Weise sieht sich – so kann man mit Walter Benjamin formulieren – »in den großen Festzügen, in den Monstreveranstaltungen, in den Massenveranstaltungen, [...] die heute sämtlich der Aufnahmeapparaturlatur zugeführt werden, [...] die Masse selbst ins Gesicht.«<sup>28</sup>

## 2. DER TAG DER ARBEIT ODER DER AKTUALISMUS DES DABEISEINS

Am 1.5.1932 notierte Joseph Goebbels in sein Tagebuch:

Die Roten demonstrieren zum 1. Mai. Aber das zieht nicht mehr. Es wird wohl auch das letztemal sein, dass sie ein Wörtchen mitzureden haben. Im nächsten Jahr werden wir ihnen wahrscheinlich zeigen, wie so etwas gemacht wird.<sup>29</sup>

Ein Jahr später, am 26.4. 1933, lautet der Eintrag: »Der 1. Mai steht auf dem Papier fertig.«<sup>30</sup> Zwei Tage später präzisiert er:

Ich arbeite tagelang an der technischen Durchführung des 1. Mai. Es soll ein Meisterwerk der Organisation und Massendemonstration werden. [...] In Tempelhof sind gigantische Anlagen aufgebaut. Der 1. Mai wird ein Massenereignis wie es die Welt noch nicht gesehen hat. [...] Jetzt kommt uns unsere vielfältige Erfahrung auf dem Gebiet der Massenföhrung zu nutze. Keine andere Bewegung als die unsere, die Massen zu dirigieren versteht, wäre zu dieser Demonstration fähig.<sup>31</sup>

Anhand der nationalsozialistischen Choreographie der Maifeier lässt sich in der Tat exemplarisch beobachten, wie tradierte politische Inszenierungsformen und deren Semantik im Zuge einer *medialen Re-Inszenierung* als vertrautes symboli-

sches Inventar zugleich beliehen und enteignet werden. Denn: Der 1. Mai als »Tag der Arbeiterbewegung« wird am 1. Mai 1933 als »Tag der Arbeit« zum nationalen Feiertag transformiert.<sup>32</sup> Das Enteignungsverfahren vollzieht sich als massenkommunikative Inszenierung von »Volksgemeinschaft«,<sup>33</sup> wobei diese semantische Überschreibung ihre Evidenz und Plausibilität aus der medialen Konstruktion gigantischer Erlebnis-Gemeinschaften bezieht: »Über den Trümmern des zusammengebrochenen liberal-kapitalistischen Staates« – so Goebbels – »erhebt sich der Gedanke der Volksgemeinschaft«<sup>34</sup> – ein Gedanke, der in den diskursiven Szenarien des Dispositivs Laut/Sprecher konstituiert wird. Das mediale Konzept zielt auf die Konstruktion eines neuen, kollektiven Erlebnisraumes, eines Erlebnisraumes, der in einer oszillierenden Choreographie von Stillstand und Bewegung einen »Aktualismus des Dabeiseins« inszeniert, wie ein zeitgenössischer Beobachter konstatiert.<sup>35</sup> Mit der Maifeier von 1933 entsteht so ein medialer Prototyp, der fortan als Ritualform lokal organisiert und über die Massenmedien distribuiert und repliziert werden kann.

Die organisatorischen, diskursiven und ästhetischen Konstituenten des Szenarios verdienen einen kurzen Blick.<sup>36</sup> Der »Tag der Arbeit« beginnt um Mitternacht; im ganzen Reich setzen sich organisierte Kolonnen Richtung Berlin in Bewegung, in der Stadt übernehmen Parolen skandierende Propagandatrupps das Wecken. An zehn Aufmarschplätzen in den Berliner Stadtteilen sammeln sich Schulkinder und Jugendliche, um in kilometerlangen Zügen zum Lustgarten zu marschieren. Dort erwartet die 200 000 Jugendlichen ein mit 1200 Sängern monumental besetzter Chor, der die Hymne »Deutschland, du mein Vaterland« singt, anschließend redet Goebbels zu den als »Avantgarde der deutschen Revolution« bezeichneten Kindern und Jugendlichen. Aus den aufgestellten Lautsprechern ertönt Marschmusik, dröhnend kreisen Flugzeugstaffeln über der Stadt. Hitler und Hindenburg treffen im Lustgarten ein, reden.

Gegen 11 Uhr setzt sich das werktätige Berlin,<sup>37</sup> von dem ein performativer Nachweis der Dankbarkeit und des Willens zur Mitarbeit eingefordert wird, in Bewegung. Sternförmig marschieren zehn Marschsäulen in Richtung Tempelhofer Feld. Durchschnittlich waren ca. 60 000 Menschen pro Marschsäule vorgesehen, realiter wurden es über 100 000. Die gestaffelten Abmarschzeiten sind so berechnet, dass alle Teilnehmer um 19.30 Uhr das Feld erreicht haben.<sup>38</sup> Unterwegs marschieren sie an den weitgeöffneten Fenstern und Türen der Wohnungen und Geschäfte vorbei, in denen die Rundfunkempfänger stehen und in voller Lautstärke jene Reportagen übertragen, die darüber berichten, dass die Kolonnen zum Feld marschieren. Simultaninszenierung: Die Massen hören sich beim Marschieren zu.

Auf dem Feld rücken die Kolonnen unter Lautsprecheranweisungen in die für sie ausgezeichneten Felder ein; Berufsgruppen in berufsspezifischer Arbeitskleidung simulieren eine ständische Schichtengesellschaft, deren soziale Binnenstruktur allerdings im volksgemeinschaftlichen Kollektiverlebnis aufgehoben ist. Aus den über dem Feld kreisenden Flugzeugen wird photographiert, in welcher Stärke die Betriebskolonnen antreten: Für schwach besetzte Betriebsstandorte werden im Anschluß an die Maifereier Erklärungen verlangt. Weitere Flugzeuge: Sie bringen Arbeiterdelegationen aus dem Reich und aus Wien nach Tempelhof.<sup>39</sup>

Gegen 18 Uhr am Nachmittag sind 1,5 Millionen Menschen auf dem Feld, eine weitere Million in den Straßen. Unterhalten werden sie durch Militärmärsche, die aus Großlautsprechern ertönen, sowie durch Reportagen, die aus dem inzwischen über Berlin eingetroffenen Luftschiff »Graf Zeppelin« direkt, über außen befestigte Lautsprecher, auf die Straßen und Plätze übertragen werden: »Aus den Außenbezirken fluten 1000 Ströme zu einem Ziel, dem Tempelhofer Feld, 1000 Ströme zum deutschen Herz und zum deutschen Volk, und alle münden dort unten in ihre große Gemeinschaft.«<sup>40</sup>

Die Stimme des Zeppelin-Reporters Alfred Ingemar Brand, die diese Sätze spricht, wird gleichzeitig auf Platte geschnitten und im Laufe des weiteren Tages vom Rundfunk deutschland- und – den »Anschluss« auf der Ebene des Rundfunks vorwegnehmend – österreichweit gesendet. Was den einzelnen Teilnehmern nicht sichtbar war, von ihnen nur fragmentiert erfasst werden konnte – die Monumentalität des Massenaufmarsches und das Bild der Massenformierung – wird ihnen hier von oben zugesprochen und vor Augen gestellt.

Aber nicht alle Feldbesucher sind zu Kolonnen formiert, nicht alle haben Tribünenkarten. Es gibt freie Bereiche – »Rasenkarten« zu zwei Reichsmark pro Stück. Der »Berliner Beobachter« bewirbt den Kartenverkauf:

Jeder Besucher der Rasenplätze hat die Möglichkeit, den Reichskanzler ständig persönlich sehen zu können, da der Reichskanzler die Front der Rasenplätze bei seiner Ankunft und bei seiner Wegfahrt abfahren wird, und da die Redekanzel, von der der Führer seine bedeutsame Ansprache halten wird, von allen Seiten aus gut sichtbar ist.<sup>41</sup>

Neben den Rasenplätzen gab es auch kostenlose Zuschauerplätze: Übergangsbereiche, in denen die alten Vergnügen der Arbeiterbewegung explizit erlaubt sind. Hier werden Picknicks gehalten, die Kinder spielen unter Beaufsichtigung, Gymnastikvereine treten ebenso auf wie Schalmeienzüge, Würstchenverkäufer und Bierkutscher bieten ihre Waren feil<sup>42</sup> – bewusst inszenierte Momente einer Popu-

larkultur. Allerdings: Wer immer das Feld betrat, der trat ein in eine Zone, die von der SA kontrolliert wurde und als Aufmarschkarree von einem Außen, das sich der Regelmäßigkeit militärischer Ordnungsfiguren entzog, deutlich geschieden war.<sup>43</sup>

Um 20 Uhr dann: Aufmarsch der 1000 Fahnen auf die Rednertribüne, Reichswehr und Schutzpolizei rücken in das Viereck vor der Tribüne ein. Zugleich flammen die Scheinwerfer auf, sieben Kapellen intonieren »Der Gott, der Eisen wachsen ließ, der wollte keine Knechte«, die »Volksgemeinschaft« singt mit. Die Zeichen für die Ankunft des »Führers« nehmen zu. Stürmische Heilrufe, SA und SS rücken ein, positionieren sich gegenüber der Rednertribüne. Adolf Hitler nähert sich dem Feld und betritt die 10 Meter hohe Bühne, von der die Redekanzel noch einmal deutlich abgesetzt ist.

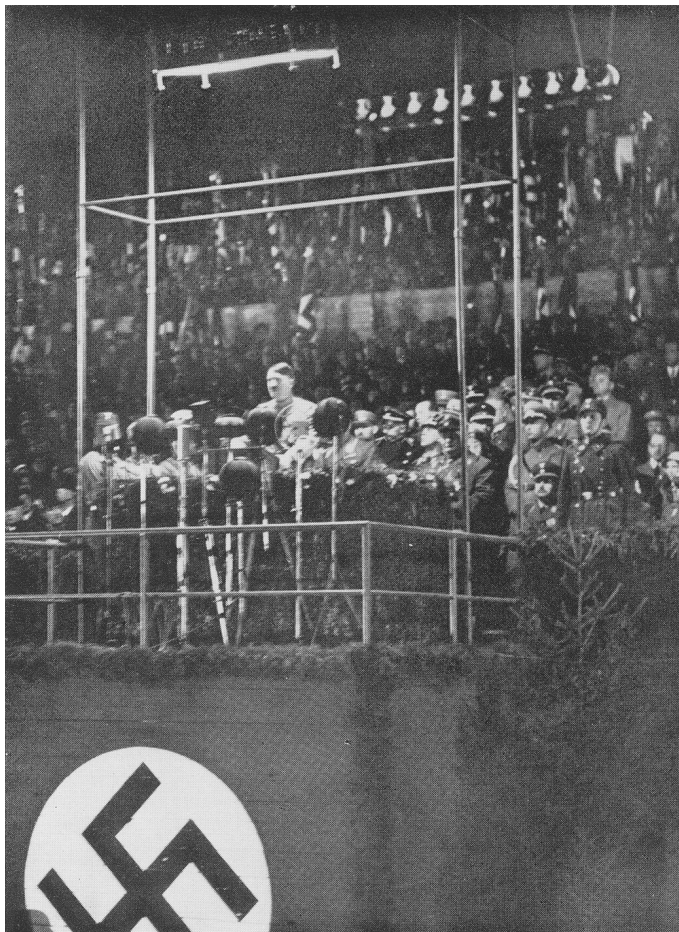


Abb. 3

»Adolf Hitler spricht«

Der Blick hinter die Kulisse: Die technische Logistik, die die Grundlage dieser Massenfeier bildete, stellte hohe Anforderungen, und es sind vor allem die Lautsprecheranlagen, die die für die Feier konstitutive zentrale Stimm-Inszenierung aller erst ermöglichten. Denn nur das Netz der funktionierenden elektrotechnischen Anlagen machte es möglich, »die Millionenmassen der Besucher zu unmittelbaren Teilnehmern an der Feierlichkeit zu machen.«<sup>44</sup>

Hier galt es einen Raum von fast 500 000 Quadratmetern mit einer einzigen Stimme zu beherrschen, den fünfviertel Millionen Zuhörern jedes einzelne Wort verständlich zu machen, jedes kleine und kleinste Beben dieser Stimme zu übermitteln. Die Techniker haben diese Aufgabe gelöst! Die Stimme Adolf Hitlers wurde auf das 50 000fache verstärkt! [...] In wenigen Tagen und Nächten ist die größte Lautübertragungs-Anlage der Welt (einschließlich Amerikas) errichtet worden.<sup>45</sup>

In den Tagen zuvor waren viele Kilometer Kabelleitungen verlegt worden.

Dann baute man ein Netz von über 100 Großlautsprechern, so dass keiner der Zuhörer weiter als 75 Meter von einem Lautsprecher entfernt war. Hätte sich irgendwo eine Lücke, eine »stumme Zone« ergeben, dann wäre rasch ein Großlautsprecherwagen herbeigeschafft und in Betrieb genommen worden ...<sup>46</sup>

Die riesige Verstärkeranlage war in einem Zelt untergebracht, und vier Umform-Maschinen von zusammen 8 Kilowatt Leistung lieferten den Gleichstrom für die Feld-Erregung der dynamischen Lautsprecher.

Zur Besprechung der Anlage dienten zehn Mikrophone modernster Konstruktion, die sich auf der Redner-Tribüne, Kommandoturm, Musik- und Zuschauertribünen verteilten. Außerdem übertrug eine direkte Leitung zum Funkhaus einen Teil des Rundfunk-Programms auf das Tempelhofer Feld.<sup>47</sup>

Die Mikrophone wurden von einem Regiestand – der »Nervenzentrale«, dem »technischen Gehirn«<sup>48</sup> – aus eingesetzt, der auch über einen Schneide-Tisch zur sofortigen Herstellung von Schallplatten verfügte, so dass die Reden aufgenommen und unmittelbar zum Funkhaus transportiert werden konnten. Kontrolliert wurden von hier aus auch die an den äußeren Längsseiten des Feldes aufgestellten

Lautsprecher mit Richtwirkung, »die den Schall an die Ohren der Hunderttausende weitergaben, die außerhalb des Feldes sich gelagert hatten.«<sup>49</sup> Hans Wendt schreibt:

Es ist als müßten sich die Seelen dieser 1,5 Millionen zu einer gewaltigen Erscheinung vereinen, einem Aufschrei, einer Bewegung, einer alles überragenden Gestalt. Aber sichtbar wird nur der ordnende Geist, der über diese Heerschar waltet, die alles übertrifft, was die Welt je an Massenaufmärschen gesehen hat.<sup>50</sup>

In Italien, so berichtete selbst die Nordschleswiger Tageszeitung stolz, werden die Reden des *Duce* in den Städten noch verlesen, in Deutschland dagegen durch Lautsprecher und Rundfunk unmittelbar der Bevölkerung übermittelt, so dass jedem die Möglichkeit gegeben wird, »selbst die Stimme des Führers zu hören.«<sup>51</sup>

In der Tat, am 1. Mai 33 gehörte auch der Rundfunk zum Medienverbund des Laut/Sprecher-Dispositivs. Erstmals wurde ein achtzehnstündiges Programm zusammengestellt, zusammengesetzt aus Redeübertragungen, Hörspielen, Hörberichten und Reportagen über die Maifeiern in anderen Städten, das den Eindruck vermittelte, die ganze Nation gliedere sich freudig in die ›Volksgemeinschaft‹ ein. Dieses Radioprogramm stellte im *Medien-Großversuch* einen *Subtest* dar: Ausgetestet wurde, welche rundfunk-spezifischen Sendeformen – etwa Konzert, Reportage, Hörspiel, Funkbericht – als Unterhaltungselemente des Gesamtevents geeignet waren. Ausgetestet wurde insbesondere die Wirkungskraft der ›disembodied voices‹ des Rundfunks, und aufschlussreich ist, dass sich gerade die im Zentrum des *Hörspiels* stehende ›nackte Stimme‹, die Stimme also als – so Richard Kolb – »körperlose Wesenheit«, als dispositiv-untauglich erwies.<sup>52</sup> Erfolgreich dagegen waren die aus den Massenveranstaltungen berichtenden *Reportagen*, deren Re-Entry in die Szenarien, die sie kommentierten, besondere Agitationswirkung entfalten konnte.

Auch für Adolf Hitler selbst hatte sich die mediale Nutzbarkeit des Rundfunks auf die Funktion des Verbreitungsmediums reduziert: Zwei Versuche, ›einsam‹ vor dem Rundfunkmikrofon zu sprechen, waren Anfang Februar 1933 gescheitert.<sup>53</sup> Von nun an ist jede gesendete Rede »verkörperte Rede«, ausgeschnitten aus Massenversammlungen und imprägniert mit akklamierender Massenresonanz. Die Stimme des ›Führers‹ wird nie mehr direkt und unmittelbar aus dem Studio an das anonyme Kollektiv der Radiohörer adressiert, sondern immer nur indirekt – als Exempel einer zuvor gelungenen Adressierung im Rahmen einer Massenveranstaltung.

Die Faszination der Lautsprechernutzung lag dabei vor allem in dem Umstand, dass die durch elektro-akustische Wandlung vom Leib des Redners getrennte Stimme in den performativen Szenarien der Massenrede die Aufgabe übernehmen konnte, Massenkommunikation als leibanesende Kommunikation in Szene zu setzen. In seiner Mairede formuliert Adolf Hitler diesen Zusammenhang so:

Wir haben die erste Aufgabe der Wiedergenesung unseres Volkes erkannt: das deutsche Volk muß sich wieder gegenseitig kennenlernen. Die von Standesdünkel und Klassenwahnsinn Befallenen, die sich nicht mehr verstehen lernten, die müssen den Weg zueinander finden.<sup>54</sup>

Das Dispositiv Laut/Sprecher schließt durch seine diskursive Ordnung jenen Kontrollverlust aus, der in das Medium Rundfunk immer dann eingeschrieben ist, wenn die Rezeption nicht als Gemeinschaftsempfang organisiert werden kann. Die Steuerbarkeit des Auditoriums ist hier vollständig. »Selbst der Beifall, der immer wieder aufdröhnt«, hört sich in den Ohren Hans Wendts an, »als wollten die Rufe nicht den Redner unterbrechen, als wagten sie es nicht. Ein Strom von Vertrauen flutet empork zu dem Kanzler des Volkes.«<sup>55</sup>

Blickt man auf die Rednerseite, auf Adolf Hitler, dann wird deutlich, dass sich eine solche Disziplinierung nur vor dem Hintergrund der technisch sichergestellten Stimmübertragung für den Gesamttraum,<sup>56</sup> aus dem Spannungsfeld von feed-back-Verarbeitung und Selbstinduktion des Redners sowie aus der Logik vorgetäuschter direkter Interaktion entfalten konnte. Das Dispositiv Laut/Sprecher ermöglicht, dass der Redner seine eigene Stimme durch ein technisch verzögertes Feedback in eins mit einer Imprägnierung durch die Massenresonanz wieder vernimmt. Da der Lautsprecher die Stimmsignale immer mit einer gewissen zeitlichen Verzögerung überträgt, wird die Selbstwahrnehmung des Redners durch die technisch vermittelte, zeitlich verschobene äußere Resonanz – durch die Stimme, die alle gehört haben – überschrieben. Die durch die Mikrofon-Lautsprecher-Konfiguration medial vom Leib getrennte Stimme schreibt sich kollektiv ratifiziert erneut in den Körper des Redners ein. Sie wird durch die Mikrofon/Lautsprecher-Konfiguration auratisch inszeniert in einem paradoxen zugleich von Nähe und Ferne. Die Stimme ist bei sich und doch technisch distanziert, sie ist Massenkommunikation und doch auch intime Kommunikation im ›Raum wechselseitiger Wahrnehmung‹. Es ist diese Synthese aus techno-akustischem Format mit neuen Formen performativer Inszenierung der Stimme, die die Wirkungsmächtigkeit des Dispositivs Laut/Sprecher hier verfestigt.





Fot. N. B. S.

In der Mittagssonne: Das Tempelhofer Feld beginnt sich zu füllen

Abb. 4

Tempelhofer Feld 1933

Die scheinbar authentische Stimme des Redners kann also als mediale Verfertigung der Rednerstimme in der Massenversammlung analysiert werden. In gewissem Sinne lässt sich daher von einer *Medienfiktion* sprechen, denn die Laut/Sprecher-Stimme setzt die Fiktion einer ›Volksgemeinschaft‹ voraus, die durch ihre Erzeugung, Übertragung und kollektive Ratifizierung aller erst hervorgebracht wird.

1933 bedeutete das auch, dass das noch nicht verfestigte Dispositiv Laut/Sprecher gegen den immer noch massiv vorhandenen Eigensinn der Rezipienten mit den Mitteln von Zwang und Gewalt verankert werden musste: Von der einen Seite des Auditoriums dröhnt die ›Führerstimme‹, auf der gegenüberliegenden Seite haben SA, SS und Wehrmacht Aufstellung bezogen. Die nicht bloß symbolische Aussage dieser Konfiguration zeigte sich am 2. Mai: Am Morgen dieses Tages brechen um 9 Uhr in allen Teilen des Reiches nationalsozialistische Einsatz-

kommandos auf, um die Häuser und Einrichtungen der freien Gewerkschaften mit Gewalt zu besetzen. Innerhalb einer Stunde ist die einstmals stärkste Arbeiterbewegung in einem nicht-sozialistischen Land, ohne dass es zu nennenswertem Widerstand gekommen wäre, zerschlagen. »Den ersten Mai« – notierte Joseph Goebbels am 17. April 1933 in sein Tagebuch – »werden wir zu einer grandiosen Demonstration deutschen Volkswillens gestalten. Am 2. Mai werden dann die Gewerkschaftshäuser besetzt. [...] Es wird vielleicht ein paar Tage Krach geben, aber dann gehören sie uns.«<sup>57</sup>

### 3. DIE FRIEDENSREDE AUS DER WAFFENSCHMIEDE

Am 7. März 1936 wurde der Reichstag in Berlin einberufen, und Adolf Hitler verkündete die Re-Militarisierung des Rheinlands, die Kündigung des Locarno-Vertrages, die Auflösung des Reichstags und Neuwahlen für den 29. März 1936.<sup>58</sup>

In der Wahlkampagne, die nun folgte – und die am Beispiel Königsbergs betrachtet werden soll –, zeigt sich das Dispositiv Laut/Sprecher in seinen konstitutiven Elementen in etablierter Form, die während der nächsten neun Jahre relativ stabil bleiben sollte. Der Laut/Sprecher wird hier zu einem *medientechnologisch-ausdifferenzierten Stimm-Dispositiv*, das die ›Stimme des Führers‹, ihre Substitute und Verlängerungen in einer Verknüpfung von Diskursen und Institutionen erfasst.

Blickt man auf die Diskurse, dann wird deutlich, dass die nicht-distribuierte Massenveranstaltung und die in sie eingebettete Massenrede nach wie vor die grundlegende NS-Diskursform darstellte. Am 11. März 1936 schreibt die Preußische Zeitung unter der Überschrift: »Wahlkampf durch Goebbels eröffnet. Großkundgebung in der Deutschlandhalle«:

Die riesige Deutschlandhalle in Berlin ist fertig geworden, doch die beiden riesigen Räume reichten nicht aus, und an nicht weniger als 230 weiteren Versammlungsorten hörten die Berliner die Rede ihres Gauleiters Dr. Goebbels. Die Deutschlandhalle mußte schon eine Stunde vor Beginn gesperrt werden. Viele Tausend die keine Karten mehr bekommen hatten, blieben auf den Vorplätzen stehen und hörten durch Lautsprecher, die schnell auf Kraftwagen aufgestellt wurden, die Versammlung mit.<sup>59</sup>

Auf der nächsten Zeitungsseite findet sich die Beschreibung des entsprechenden Königsberger Szenarios, welches das in den Maifeiern ausgetestete Bewegungsschema lokal appliziert:

Parteigenosse Dr. Ley eröffnet den Wahlkampf in Ostpreußen und sprach zu Tausenden im Schlageterhaus:

Schon lange vor Beginn traten vor den großen Betriebsstätten und auf bestimmten Sammelplätzen Männer und Frauen zusammen, um sich in geschlossenem Zuge zur Kundgebungshalle zu begeben. Musik und Lieder von Kampf und Arbeit und Freiheit erfüllten die Straßen, und Fahnen wehten über unabsehbaren Menschenmengen. Zug um Zug rückte in die lichtdurchflutete Halle des Schlageterhauses ein. [...] In kaum einer Stunde war der riesige Raum überfüllt und mußte polizeilich gesperrt werden. Tausende fanden keinen Einlaß mehr und konnten sich nur mehr vor der Halle und auf den Straßen an den Lautsprechern der Übertragungsanlage verteilen, um an der Kundgebung teilzuhaben. Wir gewannen die Arbeiter der Volksgemeinschaft wieder, einer Volksgemeinschaft, die nicht Angelegenheit des Verstandes, sondern des Herzens ist.<sup>60</sup>

Eine Bewegung ist eine Bewegung ist eine Bewegung, das Schema ist bekannt: Sammelplätze, sternförmig aufgebaute Marschkolonnen, Gesang, das Operieren mit der Knappheit der Ressource ›Führerstimme‹ in überfüllten Räumen, die über Lautsprecher nach außen übertragenen Reden, das gemeinsame Anhören dieser Reden durch die an festgelegten lokalen Versammlungspunkten stillgestellte und kontrollierte ›Gemeinschaft‹.<sup>61</sup>

Die Differenz zu 1933 aber besteht darin, dass die Distribution der Massenrede spätestens 1936 durch den Rundfunk erweitert wird, der nun vollständig in den NS-Medienverbund integriert war. In Bezug auf die Übertragung der ›Führer‹-Rede besteht die Pointe dieser Übertragungen darin, dass sich die Rundfunkhörer als eine natürliche Extension der Zuhörerschaft der Massenversammlungen begreifen sollen, auch wenn sie nicht an der Inszenierung partizipieren, deren Teil sie zu sein scheinen. Was die Rundfunkhörer hören, ist nicht einfach die Stimme des ›Führers‹, sondern die durch die affirmative Resonanz eines präsenten Massenpublikums imprägnierte Stimme des Redners. Hitlers Stimme ist nicht ›nackt‹, sondern durch die Massenakklamation armiert, deren Teil zu werden das Rundfunkpublikum genötigt ist. In diesem Sinne setzt der Nationalsozialismus also nicht auf die gleiche Wirkungsmächtigkeit von Rundfunk- und Massenrede, sondern auf den *Rundfunk als globale Distributionsform gelungener lokaler Kommunikation*. Die Preußische Zeitung formuliert diesen Zusammenhang:

Seine Worte werden uns Glauben und Verpflichtung sein. In allen Ortsgruppen der NSDAP findet groß angelegter Rundfunkgemeinschafts-

empfang statt. In den Dörfern und auf den Bauernhöfen, in den Sälen und Versammlungsräumen werden die Lautsprecher die Worte Adolf Hitlers wiedergeben. Eine ganze Provinz wird ihn hören und sich aufs Neue verpflichten. In den Straßen und der Stadt Königsberg werden Groß-Lautsprecher auch dem letzten Volksgenossen den Verlauf der Kundgebung und die Rede des Führers übertragen.<sup>62</sup>

Dass der Rundfunk hier nicht in seiner genuinen Medialität genutzt wird, zeigt sich darin, dass er allein in der Rezeptionsform des *Gemeinschaftsempfangs* Teil des Dispositivs ist. Der Rundfunk fungiert also als Endknoten eines Netzes von Verbreitungsmedien. Die Rezeption ist im Rahmen des Dispositivs Laut/Sprecher grundsätzlich nur in der Form des ›Gemeinschaft‹affirmierenden Kollektivempfangs konzeptualisiert. Die klassische Form des Dispositivs generiert genuine mediale Diskursformen – den *Gemeinschaftsempfang* oder die *Massenrezeption* –, die die Stimme des ›Führers‹ in die Stimme des Volkes transformieren. »Hier spricht nicht der Mann«, schreibt die Preußische Zeitung, »sondern in ihm vereinen sich die Stimmen von 68 Millionen, die gleichen Herzens und gleichen Sinnes sind.«<sup>63</sup> Zentral dabei ist, dass die Affirmation der Volksgemeinschaft in regelmäßigen Perioden als mediales Massenerlebnis iteriert werden muss, stabilisiert sie doch allein so den medialen Volkskörper.

Blickt man nun auf die institutionelle Seite, dann lässt sich die inszenatorische und organisatorische Perfektionierung des Dispositivs an einem Szenario verdeutlichen, das die NSDAP zum Abschluss des »Wahlkampfes« inszenierte.

Ganz Deutschland hört den Führer. Vor der Friedensrede Adolf Hitlers aus der Waffenschmiede des Dritten Reiches. Gemeinschaftsempfang aller Schaffenden.

Nur wenige Stunden trennen uns noch von dem Appell, den der Führer des deutschen Volkes heute Nachmittag aus der Waffenschmiede des Deutschen Reiches in Essen an das ganze Deutschland richten wird.<sup>64</sup>

Wenn heute Nachmittag das Sirensignal aus den Kruppwerken gegeben und von den Sirenen und Alarmgeräten aller Werke, Schiffe usw. im ganzen Reich übernommen und eine Minute lang durchgeführt wird, wenn gleichzeitig von 16.00 Uhr bis 16.01 eine Minute Verkehrsstille in ganz Deutschland eintritt, dann weiß sich das ganze Volk einig in dem fanatischen Willen diesen Appell des Führers mit seinem heißen Bekenntnis zur nationalsozialistischen Politik, die eine Politik des Friedens, der Ehre und der Gleichberechtigung ist, zu beantworten. Sämtliche schaf-

fenden deutschen Menschen der Faust und der Stirn werden in dieser historischen Stunde zum Gemeinschaftsempfang an den Lautsprecheranlagen versammelt sein, während bereits um 15.45 Uhr über alle Sender das Kommando ertönt: Heißt Flagge, und damit ganz Deutschland zum äußeren Zeichen seiner inneren Verbundenheit mit seinem Führer und Reichskanzler und mit der nationalsozialistischen Bewegung unter dem weithin leuchtenden Zeichen des Hakenkreuzes steht.<sup>65</sup>

Diese reichsweite Inszenierung der Volksgemeinschaft im Gemeinschaftsempfang, in der Massenbeschallung, setzte einen technisch und institutionell ausdifferenzierten Apparat voraus.

Die Firma Telefunken hatte 1934 den sogenannten »Pilzlautsprecher« neu entwickelt und ihn für die Maifeier der NSDAP auf dem Tempelhofer Feld zur Verfügung gestellt. Im Gegensatz zu den Richtungs Lautsprechern – deren Qualität bei zunehmender Distanz und schlechtem Wetter abnahm, im Gegensatz



Abb. 5  
Telefunken-  
Pilzlautsprecher

Abb. 6



auch zu deren Rückkopplungsanfälligkeit – erlaubte der Pilzlautsprecher eine qualitativ gleichmäßige Schallabstrahlung in Hallen und auf großen Flächen, wodurch sich die Übertragungsqualität in der Massenversammlung ebenso verbessern ließ wie die akustischen Aufnahmebedingungen für die vom Rundfunk übertragenen Reden. In Königsberg werden die Pilzlautsprecher erstmals zum Zwecke einer Verstärkung und Übertragung einer von Hitler in der Schlageterhalle gehaltenen Rede eingesetzt.

»Der Führer brachte ein besonders für ihn konstruiertes Rednerpult mit, das ihn überall hin begleitet.«<sup>66</sup> In der Halle ist »in seiner Nähe [...] ein Regiestand aufgebaut, der mittels einer Signalanlage jeweils die Zeichen zur Aufnahme gibt. Für die Einrichtung sind mehr als 1000 Meter Kabel gelegt worden«, Hilfestellung haben die Stadtverwaltung und der Arbeitsdienst geleistet. In der Halle arbeiten 25 Spezialisten: Tonmeister, Elektriker, Techniker, Filmleute; sie installierten mehr als 50 Pilzlautsprecher, »darunter aber auch drei mächtige Großlautsprecher«. In jenen Straßen der Stadt, »die der Führer auf seinem Wege berühren wird«, sind alle 200 bis 800 Meter gleichfalls Großlautsprecher aufgestellt worden. Annähernd 200 Fernmeldetechniker, Funkwarte und Ingenieure sind von der Rundfunkdienststelle der Gaupropagandaleitung für die Betreuung und Bedienung der mächtigen Anlage aufgebeten.

Schon am Vortag brachten Züge und Sonderzüge aus der Provinz Tausende nach Königsberg, und

[...] morgens verkündete dumpfes Brausen in der Luft das Nahen der beiden Zeppelin-Luftschiffe. Flugzeuge waren zu ihrem Empfang aufgestiegen. Ein Schwall grüner Wahlerinnerungszettel regnete aus den Zeppeli-

nen nieder, und aus den Lautsprechern der Riesenschiffe tönten die ernsten Mahnungen zur Wahl.<sup>67</sup>

Resümiert man die bislang charakterisierte Verknüpfung von Diskurs und Institution, dann zeigt sich der »Wahlkampf« des Jahres 1936 als ein Exempel für das entfaltete Dispositiv Laut/Sprecher, in dem die Zentrierung auf die jeweils *eine* – vom Leib des Redners durch techno-akustische Wandlung und Zerdehnung getrennte – *Stimme sich medial als Stimme des Volkskörpers etabliert*. Am Ende des Hitlerschen »Appells an Ostpreußen« rollte ein nicht enden wollender Beifallssturm durch die Halle und über die Straßen und Plätze Königsbergs. Der ganze Raum war *eine einzige jubelnde Stimme* der Begeisterung.<sup>68</sup>

Bei der Wahl am 29. März 1936 stimmten 98,9 Prozent »für die Liste und damit für den Führer.«<sup>69</sup>

- 1 Der Beitrag wurde im Rahmen des Projektes »Laut/Sprecher: Mediendiskurse und Medienpraxen in der Zeit des Nationalsozialismus« des SFB/FK 427 »Medien und kulturelle Kommunikation« erarbeitet. Das Konzept »Dispositiv Laut/Sprecher« markiert hier einen strukturellen Zusammenhang, der neben der technischen Apparatur Lautsprecher auch die Stimmen der Redner und ihre mentalen Dispositionen sowie die Szenarien ihrer performativen Prozessierung einbindet.
- 2 Dass der »fähige und erfahrene Redner« als »das Rückgrat der weltanschaulichen Schulungsarbeit der NSDAP« anzusehen ist, dazu vgl. Bundesarchiv, Archivalie NSD 9/65, 1539. 1927 – in diesem Jahr wurde auch das Redeverbot für Hitler wieder aufgehoben – begann die Propagandaabteilung der NSDAP mit einer Neuorganisation des bisherigen »Rednerwesens«. 1928 eröffnet Fritz Reinhardt in Herrsching am Ammersee die »Rednerschule der NSDAP«, in der die Rednerschüler u. a. in die Regeln einer Rhetorik unter den Bedingungen des Lautsprechers eingeführt wurden. Die von Reinhardt herausgegebenen »Rednermaterialien« liegen im Bestand NS 26 des Bundesarchivs Berlin als Gesamtkomplex vor. Zum Themenfeld der nationalsozialistischen Versammlungspraxis und Rednerausbildung vgl. Randall L. Buytwerk: Die nationalsozialistische Versammlungspraxis. Die Anfänge vor 1933, in: Gerald Diesener/Rainer Gries (Hg.): Propaganda in Deutschland. Zur Geschichte der politischen Massenbeeinflussung in Deutschland, Darmstadt 1996, S. 30–35. Auf Buytwerk und seine Übersetzungen der Reinhardt-Kursmaterialien ins Amerikanische bezieht sich jüngst auch Claudia Schmölders: Stimmen von Führern. Auditorische Szenen 1900–1945, in: Friedrich Kittler/Thomas Macho/Sigrid Weigel (Hg.): Zwischen Rauschen und Offenbarung. Zur Kultur- und Mediengeschichte der Stimme, Berlin 2002, S. 175–195 (hier: S. 175 f.). Schmölders stellt ihren Beitrag in den Kontext einer Untersuchung »charismatischer Mündlichkeit« und formuliert in diesem Zusammenhang die Auffassung, das »demagogische Charisma« des Redners Hitler habe der Technik des Lautsprechers und des Radios nicht bedurft, sondern sei über »unmittelbare Rede erlangt und stabilisiert« worden.
- 3 C. W. Kollatz: Sprachverstärkung für Massenansprachen, in: Funk-Anzeiger 3 (1925), S. 343 f.
- 4 Eine höchstwahrscheinlich aller erste elektroakustische Verstärkung einer politischen Rede, mit der eine Anzahl von Menschen in bis dato unvorstellbarer Größenordnung gleichzeitig adressiert werden konnte, fand bereits 1919 in einem Stadion in San Diego statt, als Woodrow Wilson über das Vorhaben einer *League of Nations* sprach.
- 5 F. R. Watson: Der akustische Entwurf von Kirchen, in: Die Schalltechnik (1928), S. 1 ff.
- 6 Zur Lautsprecher-Versuchsanlage im Kölner Dom Ende Februar und zu Ostern 1927 vgl. Ferdinand Trendelenburg: Über Bau und Anwendung von Großlautsprechern, in: Elektrotechnische Zeitschrift (1927), S. 1685–1691; auch: Kölnische Volkszeitung Nr. 281 (16.4.1927, 2. Morgenausgabe) und Nr. 292 (21.4.1927, Abendausgabe).

- 7 Ferdinand Trendelenburg: Experimentalbeitrag zur Raumakustik, in: Wissenschaftliche Veröffentlichungen des Siemens-Konzerns (1927), S. 276 f.; W. Zeller: Zur Theorie der Hörsamkeit im Raum, insbesondere in großen Hallen und auf Festplätzen, in: Centralblatt der Bauverwaltung (1935).
- 8 Walter Willms: Eine Schallübertragungsanlage großen Frequenzumfangs, in: Elektrische Nachrichtentechnik (1932), S. 68 f.
- 9 Hans Gerdien: Über die klanggetreue Schallwiedergabe mittels Lautsprechern, in: Telefonen-Zeitung, Nr. 43 (1926), S. 28–38.
- 10 Vgl. Alfons Arenhövel: Der Berliner Sportpalast und seine Veranstaltungen 1910–1973, Berlin 1990.
- 11 Anonym: Lautsprecher im Deutschen Reichstag, in: Funk 5 (1929), S. 19.
- 12 Über die elektroakustischen Installationen beispielsweise der AEG vgl. »Lautübertragungsanlagen bei Massenveranstaltungen« (Sonderdruck aus den AEG-Mitteilungen, Heft 8, September 1931), (Deutsches Technikmuseum Berlin, Historisches Archiv, GS 5183). Die AEG berichtet von folgenden Lautsprechereinsätzen der Jahre 1930/31: Beim Katholikentag in Münster 1930 wurden 120 000 Personen beschallt; bei der Marienkundgebung in Köln 1931 wurden Verstärker und Mikrofon im Chorrundgang von St. Aposteln installiert, die Lautsprecher waren auf dem Dach eines in der Nähe der Kirche gelegenen Hauses montiert; bei der Einweihung der Ford-Automobilwerke im Juni 1931 versorgten Lautsprecher Freigelände und Innenwerk; bei der im Zirkus Busch am 30.3.31 stattfindenden Gemeindebeamtenagung wurden die Lautsprecher am Beleuchtungsrundgang aufgehängt, Lautsprecher befanden sich auch im Restaurationssaal, damit auch diejenigen folgenden konnten, die im Saal selbst keinen Platz gefunden hatten; die Westfalenhalle in Dortmund wurde mit sechs beweglichen Holztürmen ausgestattet, auf denen Lautsprecher montiert waren; bei dem 1931 in Hamm stattfindenden Kreisturnfest wurden die Lautsprecher um den großen Exerzierplatz herum aufgestellt, damit man diesen »akustisch zu beherrschen« vermochte; im Mai 1931 erhielt schließlich der Flugplatz Düsseldorf seine Lautsprecheranlage.
- 13 Spätestens ab den frühen 30er Jahren verfügte die NSDAP sowohl über eigene Lautsprecheranlagen als auch über eigens zum Zweck der Lautsprecherinstallation ausgebildete Experten. Vgl. in diesem Zusammenhang die Hinweise bei Arno Schröder: Hitler geht auf die Dörfer. Der Auftakt zur nationalsozialistischen Revolution. Erlebnisse und Bilder von der entscheidenden Januarwahl 1933 in Lippe, Detmold 1938, S. 120, der anlässlich des Wahlkampfes in Lippe im Januar 1933 über den Einsatz einer Lautsprecher-Techniker-Truppe unter Leitung des Gaupropagandaleiters Leopold Gutterer, vom Einsatz NSDAP-eigener Lautsprecherwagen sowie über die »vorbildlich arbeitende Lautsprecheranlage des Parteigenossen Schäfer« (ebd., S. 132).
- 14 Die Tagebücher von Joseph Goebbels. Sämtliche Fragmente, hg. v. Elke Fröhlich, Bd. 2, München 1987, S. 187.
- 15 Ab 1920 übernahm die SA die Aufgabe, den Redner vor Störungen zu schützen, vgl. Adolf Hitler: Mein Kampf. Zwei Bände in einem Band, München 1939, S. 549: »[...] wurde kurzerhand festgestellt, dass die Herren der Versammlung wir seien [...] und dass jeder, der es wagen sollte, auch nur einen Zwischenruf zu machen, unbarmherzig dort hinausflöge, von wo er hereingekommen sei.«
- 16 So lautet auch der Titel des folgenden mit vielen illustrierenden Fotografien versehenen Bandes: Heinrich Hoffmann (Hg.): Hitler über Deutschland. Text von Josef Berchtold, München 1932. Dieser will die Wahlkampagne als ein »Volksaufklärungswerk von so ungeheuren Ausmaßen, wie es die Geschichte niemals gesehen hat«, verstanden wissen. (Ebd. S. 3.)
- 17 Vgl. Karlheinz Schmeer: Die Regie des öffentlichen Lebens im Dritten Reich, München 1956, S. 21; Gerhard Paul: Aufstand der Bilder. Die NS-Propaganda vor 1933, Bonn 1990, S. 205 f.
- 18 Joseph Goebbels: Vom Kaiserhof zur Reichskanzlei. Eine historische Darstellung in Tagebuchblättern, München 1934, S. 67.
- 19 Vgl. Schmeer: Regie des öffentlichen Lebens (Anm. 17), S. 21.
- 20 Hoffmann: Hitler über Deutschland (Anm. 16), S. 27.
- 21 Ebd. S. 23.
- 22 Ebd. S. 28.
- 23 Ebd. S. 29.
- 24 In den Tagebüchern sowie in Goebbels: Kaiserhof (Anm. 18, hier S. 54) lässt sich nachlesen, wie straff die Propagandaabteilung organisiert wurde, um den »in der Hauptsache mit Plakaten und Reden geführt[en] Wahlkrieg« zu gewinnen. Immer wieder berichtet Goebbels von Konferenzen, die systematisch die Fehler der vergangenen Tage auflisteten und Konsequenzen daraus zogen. Auf die Differenziertheit des Organisationsgedächtnisses der Propagandaabteilungen verweisen die im Bundesarchiv Berlin aufbewahrten Archivalien des Bestandes NS 22, die in den Variationen ver-



- schiedener Jahre die minutiösen Planungen der propagandistischen Großereignisse vor Augen führen.
- 25 Heinrich Salzmänn: Die Technik der Gestaltung der Propagandamittel, in: Unser Wille und Weg, Heft 6 (1933).
- 26 Vgl. Niklas Luhmann: Einfache Sozialsysteme, in: Zeitschrift für Soziologie I (1972), S. 51–65.
- 27 Vgl. Goebbels: Kaiserhof (Anm. 18): »Die Propagandaabteilung arbeitet auf Hochdruck. [...] Nun setzt die Technik des Wahlkampfes ein. [...] In der Auflage von 50 000 Stück wird eine Grammophonplatte hergestellt. Diese ist so klein, dass sie in einem Briefumschlag verschickt werden kann.« (Ebd. S. 54.) »In Berlin ist alles auf Draht. 10 Minuten für den Tonfilm gesprochen. Er soll auf den bekanntesten Plätzen der Großstädte abends vorgeführt werden.« (Ebd. S. 55.) »Der Tonfilm ist wunderbar gelungen. Man sieht sich zum ersten Male lebend im Bild und kommt sich ganz fremd vor.« (Ebd. S. 56.) »Wir erobern jetzt alle Kinotheater für unsere Propaganda. Zwar machen die Besitzer alle Schwierigkeiten, aber das Publikum will es so.« (Ebd. S. 57.)
- 28 Walter Benjamin: Das Kunstwerk im Zeitalter seiner technischen Reproduzierbarkeit, in: ders.: Illuminationen. Ausgewählte Schriften, Frankfurt 1980, S. 167.
- 29 Goebbels: Kaiserhof (Anm. 18), S. 90. Der Eintrag vom 4.5.32 lautet: »Ich spreche mit meiner Abteilung den Umbau des ganzen Propagandaapparates durch. Auf der Propaganda ruht in den nächsten Monaten die Hauptlast der Arbeit. Unsere Technik muß bis zum letzten verfeinert werden.« (Ebd. S. 91.)
- 30 Ebd. S. 304.
- 31 Ebd.
- 32 Am 10. April 1933 veröffentlicht das Reichsgesetzblatt den Wortlaut des Gesetzes, das den vollen Feiertagsschutz (einschl. Lohnfortzahlung) garantiert und den 1. Mai zum Reichsfeiertag erhebt. Deutlich wird damit die herausgehobene Position dieses Feiertages markiert, denn die Feiertagsgesetzgebung ist eigentlich Aufgabe der Länder.
- 33 Siehe dazu auch die Begründung des Reichsinnenministers (BA, R 43 II/1268, Bl. 19): »Die Einführung des Feiertags der nationalen Arbeit am 1. Mai soll ein sichtbares Symbol schaffen für den Anbruch einer neuen deutschen Volksgemeinschaft. An diesem Tag soll das deutsche Volk seiner unauflöslichen Schicksalsverbundenheit mit der deutschen Arbeiterschaft, der deutsche Arbeiter seiner Solidarität mit dem Volksganzen in feierlicher, von starkem Zukunftshoffen beseelter Form Ausdruck verleihen.«
- 34 Ausspruch aus der Rede, die Goebbels am Vormittag des 1. Mai 1933 im Berliner Lustgarten hält, zitiert nach: Hans Wendt: Der Tag der nationalen Einheit. Die Feier des 1. Mai, Berlin 1933, S. 18.
- 35 Zitiert nach: Faschismus und Ideologie, Argument Sonderband, Berlin 1980, S. 73. (Hervorh. C. E. J.)
- 36 Die folgende Schilderung bezieht sich auf die Angaben bei Hans Wendt: Der Tag der Nationalen Arbeit (Anm. 34).
- 37 Der Völkische Beobachter, Ausgabe vom 1.–3. Mai 1933, S. 1, nennt die folgenden Gruppen: Nationalsozialistische Betriebszellen, Fachschaften der Beamten, Ortsgruppen, die Vereine, Innungen, Zünfte und Sportorganisationen. Wendt: Tag der nationalen Arbeit (Anm. 34), S. 22, nennt »SA und Betriebszellen, Stahlhelm, Unorganisierte«. Zur Teilnahme hatten allerdings auch die freien und christlichen Gewerkschaften aufgerufen. Am Vortag des 1. Mai berichtete der Völkische Beobachter (30.4.33) über den Ansturm der Verbände, die ihre Teilnahme ankündigten. Diese reichten von der deutschen Turnerschaft über den Berliner Tennisverband, dem Boxverband, dem Verein für das Deutschtum im Ausland bis zum ADAC und dem deutschen Touring Club; von der Goldschmiedeinigung Schöneberg bis zu Fleischern, Bäckern und der Heilsarmee.
- 38 Der Aufmarschplan zum 1. Mai wird im Völkischen Beobachter, aber auch in allen Berliner Zeitungen bereits eine Woche zuvor veröffentlicht.
- 39 Die Arbeiter werden – zum Zeichen dessen, dass die Klassengesellschaft hier zur Volksgemeinschaft mutiert – z. T. in den Luxushotels der Stadt wohnen. Das »12 Uhr Blatt« veröffentlicht am 24.4.1933 (ohne Seitenangabe) eine Liste der Hotels, die die Arbeiter aufnehmen werden. Die Arbeiter – das vermerkt der Artikel besonders – »werden nach ihrer abendlichen Heimkehr im Hotel Obst und Blumen, Erfrischungen und Lektüre vorfinden.«
- 40 Vgl. Rundfunkreportage vom Tempelhofer Feld am 1. Mai 1933; Transkript der Schallplattenaufnahme des deutschen Rundfunkarchivs Frankfurt/M., Archivnummer: C 12160.
- 41 Berliner Beobachter. Beiblatt zum Völkischen Beobachter/Berliner Ausgabe vom 29.4.1933, S. 1.
- 42 Ca. 3000 »fliegende« Händler hatten bis zum 30.4.33 eine Lizenz erworben; vgl. Das 12 Uhr Blatt. Neue Berliner Zeitung am 1.5.33. Berlin, ohne Seite.

- 43 Vgl. dazu auch: Eberhard Heuel: Der umworbene Stand. Die ideologische Integration der Arbeiter in den Nationalsozialismus 1933–1935, Frankfurt/M./New York 1989, S. 121.
- 44 Artikel: »Tempelhof rüstet zum 1. Mai«, Königsberger Hartungsche Zeitung vom 27.4.33, o.S.
- 45 Eduard Rhein: Mit der Kraft von 5000 Stimmen. Die Großlautsprecher-Anlage für Hitlers Rede auf dem Tempelhofer Feld – Eine Glanzleistung der Technik, in: Berliner Morgenpost vom 3.5.1933, o.S.
- 46 Ebd.
- 47 Ebd.
- 48 Ebd.
- 49 Ebd.
- 50 Wendt: Tag der nationalen Arbeit (Anm. 34) S. 27.
- 51 Nordschleswiger Tageszeitung Flensburg vom 2.5.35, o.S.
- 52 Die medientechnologische Entwicklung der Konfiguration Mikrophon/Verstärker/Lautsprecher, die zu einer immensen Erweiterung des Hörbereiches führte, evozierte in den späten zwanziger Jahren einen intensiven »Stimmdiskurs«. Das neu entwickelte Mikrophon (vorher arbeitete man mit einem »Aufnahmehorn«) trennt das Sprechen und das Hören voneinander: hier der Sprecher, der die eigene Stimme an einen resonanzlosen Äther adressiert, dort der Hörer, der aufgefordert ist, die Botschaft der – durch die Ausschaltung der gesamten körpersprachlichen Rhetorik und den Wegfall der visuellen Präsenz des Sprechers charakterisierten – »disembodied voice« in einen Zustand glaubwürdiger Information rückzuüberführen. Die Position der Rundfunkleute ist eindeutig: »Vor dem Hörer erscheint was ist, was sich nicht spielen läßt: die Stimme, der Atem, das Wesentliche«; vgl. Iwan Heilbut: Stimme und Persönlichkeit, in: Rufer und Hörer III, 1933, S. 546–548. Diese Apotheose der Stimme wird vor allem von den Hörspieltheoretikern und -machern vertreten. So fordert Richard Kolb für dieses Genre, dass die Intensität des inneren Erlebnisses im Hörspiel die seelische Erlebnisfähigkeit des Hörers so in Gang setzen solle, dass die Figuren des Hörspiels zu Stimmen seines (des Hörers) Herzens und Gewissens werden: »Die entkörperte Stimme des Hörspielers wird zur Stimme des eigenen Ich«; vgl. Richard Kolb: Die neue Funkkunst des Hörspiels, in: Richard Kolb/Heinrich Siekemeier: Rundfunk und Film im Dienste nationaler Kultur, Düsseldorf 1933, S. 239 ff. Von Goebbels ist bekannt, dass er kein Freund eines derart konzipierten Hörspiels war, das er als »unreifes Produkt« charakterisierte, das mit seinen »gebrüllten Ekstasen auf die Hörer abschreckend« wirkte; vgl. Wolfram Wessels: Hörspiele im Dritten Reich. Zur Institutionen-, Theorie- und Literaturgeschichte, Bonn 1985. Auch die am 1.5.33 gesendete »Symphonie der Arbeit« wurde in der Berliner Morgenpost vom 3.5.33, ohne Seite, negativ, als »untönend und langatmig« kritisiert.
- 53 Am 1.2.33 hatte Adolf Hitler vor einem in der Reichskanzlei aufgestellten Mikrophon sein Rundfunkdebüt gegeben. Rundfunkhörer und -macher äußerten Kritik an Intonation und sprecherischer Professionalität, tags darauf wiederholte Hitler seine Ansprache für eine Schallplattenaufnahme. Diesmal sprach er zwar deutlicher, war gleichwohl aber nicht zufrieden, wie er in einem Interview zugab: »Aber die Möglichkeiten des Rundfunks auszunutzen, das will erst gelernt sein. Ich selber war vor dem Mikrophon fast verzweifelt.« Zitiert nach Ansgar Diller: Rundfunkpolitik im Dritten Reich, München 1980, S. 62 f.
- 54 Zitiert nach: Wendt: Tag der nationalen Einheit (Anm. 34), S. 31.
- 55 Ebd., S. 41.
- 56 Am Mittag war eine der acht Umformer-Maschinen auf dem Tempelhofer Feld ausgefallen, so dass sich die Gesamtlautstärke plötzlich um ein Drittel verringerte. Von den Zuhörern merkte das gleichwohl niemand, da in Bruchteilen von Sekunden auf Reserven zurückgegriffen wurde, und eine Stunde später war der Umformer durch einen vorsorglich bereitgestellten Gleichrichter ersetzt; vgl. Rhein: Kraft von 5000 Stimmen (Anm. 45), o.S.
- 57 Goebbels: Vom Kaiserhof zur Reichskanzlei (Anm.18), S. 299.
- 58 Ebenfalls am 7.3.1936, gegen 13.00 Uhr – zeitgleich zum Höhepunkt der Hitlerrede – näherten sich deutsche Truppen der Hohenzollernbrücke in Köln und überschritten den Rhein. Bereits am frühen Abend dieses Tages wurde deutlich, dass der Coup gelungen war, da weder Franzosen noch Briten militärisch einschränkten.
- 59 Die im folgenden zitierten Zeitungsmeldungen sind der »Preußischen Zeitung. Amtliches Nachrichtenblatt für alle staatlichen und städtischen Behörden, Königsberg« entnommen, hier der Ausgabe Mittwoch, 11. März 1936, ohne Seite.
- 60 Ebd.
- 61 Zahlreich sind allein in der Preußischen Zeitung die Artikel, die dieses Schema tradieren. So ein Bericht über die von Hitler in Karlsruhe gehaltene Rede vom 13.3.36: »Zehntausende mußten drau-

ßen bleiben, sie hörten aus den Lautsprechern die Rede des Führers.« Aus Anlass der Rede Görings in Königsberg ebenfalls am 13.3.36: Bereits eine Stunde vor Beginn musste die Halle wegen Überfüllung gesperrt werden, »und ähnlich ging es in den für Parallelversammlungen bereitgestellten Räumen, in die die Rede über Lautsprecher übertragen wurde.«

62 Ebd. 16.3.1936. Ein ähnliches Szenario wurde auch in München aufgebaut: »Auf der Theresienwiese allein, wo die beiden größten deutschen Lautsprecherfirmen in Zusammenhang mit den technischen Behörden der Stadt ein bisher in München noch nie dagewesenes Lautsprechernetz errichteten, arbeiteten insgesamt 250 Techniker und Hilfskräfte an der Aufstellung der 200 Lautsprecher und ebenso vieler Zuleitungsmasten, denen eine Gesamtsprechleistung von zweimal 3000 Watt zur Verfügung steht. Auch an den Ausfallstraßen des Bavariaringes sind Lautsprecher aufgestellt, die die Stimme des Führers noch weit in die an die Theresienwiese angrenzenden Straßen hineinragen werden.«; Ebd. 14.3.1936.

63 Ebd. 28.3.1936.

64 Der Blick nach Essen macht deutlich, dass dort ein ähnliches Szenario arrangiert wurde: »Schon am frühen Morgen hört man durch die Straßen Essens das Marschieren der Kolonnen. SA-Kolonnen ziehen zu ihren Aufmarschplätzen. Kopf an Kopf stehen die Volksgenossen den kilometerlangen Weg entlang. Die größte Halle Europas, die Kruppsche Lokomotiv-Bauhalle genügt nicht, um die Massen zu fassen, die den Führer sehen und hören wollen. Vierhundert Lautsprecher hat man aufgebaut, um ihnen auch das Wort des Führers zu vermitteln. Das Ruhrgebiet pilgert an diesem Tag nach Essen, dem Herzen der Schwerindustrie an der Ruhr.«; Ebd. 28.3.1936.

65 Ebd. 27.3.1936.

66 Die folgenden Zitate sind der Preußischen Zeitung vom 18.3.1936 entnommen.

67 Ebd. 27.3.1936.

68 Ebd. 19.3.1936.

69 Vgl. Ian Kershaw: Hitler. 1889–1936, Stuttgart 1998, S. 742.

Michel Chion

**MABUSE – MAGIE UND KRÄFTE DES »ACOUSMÈTRE«.  
AUSZÜGE AUS »DIE STIMME IM KINO«<sup>1</sup>**

**I. DAS ACOUSMÈTRE**

**1. DAS URSPRÜNGLICHE VERSTECKSPIEL**

Die visuelle Wahrnehmung (*vision*) des Menschen ist ebenso ausschnitthaft und gerichtet wie das Sehen (*vision*) des Kinos, das Hören aber ist ungerichtet. Wir sehen nicht, was sich hinter uns befindet, aber wir hören nach allen Seiten. Von allen Sinnen ist das Ohr vielleicht derjenige, der als erster zum Leben erweckt wird. Er funktioniert vom fötalen Stadium an, in dem der Mensch die *Stimme der Mutter* empfängt, die er nach seiner Geburt wiedererkennt. Setzt das Sehen hin-



Abb. 1

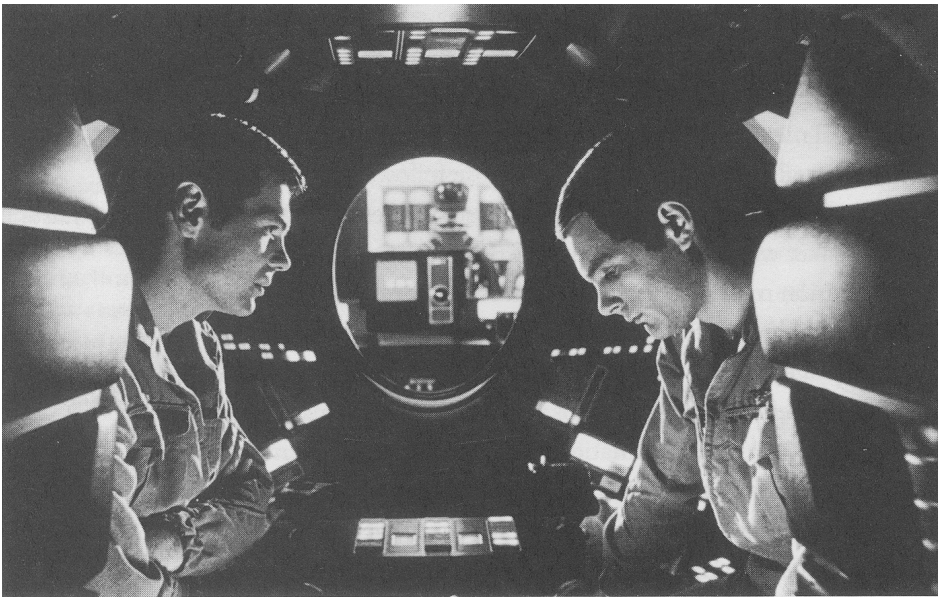
Elsie (Inge Landgut) trifft den Schatten des Mörders in M (Fritz Lang, 1931)

gegen erst nach der Geburt ein, so entwickelt es sich dafür, zumindest in unserer Kultur, zum strukturiertesten unserer Sinne. Dieser verfügt über ein beachtliches Repertoire an Formen, Vorstellungen und vor allem an Benennungen, mit dem es weder der Tastsinn noch der Geruchssinn, ja nicht einmal das Gehör auch nur im entferntesten aufnehmen können. Man findet sich vornehmlich durch das Sehen zurecht, weil es ein Benennen und Erkennen der Formen erlaubt, das tausendfach genauer und feiner ist als bei jeder anderen Art des sinnlichen Erfassens. Das Gehör hingegen ist ein ebenso scharfer wie archaischer Sinn, von dem viel im ungewissen Gebiet des Namenlosen verbleibt: *Man verspürt etwas, aber man kann es nicht sagen.*<sup>2</sup>

Der Säugling macht die Erfahrung, dass die Mutter ständig Verstecken mit ihm spielt, indem sie sich seinem Gesichtsfeld entzieht – sei es, dass sie hinter ihm vorbeigeht, sei es, dass ein Gegenstand sie verbirgt, sei es, dass er sich an sie schmiegt und sie nicht mehr sieht. Aber das olfaktorische und vokale Kontinuum, eventuell auch der taktile Kontakt, erhalten die Gegenwart der Mutter aufrecht, die nicht mehr zu sehen ist (übrigens bildet ein Minimum an Abstand und Trennung die Voraussetzung dafür, dass sie sichtbar bleibt). Manche Spiele des Kinos mit dem *Off* (eine Person verstecken, sie aber durch den Klang fühlbar halten) ver-

**Abb. 2**

Gary Lockwood und Kees Dullia in 2001: A SPACE ODYSSEY  
(Stanley Kubrick, 1968)



Michel Chion

126

wirklichen eine Art von Transposition und Reduktion dieses Auftauchens und Verschwindens, um deren Dramatik für das Kind man weiß. In gewisser Weise haben cinematographischer Schnitt und Montage mit dem Auftauchen und Verschwinden der Mutter zu tun, auch mit Spielen von der Art der ›kleinen Spule‹ (Fort! Da!), das erstmals von Freud beobachtet und von Lacan als das Modell der »repetitiven Spiele, bei denen die Subjektivität gleichermaßen die Beherrschung seiner Verlassenheit und die Geburt des Symbols anzettelt«<sup>3</sup>, analysiert worden ist. Man kann *On* und *Off* auch einen anderen Namen geben, wenn es darum geht, dass eine Person, die der Kadrierung entkommen ist, durch ihre Stimme aufrechterhalten wird, oder besser gesagt, wenn man darauf besteht, uns jemanden nicht zu zeigen, dessen Stimme man hört: Dieser Name lautet Versteckspiel.

## 2. WEDER DRINNEN NOCH DRAUßEN

Es ist hinlänglich bekannt, dass es mit der Erfindung des Sprechfilms möglich wurde, die Stimme der Schauspieler hörbar zu machen und zum Beispiel über das Bild der Garbo eine Stimme zu legen. Was man auch weiß, jedoch weniger häufig ausspricht, ist, dass es im Sprechfilm *auch* darum geht, dass uns eine geschlossene Tür oder ein undurchsichtiger Vorhang gezeigt wird und man dabei die Stimme einer Person hört, die man dahinter vermutet. Oder es geht darum, einen leeren Raum zu zeigen und die Stimme von jemandem erklingen zu lassen, den man ›da‹ vermutet, im ›hic et nunc‹ der Szene, aber außerhalb der Kadrierung. Dies heißt, dass das leerste aller Bilder oder sogar der schwarze Bildschirm von einer akusmatischen Präsenz bewohnt wird, wie es etwa bei Ophüls (*LE PLAISIR*), Welles (*THE MAGNIFICENT AMBERSONS – DER GLANZ DES HAUSES AMBERSON*) oder Duras (*L'HOMME ATLANTIQUE – DIE GRÜNEN AUGEN*) geschieht.

Akusmatisch, so ein altes Wörterbuch, »nennt man einen Klang, den man hört, ohne die ihn hervorrufende Ursache sehen zu können.« Man kann Pierre Schaeffer gar nicht genug Anerkennung dafür aussprechen, dieses seltene Wort in den fünfziger Jahren exhumiert zu haben, um eine Hörsituation zu bezeichnen, die heute ganz gewöhnlich ist, weil wir durch Radio, Telephon und Schallplatte systematisch mit ihr konfrontiert werden. Es ist eine Situation, die schon lange vor diesen Medien existierte, die aber, da sie nicht an einem spezifischen Begriff festgemacht wurde, nicht als besondere ausgemacht und erst recht nicht in ihren Auswirkungen gedacht werden konnte. Hingegen hat Schaeffer sich für seine Zwecke (es ging um die *konkrete Musik*) damit begnügt, für die andere, scheinbar triviale Situation, in der man die Quelle *sieht*, von ›direktem‹ Hören zu sprechen.

Da dieser Ausdruck eher zweideutig ist, ziehen wir es unsererseits vor, von *visualisiertem Hören*<sup>4</sup> zu sprechen.

Das Sprechkino ist natürlich vom visualisierten Klang (den man häufig »synchron« oder »on« nennt) ausgegangen, aber recht schnell wurden auch akusmatische Klangexperimente angestellt, und dies nicht nur mit der Musik, sondern vor allem mit der *Stimme*. Als Beispiel hierfür wird häufig M: EINE STADT SUCHT EINEN MÖRDER von 1930 angeführt. Dort zeichnet sich der Schatten des Kindsmörders auf dem Plakat ab, auf dem eine Belohnung auf seinen Kopf ausgesetzt ist, während seine Stimme von einem Ort außerhalb des Bildausschnitts zu dem kleinen Mädchen sagt (das sich in diesem Augenblick, und darin weicht der Film von dem bekannten Photo von den Dreharbeiten ab, auch außerhalb des Bildausschnitts befindet): »Du hast aber einen schönen Ball!« Die Koexistenz von *Stimme* und *Schatten* in dieser Einstellung, ebenso wie der Einsatz der akusmatischen Stimme zum Zwecke der Beunruhigung, sind sehr beredt. Doch schon bald sollte die Stimme ohne Schatten oder Doppelbelichtung auskommen und ihre akusmatische Person durchsetzen.

Damit wir uns recht verstehen: Der Unterschied zwischen der einen Situation (visualisiert) und der anderen (akusmatisch) besteht nicht darin, dass der Klang seine Natur, seine Präsenz, seine Distanz oder seine Farbe verändert. Er betrifft nur die Beziehung zwischen dem, was man sieht, und dem, was man hört. Die Stimme von M ist ebenso gegenwärtig und klar bestimmt, wenn man ihn nicht sieht, wie in irgendeiner der Einstellungen, in denen er zu sehen ist. Wenn man die »Tonspur« eines beliebigen Films *ohne Bild* anhört, ist es daher unmöglich, allein vom Ton ausgehend zwischen akusmatischen und visualisierten Klängen zu unterscheiden. Nur das Ansehen des Films erlaubt es, hier eine Entscheidung zu fällen, oder aber, falls man ihn schon gesehen hat, die Erinnerung an die Bilder. Das Hören ohne Bilder »akusmatisiert« alle Klänge gleichermaßen, ohne dass sie auf dieser akusmatisierten »Tonspur« (die bei dieser Gelegenheit wirklich eine Tonspur wird, ein Ganzes) irgendeine wahrnehmbare Spur ihrer anfänglichen Beziehung zum Bild bewahren.

Um zu verstehen, was bei diesem Unterschied auf dem Spiel steht, lohnt es sich, auf die alte Bedeutung des Worts »Akusmatik« zurückzukommen: So lautete, wie es heißt, der Name, den man einer pythagoräischen Sekte gegeben hatte, deren Anhänger ihren Meister *hinter einem Wandbehang* sprechen hörten – dies angeblich, damit der Anblick des Senders sie nicht von der Nachricht ablenke (so wie man im Fernsehen leicht von den Reden einer Person abgelenkt wird, wenn man sie dabei betrachtet, wie sie ihre Augenbrauen zusammenzieht oder mit ihren Händen spielt – was die Kameras mit Vorliebe in allen Einzelheiten zeigen).

Michel Chion

128

Aber diese Regel des Seh-Verbots, die aus dem Meister, dem Gott oder dem Geist eine *akusmatische Stimme* macht, findet sich natürlich in einer Vielzahl von Riten und Religionen wieder, insbesondere der islamischen und der jüdischen Religion. Man trifft auch in der Anordnung der Freudschen Kur auf sie, wo der Patient den Analytiker nicht sieht, der ihn seinerseits nicht anschaut. Und schließlich findet man sie im Kino, wo die Stimme des akusmatischen Meisters, der sich hinter einer Tür, einem Vorhang oder außerhalb des Bildfeldes versteckt, im Zentrum mehrerer wichtiger Filme steht: DAS TESTAMENT DES DOKTOR MABUSE (Stimme des bösen Genies), PSYCHO (Stimme der Mutter), THE MAGNIFICENT AMBERSONS – DER GLANZ DES HAUSES AMBERSON (Stimme des Regisseurs).

Wenn es sich bei der akusmatischen Präsenz um die einer Stimme handelt, und vor allem dann, wenn die Stimme nicht visualisiert wurde – wenn man noch kein Gesicht über sie legen konnte, hat man es also mit einem Wesen zu tun, das einer besonderen Spezies angehört. Es handelt sich gewissermaßen um einen sprechenden und handelnden Schatten, dem wir den Namen *acousmètre* verleihen – akusmatisches Wesen. Eine Person, die über das Telephon mit einem spricht und die man noch nie gesehen hat, ist ein *acousmètre*. Wenn man diese Person aber schon gesehen hat, oder, um vom Kino zu sprechen, wenn man sie weiterhin hört, nachdem sie bereits aus dem Blickfeld verschwunden ist, wo sie zuvor zu sehen war – handelt es sich dann immer noch um ein *acousmètre*? Wahrscheinlich ja, aber es gehört einer anderen Spezies an, den bereits visualisierten *acousmètres*. Es wäre natürlich ein Leichtes, die Neologismen zu vermehren, und sei es nur, um *wenigstens* zwei Fälle zu unterscheiden, je nachdem, ob man ein Gesicht über die unsichtbare Stimme legt oder nicht.

Dennoch ziehen wir es vor, die Definition des *acousmètre* offen zu halten und vermeiden es, Unterteilungen zu fabrizieren, die zwangsläufig mehr oder weniger schulmäßig wirken würden. Wir belassen diesen Begriff in seiner Unbestimmtheit, seiner Allgemeinheit. Doch soll es hier eher um jenes Wesen gehen, das man als vollständiges *acousmètre* bezeichnen kann, dasjenige, das noch nicht zu sehen war, das aber weiterhin jeden Moment im Blickfeld auftauchen kann. Vertrauter und beruhigender ist das *acousmètre*, das man schon gesehen hat und das nur zeitweise das Blickfeld verlassen hat, auch wenn in den dunklen Bezirken des akusmatischen Feldes, welches das Sichtfeld einhüllt und umfasst, auch dieses *acousmètre* wie durch Ansteckung einige der Kräfte des vollkommenen *acousmètre* erlangen kann. Uns ebenfalls vertrauter ist der freundliche akusmatische Kommentator, der sich nie zeigt, der sich aber auch nicht ins Bild hineinwagen muss. Um welche Kräfte und welche Wagnisse es sich hier handelt, werden wir später sehen.



Doch was ist mit den *acousmètre* des Radios und mit den Stimmen aus der Kulisse im Theater oder der Oper? Handelt es sich hier nicht um die gleiche Spezies, und sind wir nicht dabei, mit großem Aufwand den in den Kulissen versteckten Sprecher oder Schauspieler neu zu erfinden ?

*Das acousmètre des Radios.* Man möge uns diese Binsenwahrheit nachsehen: Das Radio ist in seiner Essenz akusmatisch. Die im Radio sprechenden *acousmètres* sind *acousmètres* kraft des Prinzips des Mediums selbst: Man riskiert also nicht, sie zu sehen, und darin besteht ihr essentieller Unterschied zum *Kino-acousmètre*. Es gibt keinen Effekt, den man mittels des Spiels zwischen ›zeigen‹, ›zum Teil zeigen‹ und ›nicht zeigen‹ erzielen könnte.

Im Kino hingegen erweist sich die akusmatische Zone als schwankend und permanent der Möglichkeit ausgesetzt, durch das in Frage gestellt zu werden, was man als nächstes sehen wird. Selbst in einem Extremfall wie *SON NOM DE VENISE DANS CALCUTTA DÉSSERT – SEIN NAME AUS VENEDIG IM VERLASSENEN KALKUTTA* von Marguerite Duras, wo man im ausgestorbenen Bild die Gesichter und Körper der die Tonspur (dieselbe wie in *INDIA SONG*) bevölkernden *acousmètres* praktisch nicht auftauchen sieht, bleibt es das Prinzip des Kinos, dass diese Gesichter und diese Körper jederzeit in Erscheinung treten und die Stimmen entakusmatisieren könnten. Anders als im Radio übrigens lässt im Kino das, was man gesehen und gehört hat, darauf schließen, was man nicht sieht, wobei die Möglichkeit der Täuschung immer bestehen bleibt. Es gibt einen Rahmen, dessen Ränder sichtbar sind: *Man sieht, wo abgeschnitten wird* und folglich auch, wo der Raum außerhalb des Bildfeldes beginnt. Im Radio, und darin besteht der große Unterschied, sieht man dies nicht.

*Das acousmètre des Theaters.* In seiner *Histoire du cinéma* erliegt Georges Sadoul der Versuchung, die in den Anfängen des Sprechfilms zum »audiovisuellen Kontrapunkt« angestellten Versuche den »klassischen Geräuschen aus der Kulisse im Theater« anzunähern.<sup>5</sup> Stimmt dies tatsächlich? Zwischen einer Stimme aus der Kulisse und der Stimme eines *Kino-acousmètre* liegt mehr als nur eine Nuance.

Im Theater wird die Stimme aus der Kulisse wirklich anderswo lokalisiert; man hört deutlich, dass sie von einem anderen Ort als der Bühne herrührt. Das Kino setzt keine *Bühne* ein, selbst dann nicht, wenn es diese einmal simuliert. Vielmehr setzt es einen *Rahmen* mit variablen Standpunkten ein, und in diesem Rahmen findet die Aufteilung in visualisierte und akusmatische Stimmen lediglich im Kopf des Zuschauers statt, der dabei stets dem folgt, was er im Verlauf des Filmes sieht und hört. In der Mehrzahl der Fälle (denn der Stereoklang bleibt noch die Ausnahme, und die Art, wie er eingesetzt wird, verändert nicht immer merklich die Spielregeln) kommt der *Off-Klang vom selben realen Ort* wie die anderen

Michel Chion

130

Klänge, das heißt aus dem gleichen Lautsprecher. Natürlich gibt es zweideutige Fälle, in denen zwischen *off* und *on* nicht immer sehr klar unterschieden wird, wie etwa bei Fellini, aber ist es nötig, daran zu erinnern, dass es nur Zweideutigkeit geben kann, wo es keine Identität gibt?

Wir sind also weit von der Stimme aus der Theaterkulissee entfernt, die wirklich als hinter der Bühne liegend wahrgenommen wird, und die einen, anders als das cinematographische Bildfeld, nicht von einem Standpunkt zum nächsten, von einer Großaufnahme zur Totalen springen lässt. Das *acousmètre* des Kinos ist ›außerhalb der Bildfläche‹ und folglich für den Zuschauer außerhalb des Bildes, aber gleichzeitig ist es *in diesem Bild*, hinter dem es hervorkommt, real (klassisches Kino) oder imaginär (Fernsehen, Autokino, etc.). So, als ob seine Stimme lauern und rastlos an der Oberfläche umherstreifen würde, *zugleich draußen und drinnen*, und um einen Ort verlegen wäre, an dem sie sich niederlassen könnte: Vor allem dann, wenn noch nicht gezeigt wurde, welchen Körper diese Stimme normalerweise bewohnt. Weder drinnen noch draußen zu sein, das ist das Geschick des *acousmètre* im Kino.

### 3. WAS VERMAG DAS »ACOUSMÈTRE«?

Je nachdem, ob das *acousmètre* bereits zu sehen war oder nicht, stellt sich die Lage unterschiedlich dar. War sie noch nicht zu sehen, wird selbst die harmloseste akusmatische Stimme, sobald sie (und sei es auch noch so geringfügig) in das Bild eingreift, mit magischen Kräften ausgestattet, die meistens unheilvoller, seltener beschützender Natur sind. Dass die Stimme in das Bild eingreift bedeutet, dass sie sich nicht damit begnügt, als Beobachterin zu sprechen, sondern dass sie mit diesem Bild ein Verhältnis der *möglichen Inklusion* unterhält, ein Macht- und Besitzverhältnis, das in beide Richtungen funktionieren kann.

Die noch nicht gesehene Stimme (zum Beispiel die Stimme Mabuses im TESTAMENT oder die Maupassants in den beiden ersten Sketchen von Ophüls' PLAISIR) besitzt eine Art Jungfräulichkeit, weil der Körper, von dem man annimmt, dass er die Stimme ausstrahlt, noch nicht ins Blickfeld gerückt wurde. Ihre *Entakusmatisierung*, oder anders gesagt, die simple Tatsache, dass man schließlich denjenigen zeigt, der spricht, ist immer wie eine Defloration, die den Verlust der an die Jungfräulichkeit des *acousmètre* geknüpften Kräfte nach sich zieht, die es gleichzeitig aber in die Reihen der Menschen zurückkehren lässt.

Die symmetrische Entsprechung der noch nicht gesehenen Stimme ist der Körper, der noch nicht gesprochen hat, die *stumme Person* (die nicht mit der Per-

son des Stummfilms zu verwechseln ist). Tatsächlich ähneln diese beiden symmetrischen Personen einander und stimmen in vielen Punkten überein. Ein ganzes Bild, eine ganze Fiktion, ein ganzer Film kann so von der Epiphanie eines *acousmètre* abhängen. Was auf dem Spiel steht, ist eine Suche, die darauf abzielt, dieses Wesen ans Licht zu bringen. Unschwer erkennt man hierin die treibende Kraft von Filmen wie MABUSE, PSYCHO, aber auch von zahlreichen fantastischen Filmen und Kriminalfilmen (*fictions*), wo es um die Entschärfung des *acousmètre* geht, welches ein verborgenes Monster, der Big Boss, das Genie des Bösen oder in selteneren Fällen ein Weiser ist . . . Dafür darf sich natürlich, wie wir bereits gesagt haben, das *acousmètre* nicht in der zurückgezogenen Position des Marktschreiers, des Kommentators, oder der Stimme der *Laterna magica* befinden, sondern die Stimme muss, wie zaghaft auch immer, *einen Fuß ins Bild*, in den Ort des Films gesetzt haben, sie muss in einem Übergangsort umhergeistern, der weder der Innenraum des Filmschauplatzes noch das Proszenium ist – ein Ort, der keinen Namen hat, mit dem das Kino aber ständig spielt.

Weil es auf dem Bildschirm ist, ohne dort zu sein, weil es auf der Oberfläche des Bildschirms umherirrt, ohne in sie einzutreten, ist das *acousmètre* ein destabilisierender, ein Spannungsfaktor, ist es eine Einladung, *nachschauen zu gehen*, und kann es auch eine Einladung sein, sich *darin zu verlieren*. Aber was hat man vom *acousmètre* zu befürchten? Und worin bestehen seine Kräfte?

Sie lassen sich in vier Aspekten zusammenfassen: überall zu sein, alles zu sehen, alles zu wissen, alles zu können. Anders gesagt: Allgegenwärtigkeit, Panoptismus, Allwissenheit, Allmacht.

*Das acousmètre ist überall*, seine Stimme dringt aus einem substanzlosen, nicht lokalisierten Körper hervor, und scheint sich von keinem Hindernis aufhalten zu lassen. Medien wie das Telephon oder das Radio, welche die akusmatischen Stimmen auf die Reise schicken und sie gleichzeitig ›hier‹ wie ›dort‹ existieren lassen, dienen in Filmen häufig dazu, dieser Allgegenwärtigkeit Gestalt zu verleihen. In 2001. A SPACE ODYSSEY – 2001. ODYSSEE IM WELTRAUM bewohnt HAL, der sprechende Computer, das gesamte Raumschiff.

*Das acousmètre sieht alles*, sein Wort ist wie das Wort Gottes: »Kein Geschöpf ist vor ihm verborgen.« Wer sich nicht im Bildfeld befindet, ist am besten positioniert, um alles zu sehen, was vorgeht; wer nicht gesehen wird, ist am besten positioniert, um einen zu sehen, wenigstens ist das die Kraft, deren Besitz man ihm unterstellt. Wenn man sich umdreht, um ihn zu überraschen, dann deshalb, weil er immer hinter einem sein könnte, und so weiter. Dies ist das obsessive oder paranoide panoptische Phantasma der totalen Beherrschung des Raums durch den Blick.

Viele Filme basieren auf der Stimme, die alles sieht: Ob es sich um den ›Chef‹ in Fritz Langs *TESTAMENT DES DOKTOR MABUSE* handelt, dessen Blick die undurchlässige Dichte eines Vorhangs durchdringt (dieser Vorhang, der dazu dient, den Blick zu verbergen, wie Barthes sagt); oder um den Computer HAL in 2001, ein ›Stimmwesen‹, das selbst als sich die Astronauten seinem Gehör entziehen, *auf ihren Lippen sehen kann*, was sie sagen; oder um die klassische Stimme des Erzählers, der von seinem zurückgezogenen Ort aus in der Lage ist, alles zu sehen; um jene klassischen Stimmen der unsichtbaren Phantome, die sich überall dorthin bewegen, wohin sich die Handlung verlagert, und denen, wie dem ›Hinkenden Teufel‹, nichts verborgen bleibt (*TENDRE ENNEMIE* von Ophüls); oder schließlich um jene Anrufer am Telefon im Thriller, die ihre Opfer mit Variationen über das Thema »Sie sehen mich nicht, aber ich sehe Sie« terrorisieren.

Ein Horrorfilm von John Carpenter, *THE FOG – THE FOG, NEBEL DES GRAUENS*, inszeniert das panoptische Phantasma auf besonders einfallsreiche Weise. Die Heldin (Adrienne Barbeau) arbeitet als Discjockey einer lokalen Radiostation in einem stillgelegten Leuchtturm, von dem aus sie die ganze Stadt überblickt. Die anderen Personen des Films kennen sie nur als *eine Stimme*, die als einzige in der Lage ist, die Situation *zu sehen*, in der sie alle befangen sind (die Stadt wird von einer unheilbringenden Wolke heimgesucht). Sie verlieren jeden Anhaltspunkt im Nebel, den allein die von den Radiowellen getragene Stimme durchdringt: Ihre panoptische Kraft strahlt von dem Leuchtturm aus, der ihr einen materiellen Ort verleiht.

Das *acousmêtre*, das alles sieht, scheint die Regel zu sein: Die Ausnahme, das Sonderbare, ist die Stimme des *acousmêtre*, das nicht alles sieht, eine Ausnahme, in der sich das panoptische Motiv in invertierter, aber expliziter Form wiederfindet. In Josef von Sternbergs Film *DIE SAGE VON ANATAHAN* spielt die Handlung auf einer Insel, auf der japanische Soldaten gestrandet sind, die man in ihrer Muttersprache reden hört. Für den westlichen Zuschauer wird diese Handlung, anstatt synchronisiert oder untertitelt zu werden, auf Englisch von Sternbergs eigener *Off-Stimme kommentiert*, die ›im Namen‹ der Gruppe von Soldaten spricht und dabei ein ungewöhnliches ›wir‹ verwendet. Dieses ›wir‹ schließt nicht die gesamte Gruppe ein, sondern eine Mehrheit, gebildet von denen, die von jeder Beziehung zur *einzigsten Frau auf der Insel* ausgeschlossen bleiben. Wenn das Bild und der es synchron begleitende japanische Ton uns Zutritt zu der Hütte gewähren, in der wir die Frau mit ihrem momentanen Gefährten entdecken, spricht die Stimme des Erzählers tatsächlich als die Stimme von jemandem, *der nicht sieht*, was vor unseren Augen abläuft und der nur Vermutungen anstellen kann (»wir haben es nicht in Erfahrung bringen können«). Anders als das Auge der Kamera ist

*sie nicht in die Hütte eingetreten.* Diese Trennung der Stimme des Erzähler-*acousmètre* vom indiskreten Blick der Kamera ist umso verstörender, als diese Stimme, die nichts zu sehen behauptet, sich dort befindet, von wo aus sie eigentlich in der Lage sein sollte zu sehen, nämlich außerhalb der Bildfläche. Daher fällt es schwer zu glauben, dass sie der Handlung, die sich innerhalb des Bildausschnittes abspielt, nicht beiwohnen soll. Man zieht es vor, in ihrer teilweisen Blindheit eine Unaufrichtigkeit zu vermuten. Das »Wir«, in dessen Namen sie spricht, verweist auf niemand bestimmtes, wie man hinzufügen muss: Nie ist es möglich, in der Masse der Soldaten das einzelne Individuum ausfindig zu machen, das die Erzählung übernimmt.

In gleicher Weise spricht ein Gutteil der akusmatischen Stimmen, die man beim Betrachten von Marguerite Duras' *INDIA SONG* hört, als Stimmen, die *nicht alles sehen*. Dieses nicht-alles-Sehen, dieses nicht-alles-Wissen, bezieht sich vor allem auf das vom Vizekonsul und Anne-Marie Stretter gebildete Paar, so wie sich dasjenige des »Wir« auf das Paar in der Hütte bezog. Wodurch uns – und es gibt andere Beispiele, auf die wir später eingehen werden – nahegelegt wird, dass das *acousmètre mit eingeschränkter Sicht* etwas mit der »Urszene« zu tun hat. *Was es nicht zu sehen behauptet ist, was das Paar tut.* So dreht sich Bertoluccis Film *LA TRAGEDIA DI UN UOMO RIDICOLO – DIE TRAGÖDIE EINES LÄCHERLICHEN MANNES* gerade um die perverse Umkehrung der Urszene (hier ist es der Vater, der nicht sieht, was der Sohn ... und – unserer Interpretation zufolge die Mutter – miteinander tun), und der Vater hat vom Sohn ein *Fernglas* geschenkt bekommen, mit dem er sich zu Beginn des Films auf dem Dach seiner Fabrik dabei vergnügt, »Ich sehe alles« zu spielen. Nun hat Bertolucci dem von Ugo Tognazzi gespielten Vater eine merkwürdige »innere Stimme« gegeben, bei der man nie weiß, wo ihr Sehen und ihr Wissen enden, was insbesondere in Bezug auf den Sohn gilt, dessen Entführung sie sieht, sowie in Bezug auf das, was sich hinter ihrem Rücken abspielt und wovon sie nichts sieht.

So ist es tatsächlich nicht am verstörendsten, unterstellt man dem *acousmètre* ein *unbegrenztes* Wissen im Sinn von »ohne jede räumliche Grenze« – sondern vielmehr, wenn dieses Sehen und dieses Wissen Grenzen haben, von denen wir nichts wissen. Die Vorstellung eines Gottes, der alles sieht und weiß (dieser Gott ist in den drei Buchreligionen, der jüdischen, christlichen, islamischen, ein *acousmètre*), diese Vorstellung mag vielleicht »anstößig« sein, wie das kleine Mädchen sagte, von dem Nietzsche spricht, aber sie ist beinahe natürlich. Viel störender ist die Vorstellung eines Gottes, eines Wesens, dessen Kräfte und dessen Sicht eingeschränkt sind, ohne dass man weiß, wo sie enden.

*Allwissenheit und Allmacht des acousmètre:* Wenn wir das *acousmètre* als ein

Wesen beschrieben haben, das mutmaßlich in der Position ist, alles zu sehen, haben wir ganz natürlich die Kräfte vorweggenommen, die sich daraus ableiten. Alles zu sehen impliziert in der Logik des magischen Denkens, in der wir uns bewegen, alles zu wissen – denn das Wissen wird einem Hineinsehen-ins-Innere gleichgesetzt – und in geringerem Maße auch, alles zu können, oder zumindest *gewisse Kräfte* zu besitzen, deren Natur und Ausmaß variieren (Unverwundbarkeit, Kontrolle über zerstörerische Kräfte, hypnotische Handlungen, etc.).

Warum all diese Kräfte für eine Stimme? Vielleicht deshalb, weil diese *Stimme ohne Ort*, denn das ist die Stimme des *acousmètre* mehr oder weniger, uns zurückversetzt in ein archaisches, ursprüngliches Stadium der ersten Lebensmonate und sogar der Zeit vor der Geburt, als die Stimme alles und überall war (doch darf man dabei nicht vergessen, dass es das Konzept dieses ›Überall‹ nur rückblickend gibt: Plötzlich erscheint es demjenigen, der nicht mehr im ungeschiedenen All-Überall des Ursprungs ist).

Der Sprechfilm ist also nicht nur ein Schauplatz, der wie in Bioy Casares' Roman *La invención de Morel – Morels Erfindung* von sprechenden Simulakren bewohnt wird; er ist auch ein Gebiet außerhalb der Bildfläche, das von akusmatischen Stimmen bevölkert werden kann, von Stimmen, die ihn begründen und für ihn bestimmend sind; von Stimmen, die dem Bild befehlen, in es eindringen und es aussaugen wie Vampire; von Stimmen, die oft alle Macht haben, die Handlung zu leiten und sie herbeizurufen, sie hervorzubringen und manchmal auch, sie auf der Grenze zwischen Erde und Wasser untergehen zu lassen. Streng genommen begründet der Sprechfilm natürlich nicht das *acousmètre*, denn das große ursprüngliche *Acousmètre* ist Jehovah, und weiterhin, für jeden von uns, Die Mutter. Aber der Sprechfilm erfindet für das *acousmètre* ein Wirkungsfeld, das ihm noch keine andere dramatische Ausdrucksform eröffnen konnte; und dies, seitdem das Aufkommen des Sprechfilms das Kino *der Stimme ausgeliefert hat*.

#### 4. DIE ENTAKUSMATISIERUNG

Dies sind die Kräfte des *acousmètre*. Natürlich genügt es, dass dieses sich zeigt, dass die Person, die spricht, ihren Körper in die Kadrierung, in das Blickfeld einfügt, damit sie ihre Macht, ihre Allwissenheit und natürlich auch ihre Allgegenwärtigkeit verliert. Dies ist eine Art symbolischer Akt, den man *Entakusmatisierung* nennen kann, die *Inkarnation der Stimme*, die das *acousmètre* dem Geschick der Sterblichen überantwortet, denn es setzt einen Ort für es fest und sagt ihm: »Hier ist dein Körper, dort wirst du sein, und nirgendwo anders«, in der gleichen

Weise, wie die Beerdigungsriten dazu da sind, der Seele des Toten zu sagen: »Du sollst nicht mehr umherirren, hier ist dein Grab.«

In wie vielen fantastischen oder Kriminalfilmen sieht man, wie das *acousmètre* zu jemandem wird, der wie alle anderen ist, sobald seine Stimme einem bestimmten Körper zugewiesen wird, dessen Konturen deutlich sichtbar werden. Meistens wird es dann, wenn nicht ungefährlich, so doch wenigstens menschlich und verwundbar. Wenn der verborgene »Big Boss« im Bild erscheint, dann lässt er sich in der Regel fangen oder umbringen »wie irgendein Idiot« (wie Pascal Bonitzer sagt und dafür das Beispiel von Aldrichs *KISS ME DEADLY – RATTENNEST* anführt<sup>6</sup>).

Die Entakusmatisierung, die gleichzeitig Enthüllung eines Bildes und eines Ortes ist, nämlich des menschlichen und sterblichen Körpers, in dem die Stimme einen Aufenthaltsort zugewiesen bekommt, ähnelt in vielerlei Hinsicht einem *Striptease*. Sie geschieht nicht unbedingt auf einen Schlag, sondern kann Schritt für Schritt erfolgen: So wie das weibliche Geschlecht der letzte Punkt ist, der in der Entkleidung enthüllt wird (der Punkt, nach dem es nicht mehr länger möglich ist, die Abwesenheit des Penis zu verneinen), gibt es einen letzten Punkt der Entakusmatisierung, und dies ist der *Mund*, aus dem die Stimme hervordringt. Wir können es also mit halben *acousmètres* zu tun haben, oder umgekehrt mit teilweisen Entakusmatisierungen, wenn man noch nicht den Mund einer Person gesehen hat, die spricht, dann aber ihre Hand sieht, ihren Rücken, ihre Füße, ihren Nacken, etc. Man kann es sogar mit einem Wesen zu tun haben, das zu einem Viertel akusmatisch ist, wenn sein Kopf von vorne als verhüllter gezeigt wird! Solange das Gesicht und der Mund nicht völlig enthüllt sind und das Auge des Zuschauers die Übereinstimmung der Stimme mit dem Mund nicht »überprüft« hat (übrigens eine ungefähre und nachsichtige Überprüfung), ist die Entakusmatisierung unvollständig und die Stimme bewahrt eine Aura der Unverwundbarkeit und der magischen Kraft.

Victor Flemings *ZAUBERER VON OZ* mit Judy Garland enthält eine hübsche Entakusmatisierungsszene, die dies äußerst gut demonstriert.

Der »Große Oz« ist der Name des von Frank Baum erfundenen Zauberers, der mit dröhnender Stimme in einer Art Tempel spricht und sich dabei hinter einem Dispositiv aus Vorhängen, fratzenhaften Masken und Rauch verbirgt. Diese donnernde Stimme tritt als Stimme auf, die alles sieht und alles weiß, denn sie ist dazu in der Lage, Dorothy und ihren Freunden zu sagen, was sie suchen, bevor diese auch nur dazu kommen, den Mund zu öffnen. Aber als sie nach Erfüllung ihrer Mission zurückkehren um einzufordern, was ihnen zusteht, weigert sich Oz, sein Versprechen zu halten und versucht, sie hinzuhalten; Entrüstung auf Seiten Dorothys, deren Hund Toto auf die Stimme zuläuft, den Vorhang herun-

Michel Chion

136

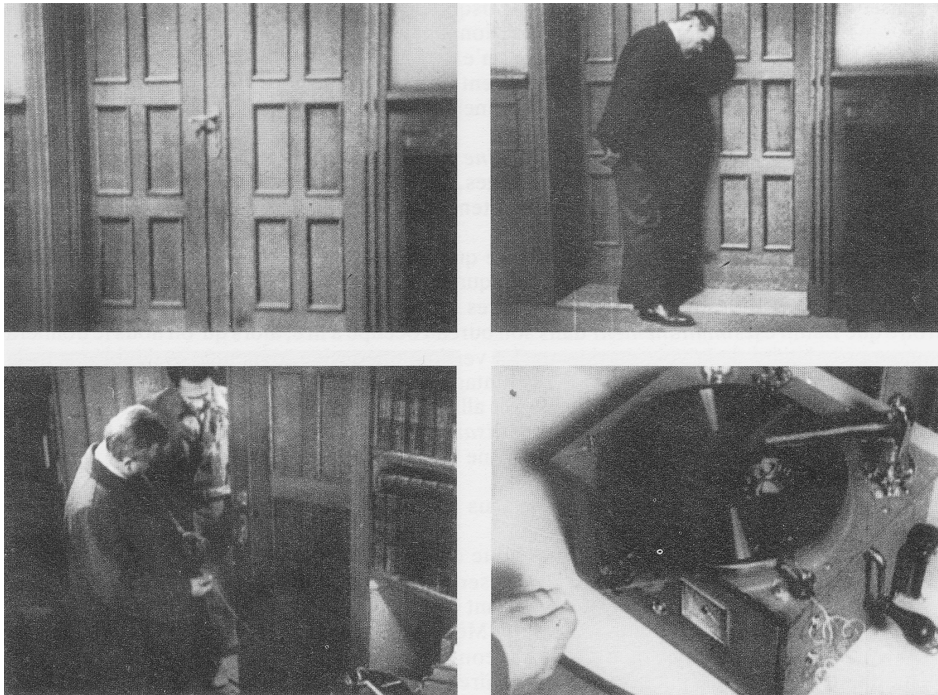


Abb. 3

Georg John, Otto Wernicke und Gustav Diesl in  
*Das Testament des Dr. Mabuse* (Fritz Lang, 1932)

terreißt, hinter dem diese sich versteckt: Er bringt ein ganz gewöhnliches Individuum zum Vorschein, das vor einem Mikrophon spricht und dabei eine naive Apparatur zur Verstärkung und Raucherzeugung bedient. Der ›Große Oz‹ ist nur ein Mensch, der sich daran ergötzt, sich zu verstecken und dabei eine dröhnende Stimme annimmt, um sich gottgleich zu machen. Und von dem Augenblick an, in dem diese Stimme ›in einen Körper gesteckt wird‹, kann man hören, wie sie ihre gigantischen Dimensionen verliert, wie sie zu einem dünnen Stimmchen wird, und wie sie so schließlich als menschliche Stimme spricht. »You're a very bad man«, sagt Dorothy – »Oh no, my dear«, sagt der Ex-Zauberer schüchtern, »I am a very good man, I'm just a bad wizard.« Die Entakusmatisierung markiert für Dorothy das Ende ihrer Initiation, jenen Augenblick, in dem sie mit der elterlichen Allmacht abschließt und den sterblichen und fehlbaren Vater entdeckt.



## II. MABUSES GESAMMELTES SCHWEIGEN

Als Fritz Lang und Thea von Harbou 1932 eine Fortsetzung ihres Stummfilms von 1922, *MABUSE DER SPIELER*, entwerfen und ihren Helden – das psychopathische Genie des Bösen – für neue Abenteuer ins Rennen schicken, scheint ihr Weg vorgezeichnet. Da das Kino seit kurzem über die Stimme verfügt, wäre nichts einfacher, als dem teuflischen Hypnotiseur eine großartige Stimme zu verleihen und dem Publikum anzukündigen: »Mabuse spricht«, so wie die Werbung für den Film *KÖNIGIN CHRISTINE* auf dem Slogan beruhte: »Die Garbo spricht!«

Statt dessen tun sie das Gegenteil, was selbst dann eine großartige Idee bleibt, wenn sie ihnen von den Umständen eingegeben wurde: Sie werden sich des Sprechfilms gerade dazu bedienen, Mabuses Stimme *nicht* erklingen zu lassen, oder genauer gesagt dazu, diese Stimme niemals mit einem menschlichen Körper zu verbinden. Die Mabuse zugeschriebene Stimme, die sich schließlich als die eines *anderen* erweisen wird, wird nur hinter einem Vorhang zu hören sein. Was den wahrhaftigen, den authentisch-ursprünglichen Mabuse anbelangt, dem Rudolf Klein-Rogge für immer seine adlerhaften Züge verleiht, so wird er bis zu seinem Tod beharrlich stumm bleiben. Man wird ihn nur als Phantom einer Doppelbelichtung sprechen hören, das die irreale und unmenschliche Stimme einer alten Hexe hat. Und weil der lebendige Mabuse schweigt, kann die Stimme, die abgelöst von jedem körperlichen Träger in seinem Namen hinter einem Vorhang spricht, die teuflischen Mächte ausüben, die man ihm nachsagt. Ein stummer Körper, eine körperlose Stimme: So teilt sich der schreckliche Mabuse auf, um besser herrschen zu können.

Man kann immerhin annehmen, dass die Autoren des *TESTAMENT DES DOKTOR MABUSE*, denn so lautet der Titel dieser neuen Abenteuer, einen guten Grund dafür hatten, aus Mabuse einen Stummen zu machen: Rudolf Klein-Rogge sprach kein Französisch! Tatsächlich wurde *DAS TESTAMENT*, wie es gängige Praxis war, in zwei Sprachen, auf Deutsch und Französisch gedreht – in den gleichen Kulissen und mit anderer Besetzung. Da die Entwicklung der Synchronisation noch nicht ganz abgeschlossen war, hatte man dieses kostspielige Mittel gefunden, um einen Film in einer Fremdsprache zu verbreiten. Aber ganz offensichtlich bestanden Lang und von Harbou darauf, den Star-Interpreten des Stummfilm-Mabuse auch für die französische Fassung beizubehalten, daher der Kniff, so zu tun, als sei Mabuse, der in seinen ersten Abenteuern im *tauben Kino*<sup>7</sup> alles andere als stumm war (aber der Zuschauer hörte seine Stimme eben nicht), zu Beginn seiner »sprechenden« Heldentaten dem Schweigen anheim gefallen. Die Stimme, die in seinem Namen spricht, wird von seinem Körper abgelöst, also

wird nur sie allein in beiden Fassungen von verschiedenen Schauspielern inkarniert. Deshalb also, so schließt der gesunde Menschenverstand, ist Mabuse stumm! Tatsächlich? Mag die Idee auch einfallsreich sein, zum Geniestreich wird sie erst dadurch, dass aus ihr die verrücktesten Konsequenzen für den Aufbau des Drehbuchs selbst gezogen wurden.

a) Auf den ersten Blick erinnert die narrative Struktur des TESTAMENTS an eine Halskette aus kleinen Szenen, die, aufgereiht wie Perlen, anfangs anscheinend ohne Zusammenhang bleiben, die sich nach und nach aber zusammenschließen, ja einen geschlossenen Kreis zu bilden scheinen. Wir haben am Anfang auf der einen Seite einen stummen Mabuse und auf der anderen eine Stimme, die in seinem Namen spricht und die seinen Tod überlebt, der sich ungefähr am Ende des ersten Drittels der Erzählung ereignet. Der Ausgang liefert uns im Prinzip die rationale Auflösung all dessen, was geheimnisvoll erscheint: Die Stimme ohne Körper, die man hinter dem Vorhang hörte, wäre demnach die des von Mabuse hypnotisierten Anstaltsleiters Baum. Anscheinend wird die ›richtige Stimme‹ wieder auf den richtigen Körper übertragen, und der sprechende Körper, das *anacousmètre*, ist wieder zusammengesetzt. Vorhang.

Befriedigt uns ein erstes Betrachten des Films annähernd mit einem Ausgang, der auf jede Frage eine rationale Antwort zu geben scheint, so fällt bei mehrmaligem Ansehen auf, dass sich der Kreis nicht wirklich schließt. Niemals sehen wir Baum dabei, wie er Mabuse seine Stimme leiht, sondern wir schließen nur darauf, dass er es tut, weil Kent die Stimme identifiziert, als er sie hinter der Tür zu Baums Arbeitszimmer hört. Anders gesagt, Kent erkennt nur in einer (von einer Platte) übertragenen Stimme eine andere (von einem Lautsprecher) übertragene Stimme wieder. Nicht nur, dass wir Baum nie diese Stimme ›aussenden‹ sehen, wir wissen auch nicht, wo er ist, nur, dass er nicht dort ist, wo man ihn vermutet. Wie kommt es, dass der Zuschauer, dem mehr Dinge zu sehen und zu hören zugestanden wird als irgendeiner Person des Films, aufgrund der Identifizierung der Stimme durch Kent die Stimme Baums nicht sicher in ihren drei Erscheinungsformen wiederzuerkennen vermag: in vivo, hinter dem Vorhang, hinter der Tür zum Arbeitszimmer? Einfach deswegen, weil die Identifizierung einer Stimme eine verwirrende Angelegenheit ist, die unter sehr wechselhaften Bedingungen erfolgt. Eine Stimme muss sehr ausgeprägte persönliche Merkmale haben (einen Akzent, eine gewisse Rauheit, oder jede andere Form von Charakter *sui generis*), um in einem Film zweifelsfrei identifiziert werden zu können. Daher auch ihr Einsatz zum Zwecke der Täuschung und die Leichtigkeit, mit der man im Namen eines anderen sprechen kann.

b) Was sieht man hinter dem von Kent zerrissenen Vorhang, als er das *acousmètre* enthüllen will? Anstelle des Mannes aus Fleisch und Blut, den man zu sehen erwartet, eine technisches Dispositiv: ein Mikrophon, den Schalltrichter eines Lautsprechers und die ausgeschnittene Silhouette, die im Gegenlicht hinter dem undurchsichtigen Vorhang zu sehen war (wenn dieser kurz von hinten erleuchtet wurde) und die dazu diente, die Anwesenheit des Meisters zu beglaubigen. Fehlt denn nichts an dieser Vorrichtung, wenn sie alles erklären soll? Doch, das Auge der Kamera.

Denn die Stimme verhielt sich wie eine *Stimme, die sieht* (*une voix qui voit*), und in Echtzeit auf die kleinste Geste derjenigen reagiert, die vor dem Vorhang erschienen. Wo ist also die optische Maschinerie, die Relaisstation des Auges, da man uns ja alle Relaisstationen der Stimme und des Ohres zeigt? Nirgendwo. Das zweite Gesicht der Stimme bleibt der Magie überlassen, den Kräften der Fernhypnose. Natürlich war das Fernsehen noch nicht erfunden, doch als Idee existierte es bereits. METROPOLIS zeigte 1926 eine Vorwegnahme, doch sah man den *Schirm*, während man hier die problematische *Kamera* hätte sehen müssen. Den Schirm zu zeigen oder die Kamera, stellt uns vor zwei unterschiedliche Probleme; im zweiten Fall muss man gleichzeitig ein *Auge* und eine *Maschine* zur Anschauung bringen, man muss also ihr Aussehen erfinden. Der letzte MABUSE von 1961 wird schließlich *die Kamera* in der vereinfachten Form von *Augen zeigen*, die in den Holzleisten der Decke verteilt sind.<sup>8</sup> Das Frappierendste aber ist, dass dieses Fehlen der Kamera alle völlig ungerührt lässt: Kent und Lilli, die niemals die Sprache darauf bringen (genauso wie sie es unterlassen, auch nur ein Wort über die von ihnen entdeckte Vorrichtung zu verlieren und sie weder unter sich noch vor Lohmann erwähnen); aber auch den Zuschauer. Diese Kamera, die nicht gezeigt werden kann, die vielleicht vor den verblüfften Augen von Kent und Lilli erscheint, wenn sie in unsere Richtung blicken (mit einem Blick, der fast ein Kamerablick ist, wie man zu recht bemerkt hat), die aber nicht im Gegenschuss gezeigt wird, der uns die Sicht der beiden vor Augen führt – wäre diese Kamera nicht die Kamera der Dreharbeiten selbst, die per definitionem unfilmbar ist, da sie ja höchstens ihren Widerschein in einem Spiegel filmen kann?

Wir würden gerne behaupten, dass es nicht im geringsten eines Organ-Auges bedurfte, um in die Ferne zu sehen, denn *dieses Auge war in der Stimme selbst des acousmètre*.

c) Und was entdeckt man schließlich in Baums Arbeitszimmer, nachdem die verbotene Tür eingeschlagen worden ist, hinter der immer die gleiche Stimme ertönte, wenn man die Klinke drückte um einzutreten: »Ich möchte jetzt nicht ge-

stört werden.«? Noch eine Vorrichtung, diesmal eine rein mechanische: Ein trickreiches System, um die Platte in Bewegung zu versetzen, auf welcher der fragliche Satz aufgezeichnet wurde. Der Eindringling selbst war es, der durch das Betätigen der Klinke zum *Agens der Stimme* wurde, die ihm den Zutritt untersagte. Man kann daraus ableiten, dass kein anderes Hindernis der Überwindung der akusmatischen Barriere im Wege stand als die freiwillige Blendung, das Begehren, an die Macht des Meisters zu glauben. Wie in *OZ* erweist sich alles als Maschinerie, als Ohrentäuschung (*trompe-l'oreille*), die nur dafür taugt, die zu täuschen, die sich drankriegen lassen.

Aber was *die Enthüllung der Maschine* vor uns versteckt, ist gerade das Wichtigste, ist das, was ins Auge sticht und das Personen wie Zuschauer völlig ungerührt lässt. Es ist die Frage: Wo sind und was machen während dieser Zeit diejenigen, die sich entfernen und die weder hinter der Tür noch hinter dem Vorhang sind?

Man sieht nicht, von welchem Zentrum, von welchem Hauptquartier aus Mabuse-Baum oder wer auch immer seine Befehle gibt, wenn die Mitglieder seiner Bande vor dem Vorhang erscheinen, man wird es niemals sehen. Noch besser, in einer dieser Szenen lässt uns die Montage glauben, dass Baum *gleichzeitig* in seinem Arbeitszimmer damit beschäftigt ist zu lesen, während man ihn doch als denjenigen ausweist, dem die Vorhangstimme gehört. (Die amerikanische Fassung, die unter der Aufsicht Fritz Langs synchronisiert und neu montiert wurde, stellt die Ordnung der Montage und der Dialoge auf den Kopf, um genau diese logischen Brüche auszulöschen, aber hier sprechen wir von der deutschen Originalversion.)

So funktionieren die beiden *Entakusmatisierungsszenen*, die hinter dem Vorhang und der Tür anstelle des Meisters, der sich dort befinden sollte, eine Maschinerie zur Stimmübertragung enthüllen, wie eine *Maske (cache)*, und dies in doppelter Hinsicht:

Sie verschieben diese Enthüllung des *acousmètre* auf später, auf einen ferneren Ort, das heißt auf niemals und nirgendwo; sie verbergen aber vor allem vor uns, dass dieses Anderswo nicht existiert.

Das Ergebnis: Wenn der Film mit einem Baum endet, der sich mit dem toten Mabuse identifiziert und seinen Platz einnimmt (so wie am Ende von *PSYCHO* Norman, von dem der Geist seiner Mutter vollständig Besitz ergriffen hat, in einer Zelle zu enden bestimmt ist), scheint sich der Kreis zu schließen. Tatsächlich aber bleiben die ganzen disparaten Elemente des Films wie die vermengten Teile eines Puzzles zurück, die sich nicht zusammenfügen lassen, auch wenn die Handlung darin besteht, ihre Zusammenfügung vorzutäuschen. Der für immer stumme Körper eines Doktor Mabuse, der sogleich nach der Abfassung seines Testaments zum Leichnam wird; dieses Testament, das er in einer *écriture automatique* abfasst



Abb. 4

Rudolf Schündler, Gustav Diessl und Wera Liessem ...

und das er in fast monströser Weise in die Welt setzt; der Phantom-Mabuse mit seinen hervorquellenden Augen, der Baum zweimal mit den geschminkten Zügen des Doktors, aber der androgynen Stimme eines Zauberers in Trance erscheint; die andere, männliche und feste Stimme des *acousmètre*, das in seinem Namen hinter dem Vorhang spricht; Baum, der abwesend ist, wenn man ihn in seinem Arbeitszimmer wähnt, der sich aber immer dann dort befindet, wenn er nach den Gesetzen der Logik anderswo sein müsste; und schließlich Mabuses Name, der in

verschiedenen geschriebenen und gesprochenen Formen zirkuliert – all diese Elemente werden uns niemals so geliefert, dass sich daraus die Zusammenfügung eines Körpers, einer Stimme und eines Namens, also eines *anakusmatischen Wesens* rekonstruieren ließe, das dann eingegrenzt, umfasst, beherrscht werden könnte und das folglich ein sterblicher und verwundbarer Mensch wäre.

Die Verkehrung der Verhältnisse geht so weit, dass der Zweifel darüber bestehen bleibt, ob der Chef hinter dem Vorhang mit Mabuse identisch ist, dessen Namen er annimmt. Und als Kent hinter der Tür zum Arbeitszimmer die Stimme erkennt, die er hinter dem Vorhang gehört hatte, ruft er aus: »Das ist die Stimme des Chefs!«, ohne den Namen auszusprechen, der die ganzen auseinandergenommenen Elemente miteinander verbinden würde. Zwischen dem ausgesprochen und dem geschriebenen Namen; zwischen dem stummen Körper des Doktors und dem sprechenden Körper Baums; zwischen Baums in die Stimme des *anacousmètre* und die des Chefs aufgespaltene Stimme organisiert Fritz Langs Film eine phantastische Schiebepuzzlepartie. Macht man einen weiteren Zug und glaubt, das Gesamtbild wiederhergestellt zu haben, bringt man anderswo ein weiteres Loch zum Vorschein, und so weiter. Fast könnte man glauben, die Autoren des TESTAMENTS, die sich den Luxus geleistet haben, das *acousmètre* Mabuse zu schaffen, welches man nicht zu fassen bekommt, welches buchstäblich »unentakusmatisierbar« ist, haben sich auch darin gefallen, in ein und demselben Film alles zu versammeln, was man mit einer Stimme ohne Körper anstellen kann.

Was also ist ein *acousmètre* im Film? Es kann ein Meister sein, der wie Pythagoras hinter einem Vorhang spricht, aber es ist auch:

- jemand, der sich für jemand anderen ausgibt,
- jemand, der nicht sagt, von wo aus er spricht (wie am Telephon),
- die Stimme eines Toten, der spricht (wie in SUNSET BOULEVARD),
- eine früher aufgezeichnete Stimme, die aus einem Mechanismus kommt,
- oder die eigene Stimme eines Maschinen-Wesens.

DAS TESTAMENT summiert alle Möglichkeiten des akusmatischen Wesens, ohne sich um die dabei aufgehäuften logischen Widersprüche zu kümmern.

Letztlich erscheint es als offensichtlich, dass Mabuse selbst nichts ist, dass es die Menschen sind, die ihn produzieren; dass er nur dadurch existiert, dass er nirgendwo lokalisiert werden kann. Wenn es einen Mabuse gibt, dann in diesem Namen ohne Identität, diesem Körper ohne Stimme, dieser Stimme ohne Ort, im allgemeinen Wahnsinn dieser auseinandergenommenen Elemente. Er ist gleichzeitig alle möglichen *acousmètres* und niemand, und letztendlich eine *Akusmaschine*.

## 1. AUF DER SUCHE NACH DER AKUSMASCHINE

Greifen wir noch einmal jenen entscheidenden Augenblick im TESTAMENT DES DOKTOR MABUSE auf, als Kent den Mechanismus enthüllt, der die Anwesenheit des Chefs simuliert. Gerade ist er gemeinsam mit Lilli in dem Zimmer mit dem Vorhang eingeschlossen worden, um dort seine Strafe zu erhalten (er hat sich Anordnungen des Chefs entzogen). Da ertönt die Stimme:

Mabuses Stimme: *Kent ! Sie waren beauftragt, die Aktion gegen die Überseebank vorzubereiten. Sie haben den Befehl nicht ausgeführt. Gehorsamsverweigerung ist gleichbedeutend mit Verrat. Auf beides steht der Tod.*

Kent: *Lassen Sie die Frau frei! Machen Sie mit mir was Sie wollen, aber lassen Sie die Frau frei!*

Die Stimme: *Sie und diese Frau werden diesen Raum lebend nicht mehr verlassen.*

Kent zieht seine Waffe und schießt auf den Vorhang, sofort wird es dunkel. Der Zuschauer sieht Blitze in dieser plötzlichen Nacht, *er weiß nicht mehr, auf welcher Seite des Vorhangs er sich befindet*: dahinter, Kent und Lilli gegenüber, oder im Gegenteil davor, so dass er mit ihren Augen sieht? Dann folgt das Geräusch des Vorhangs, der zerrissen wird, und Kent und Lilli erscheinen vor uns, wie betäubt von dem, was vor ihnen erscheint. Es ist ganz so, als würden sie das Reale entdecken: den Kinosaal mit seinen Zuschauern! Ein 180°-Gegenschuss enthüllt uns, was sie sehen: die bereits erwähnte Vorrichtung. Kent und Lilli nennen das Tableau nicht beim Namen, nur ein leises »Großer Gott!« entfährt der jungen Frau. Und die Stimme spricht weiter, *als ob nichts geschehen wäre*; aus dem jetzt sichtbaren Schalltrichter des Lautsprechers tönt es: »Sie werden diesen Raum lebend nicht mehr verlassen. Es bleiben Ihnen noch drei Stunden, um zu sterben.« Dann schweigt sie endgültig, und auf ihr Verstummen folgt ein tickendes Geräusch, das so gleichmäßig und unaufdringlich ist wie das eines Metronoms: »Was es immer auch ist,« sagt Kent, »wir müssen es finden, finden und unschädlich machen.«

Was dieses Ticken der Zeitzünderbombe betrifft, unterläuft Luc Moulet ein bezeichnender Irrtum, wenn er es mit dem ohrenbetäubenden Lärm der unsichtbaren Maschine gleichsetzt, den man in der ersten Szene des Films hört, als Hofmeister sich in einem von Erschütterungen erbebenden Raum versteckt.<sup>9</sup> Denn diese beiden Geräusche sind so unähnlich wie nur eben möglich, das erste gewaltig und andauernd wie ein gigantisches Pumpengeräusch, während Kent und Lilli einen winzigen und unbeständigen Ton hören. Im Lapsus, der seiner Erinnerung

Michel Chion

144

unterläuft, hat Moulet dennoch recht: Was diese beiden Geräusche einander annähert, ist ihre unerbittliche Regelmäßigkeit und dieselbe Unmöglichkeit, *das Herz* dieser schrecklichen Akusmaschine *schlagen zu sehen*.

Aber woraus entsteht diese Akusmaschine? Genau aus dem Verstummen der Stimme. Als ob es das *acousmètre* selbst wäre, das zur Akusmaschine würde. Kurz zuvor, als der Vorhang das Hindernis zwischen der Stimme und uns bildete, sprach sie mit Kent wie ein lebendiges Wesen, oder zumindest schien es doch so. Anschließend scheint sie von der Profanierung des Vorhangs und von dem Schuss nichts gehört oder gesehen haben und wiederholt wie ein Papagei ihre Todesankündigung. Hier nutzt die amerikanische Fassung die Synchronisation, um den Sinn der Szene zu verändern: »You stupid idiot!«, sagt die Stimme, reagiert damit auf das Ereignis und fährt fort, sich wie das Relais eines lebendigen Wesens zu benehmen. Mit Fug und Recht kann man das *Wechseln zum Automatischen*, das sich in der deutschen Originalfassung ereignet und das *acousmètre* als Mechanismus enthüllt, als wesentlich stärker und verstörender empfinden.

Mann kann dann in der ganzen Episode eine Metapher des Films und der Inszenierung enthalten sehen: Was das Paar bei der Transgression des Vorhangs entdeckt, würde der Zuschauer entdecken, nähme er die Fiktion ernst und überschritte er die Barriere, die seinen Glauben an diese Fiktion bedingt, zerrisse er den Vorhang, *diesen Vorhang, der sich selbst versteckt*, und träte er in den Raum dieser falschen Tiefe ein. Was würde er sehen? Einen Lautsprecher und flache Schatten, so platt wie die ausgeschnittene Silhouette; einen Mechanismus, nichts Lebendiges. In gewisser Weise würden Kent und Lilli entdecken, dass sie in einer Aufzeichnung, einem mechanischem Bandlauf gefangen sind. Aber nein: Der Mechanismus ist da vor ihren Augen, sie erkennen nicht, dass sie in ihn verwickelt sind, ebenso wenig wie der Zuschauer die *Desillusion* auf sein eigenes Sehen bezieht, denn er betrachtet sie als diejenige Kents und Lillis angesichts der Entblößung des akusmatischen Wesens als aufgezeichnetes Simulakrum. Wir sehen, wie sie von Erstaunen ergriffen werden beim Anblick von etwas, das auch wir sehen wollen: Der 180°-Gegenschuss, eine typisch Langsche Figur, die der *Umkehrbarkeit des Blicks* entspricht, verwirklicht unser Begehren und zeigt uns, was sie sehen, *wenn sie in unsere Richtung sehen*. Damit nähert man sich dem an, was man das Trugbild des *absoluten Gegenschusses* nennen könnte: dass die Personen des Films uns sehen könnten, wie wir sie sehen, und dass wir, die Zuschauer im Saal, weil wir von ihnen gesehen werden, in die Leinwand hineinstürzten. Finden wir nicht einen der ersten Gegenschüsse des Kinos in einem Kurzfilm von Griffith über eine Theateraufführung und zeigt dieser Gegenschuss nicht den Saal und die Zuschauer?

Aber Kent und Lilli sehen eine Vorrichtung. Sie bleiben taub und blind für





Abb. 5

... im Zimmer mit Vorhang in DAS TESTAMENT DES DR. MABUSE

diesen Saal, weil sie Simulakren sind, ohne es zu wissen – so wie die Akusmaschine sich wie eine Projektion benimmt und taub und blind für das bleibt, was die beiden gerade getan haben. Die Vorrichtung wirft ihnen ihr eigenes Simulakrenbild zurück und dem Zuschauer das Bild des Films selbst als Simulation (die Metapher des Vorhangs als Vorhang des Schauplatzes ist offensichtlich). Was Kent, Lilli und den Zuschauer nicht davon abhält, weiterhin im Modus des »ich weiß ja ... aber trotzdem« zu »funktionieren«.

Michel Chion

146

Jene andere ›Stimmenmaschine‹, die Platte mit der Endlosrille, die bei Baum gefunden und von denen in Gang gesetzt wurde, die einzutreten versuchten, war keine *Akusmaschine* – denn diese Vorrichtung wird entdeckt, benannt, fügt sich ganz in das Blickfeld ein. Hingegen ist es in dem Zimmer, in dem Kent und Lilli eingeschlossen sind, ganz klar, dass man niemals den Ort sehen wird, der nicht existiert, obwohl auf dessen Existenz das Mikrophon und der Lautsprecher schließen lassen, die den Helden enthüllt werden. Denn auch wenn es ein Mikrophon gibt, wo führt es hin, und auch wenn es einen Lautsprecher gibt, von wo aus sendet er? Ebenso wenig sieht man den zentralen Ort, an dem sich die versteckte Bombe befindet oder die Maschine, die Hofmeister halb taub machte (dass diese Maschine die Rotationspresse sein könnte, die man in einer anderen Szene stillstehen sieht, ist eine deduktive Rekonstruktion, aber der Film zeigt es nicht).

So besteht die Perfektion der Akusmaschine darin, dass ihr Zentrum unerschließbar und unmöglich zu entschärfen ist. In den Geschichten von Edgar P. Jacobs gibt es immer einen unterirdischen Schlupfwinkel, ein Zentrum, zu dem Blake und Mortimer schließlich gelangen, ein lebendiges Herz der teuflischen Maschinerie, das man nur erreichen muss, um sie mattzusetzen. In diesem Schlupfwinkel wird sich das Paar des letzten MABUSE-Films schließlich wiederfinden und damit das Ende des Zyklus markieren (DIE 1000 AUGEN DES DOKTOR MABUSE). Indessen ist der Mabuse des TESTAMENTS noch kein ›Olrik‹, und seine Akusmaschine ist noch kein irgendwo vergrabenes Zentrum mit einem Befehlshaber. Er ist ein *Geräusch*, das unauffindbare Betriebsgeräusch einer Maschine. Darin kann man eine Allegorie der metronomischen Zeit sehen, des Reigens der Stunden, der zum Tod hinführt. Aber vielleicht auch eine Allegorie der Maschinerie des in seinem Inneren belauschten menschlichen Körpers mit dem Pulsschlag seines Funktionierens, so wie sie der Fötus ›hört‹. Man rufe sich das erste Auftauchen Hofmeisters ins Gedächtnis, wie er zusammengekauert horcht, während die gigantische Akusmaschine in einem Rhythmus dröhnt, welcher eben annähernd der eines Erwachsenenherzens ist. Diese Akusmaschine hat keinen Ort, sie ist in jenem *Ringsum*, das vor der Entdeckung des Orts ist.

## 2. EIN ORTLOSER TOD

Die Hauptperson von Kubricks Film 2001 ist der Computer HAL, dessen Stimme im ganzen Raumschiff *Discovery* erklingt. Kubrick zeigt uns ein oder zwei dieser ›Augen‹, mit denen HAL sehen kann: Es sind bloße Kontroll-Lampen, die man

sich in unzähligen Exemplaren überall im Raumschiff verteilt vorstellt. Aber Kubrick hat es nicht für dienlich gehalten, jedes Mal visuell auf HALs Auge zu insistieren, wenn dieser eingreift oder man annimmt, dass er etwas sieht. Warum nicht? Weil HAL vor allem eine Stimme ist, und das genügt, denn wie Mabuse hinter dem Vorhang hat er ein Auge in seiner Stimme, einer sanften, gleichmäßigen und kalten Männerstimme, und es genügt, dass HAL mit dieser Stimme den *gesamten* Raum des Schiffs bewohnt, damit man ihm Allgegenwart, die Fähigkeit, alles zu sehen, Allwissenheit und große Kräfte unterstellt. Die Schönheit von 2001 beruht ebenso auf der großartigen Sparsamkeit dieser Mittel wie auf den überragenden Spezialeffekten.

Und wie wird der einzig überlebende Astronaut Dave Bowman nach der Eliminierung seiner Kameraden durch den verrücktgewordenen HAL diesen unschädlich machen? Indem er an den Ort geht, der das Herz und das Gehirn des *acousmètre* ist.

Die banale Lösung für die Darstellung von HALs Ende hätte darin bestanden, Explosionen mit sehr viel Lärm, Rauch und durch die Gegend geschleuderten Trümmern zu filmen. Kubricks Lösung ist unendlich ökonomischer und ausdruckstärker: HAL existiert als *Stimme*; über seine Stimme, in seiner Stimme wird er sterben. In dem Maße, wie Dave die Verbindungen seiner Schaltkreise trennt und HAL ihn anfleht, dies nicht zu tun, verändert sich seine Stimme, löst sie sich auf, stürzt sie schließlich grotesk in den Bassbereich ab wie eine Platte, die langsamer wird. Dieses Gleiten der Stimme in die Stille, die durch den Anblick des erlöschenden roten Auges verstärkt wird, ist der berührendste Tod eines *acousmètre* ...

Aber im Verlauf dieser Agonie, die live von seinem Pseudo-Bewusstsein kommentiert wird (»Dave stop, stop, will you? Dave, my mind is going, I can feel it, ... my mind is going ...«<sup>10</sup>), hat es einen präzisen Augenblick des *Übergangs zum Automatischen* gegeben, in dem dieses Subjekt HAL wieder zum Nicht-Subjekt und vom lebendigen *acousmètre* zur Akusmaschine wird, ganz so wie nach dem Durchschreiten des Vorhangs das *acousmètre* Mabuse zum Mechanismus wird. Dies geschieht, als HAL nach herzerreißenden, auf taube Ohren stoßenden Anrufungen den Ton wechselt und beginnt, mit der munteren Stimme eines jungen Computers die vorgefertigte Selbstanpreisung aufzusagen, die man ihm bei seiner »Geburt« beigebracht hatte: »Good afternoon, gentlemen, I am a HAL 9000 computer.«<sup>11</sup> Von nun an ist er nichts mehr weiter als eine Aufzeichnung, die »Daisy, Daisy, give me your answer, do, I'm half crazy ...« singt.<sup>12</sup> Und sobald er schweigt, tritt eine andere Aufzeichnung an seine Stelle, die nicht mehr von HAL kommt und Dave das Ziel seiner Mission offenbart.

Michel Chion

148

Es ist bemerkenswert, dass in 2001 wie in MABUSE der Moment des Umschlagens von akusmatischen Lebewesen zur Akusmaschine etwas Ungreifbares und Undenkbares ist, das nur durch sein Vorher und Nachher fassbar wird. Es gibt keinen Übergang vom einen zum anderen, sondern eine Collage.

Der Wechsel zum Automatischen ist vielleicht auch der Augenblick, in welchem sich im Simulakrum das Bild vom Lebendigen ›ablöst‹ und wo das Lebendige stirbt, damit das Bild leben kann, welches rein mechanische Aufzeichnung ist, so wie in Edgar Poes *Oval Portrait* oder in Augusto Geninas verkanntem Film MISS EUROPA. Louise Brooks verkörpert darin ein junges Filmsternchen, das von seinem eifersüchtigen Ehemann getötet wird, während es in einem Vorführraum seinen ersten Film, DIE VERZWEIFELTE SÄNGERIN, ansieht. Nachdem sie umgebracht wurde, nimmt der Film, der sie in Bild und Ton wiedergibt, den ganzen Bildschirm ein, als ob der Tod des Originals dem Simulakrum zu seinem vollen Leben als Mechanischem verhelfen würde ...

Doch das vielleicht Verstörendste am Tod des akusmatischen Wesens HAL ist, dass dieser Tod keinen Ort hat; die Stimme selbst ist der Ort, in den sich das Mechanische einfügt, welchen den Auftakt des Verscheidens markiert. Wortwörtliche Wiederholung der Stimme Mabuses, die von einem Ticken abgelöst wird, Entgleiten von HALs Stimme ... Seltsamer Tod, ohne Spuren, ohne Leiche.

### 3. DER TOD SPRICHT

Was für eine wirkliche Anwesenheit dieser so nahen Stimme – in der tatsächlichen Trennung! Aber auch Vorwegnahme einer ewigen Trennung! Häufig, wenn ich so lauschte, ohne die zu sehen, die von ferne zu mir sprach, schien es mir, als ob diese Stimme aus den Tiefen zu mir rufe, aus denen man nicht wieder hinaufsteigt, und ich habe die Bangigkeit kennen gelernt, die mich eines Tages umklammern sollte, wenn eine Stimme so zurückkehrte (allein, und nicht mehr länger mit einem Körper verbunden, den ich nie wiedersehen würde), um mir Worte ins Ohr zu flüstern, die ich in ihrem Vorüberhuschen auf Lippen hätte küssen wollen, die auf immer zu Staub zerfallen waren.<sup>13</sup>

So spricht Proust in diesen Zeilen aus der *Suche nach der verlorenen Zeit* über die akusmatische Stimme seiner Großmutter, die er am Telephon gehört hat. Seit die Stimme durch das Telephon und das Grammophon vom Körper isoliert werden

konnte, lässt sie ganz selbstverständlich an die Stimme des Todes denken, und mehr als wir waren diejenigen, die das Aufkommen dieser Techniken erlebt haben, für ihre wesenhafte Nähe zum Grab empfänglich. Im Kino ist die Stimme des akusmatischen Wesens häufig die Stimme des Todes: William Holdens Stimme in *SUNSET BOULEVARD*, die ihre eigene Geschichte bis zu dem Schuss erzählt, der sie in den Swimmingpool fallen lässt, und die sich als Leichnam ›sieht‹, an dem sich Lebende zu schaffen machen; Rex Harrisons Phantomstimme in Man-kiewicz's *HONEYPOT – VENEDIG SEHEN UND ... ERBEN* (ohnmächtig kommentiert sie das Scheitern ihrer posthumen Pläne); Maupassants Stimme vor der schwarzen Leinwand in Ophüls' *LE PLAISIR*, und viele andere. Auf diese Stimmen könnte man den berühmten Satz aus Poes Erzählung *The Facts in the case of M. Valdemar – Die Tatsachen im Falle Waldemar* anwenden: »Ich sage ihnen, dass ich tot bin.« Ein oxymorischer Satz, der in sich selbst seine eigene Widerrede enthält und der dennoch nichts von einem eleganten Paradox hat. Jeder erfasst unmittelbar, was er bedeutet.

Was wäre natürlicher, als dass in einem Film ein Toter als Stimme ohne Körper weiterspricht, die auf der Oberfläche des Bildschirms umherirrt? Hat die Stimme, insbesondere im Kino, nicht eine enge Beziehung zur Seele, zum Schatten, zum Double, zu jenen substanzlosen, von diesem ablösbaren Nachbildungen des Körpers, die ihn nach seinem Tod überleben und ihn manchmal sogar verlassen, während er noch am Leben ist?

Ist sie nicht die Stimme des Toten, dann gehört die *Off*-Stimme des Erzählers häufig demjenigen, der so gut wie tot ist, demjenigen, der seine Lebensbahn vollendet hat und nur noch den Tod erwartet.

### III. ÜBER DIE ICH-STIMME

Häufig geschieht es in Filmen aber auch, dass jemand, der weder tot noch so gut wie tot ist, etwas zu erzählen anhebt, woraufhin die jeweilige Handlung angehalten wird; die Stimme des Erzählers löst sich von ihrem Körper und kehrt als *acousmètre* zurück, um durch Bilder der Vergangenheit zu geistern, die ihre Worte hervorrufen. Sie spricht von einem Punkt aus, an dem die Zeit für einige Zeit aufgehoben ist. Es ist nicht nur die Verwendung der ersten Person Singular, die sie zu einer ›Ich-Stimme‹ macht. Es ist vor allem eine bestimmte Art und Weise zu erklingen und den Raum einzunehmen, eine gewisse Nähe zum Ohr des Betrachters, eine bestimmte Art, diesen für sich einzunehmen und ihn dazu zu verleiten, sich mit ihr zu identifizieren.

Man spricht auf deutsch wie auf französisch – wenn man so sagen kann – von der *Off*-Stimme, um alle akusmatischen Stimmen zu bezeichnen, die Stimmen ohne Körper, die im Film erzählen, kommentieren und die Erinnerung an die Vergangenheit auslösen. Wenn wir ›ohne Körper‹ sagen, soll das heißen, dass sie vorübergehend außerhalb des Körpers platziert, vom nun nicht mehr sichtbaren Körper losgelöst und in ein Orbit im peripheren akusmatischen Feld geschickt worden sind. Diese Stimmen wissen alles, erinnern sich an alles, sehen sich aber schnell von der Präsenz der sichtbaren und hörbaren Vergangenheit überwältigt, die sie wiedererstehen lassen: Dies ist der *Flashback*.

Das Kino hat natürlich nicht die Stimme des Erzählers erfunden. Ebenso wie es zur Begleitung seiner Geschichten die Opernmusik, die ›Orchestergraben-Musik einbezog, hat sie die Stimme des *Bänkelsängers* (*montreur d'images*) integriert, den es schon viel früher gab. Hierzu sei auf das Buch von Jacques Perriault, *Mémoires de l'ombre et du son*, verwiesen, das von den in fester Abfolge gezeigten Ansichten handelt, die im 18. und 19. Jh. über Land getragen und von Texten begleitet wurden, die für das laute Vorlesen geschrieben waren und manchmal die ›gesprochene Zeitung‹ genannt wurden. Aber wir müssen noch weiter zurück.

Schon seit grauer Vorzeit *sind es die Stimmen, welche die Bilder zeigen* und der Welt eine Ordnung der Dinge geben, diese beleben und benennen. Die erste Bildvorführerin (*montreuse d'image*) ist Die Mutter, deren Stimme noch vor jedem (eventuellen) Erlernen der geschriebenen Zeichen bewirkt, dass sich die Dinge in einer lebendigen und symbolischen Zeitlichkeit abheben. In den Funktionen des Marktschreiers und des Geschichtenerzählers wie in der traditionellen *Off*-Stimme des Kommentators besteht immer noch etwas von dieser ursprünglichen Funktion fort.

Wir haben gesagt, dass der *Ort*, von dem aus die Stimme im Kino spricht, häufig wie ein Ort *abseits* des Bildes, des Schauplatzes ist, demjenigen vergleichbar, an dem sich der Vortragende aufhält, der in Fleisch und Blut gegenwärtige Bergsteiger, der seine Heldentaten kommentiert.

Solange die Filmstimme vom Platz des abseits stehenden Bildvorführers aus zu uns spricht, sei es nun, dass der Erzähler tatsächlich anwesend ist, sei es, dass er auf der Tonspur aufgezeichnet wurde, liefert uns das Kino nichts Eigenständiges im Vergleich zur guten alten Stimme der *Laterna Magica*, zur Stimme der Mutter und des Vaters, die zu dem Kind auf ihrem Schoß sprechen, das sie über seinem Kopf reden hört und von ihren Stimmen eingehüllt wird wie in einen großen Schleier. Vielleicht lässt das Kino uns so diese intensive Nähe und Präsenz der Eltern wiederfinden, aber wahrscheinlicher ist es, dass es uns des Lebens, des Einverständnisses und des möglichen Austauschs verlustig gehen lässt.

Erst wenn die Stimme mehr oder weniger in den Ort der Leinwand ›hinein-gezogen‹ wird, wenn sie sich auf eine lebendige und dramatische Beziehung mit dem Bild einlässt, wenn sie die Inklusion in diesem Feld streift oder sich dem Auge der Kamera entzieht, verändert sich etwas. Man denke etwa an Welles' Stimme in *THE MAGNIFICENT AMBERSONS – DER GLANZ DES HAUSES AMBERSON* mit dem Mikrophon, das in der letzten Einstellung im leeren Bild auftaucht und den Platz außerhalb des Blickfelds bezeichnet, von dem aus der Erzähler spricht. Es fehlt nicht viel, dass dieser sich zeigen würde, und dies reicht aus, damit diese Stimme tiefer in den Film verwickelt wird als eine klassische *Off*-Stimme.

Zwischen dem Punkt, wo die Stimme ›in Deckung gegangen ist‹, wie Serge Daney sagt, und dem, wo sie sich ins Bild hineinwagt, gibt es keine besonders markierte Auflösung der Kontinuität: Eine Kleinigkeit genügt, um sie umschlagen zu lassen.

Aber eine Stimme wird nicht einfach so zur Stimme eines *Off*-Erzählers. Der Sprechfilm hat die Kriterien der Klangfarbe, des Raums, des Timbres kodifiziert, denen sich eine Ich-Stimme zu beugen hat, um als solche zu funktionieren. Diese Kriterien entsprechen echten Normen, von denen man selten abweicht: dramatische Normen der Interpretation, technische Normen der Aufzeichnung. Sie sind weit davon entfernt, willkürlich zu sein, und um Verwirrung in der Erzählung zu stiften genügt es, eine von ihnen nicht zu beachten.

Die Ich-Stimme im Kino ist nicht einfach wie die Stimme, die in einem Roman »ich« sagt. Um die Identifikationsleistung des Zuschauer hervorzurufen, damit sie sich dieser also mehr oder weniger aneignet, muss eine bestimmte Art und Weise der Positionierung im Raum und der Aufnahme gewählt werden. Erst diese erlaubt es ihr, wie der *Angelpunkt der Identifikation* zu funktionieren und so in uns zu erklingen wie eine Stimme in der ersten Person, so als handele es sich um unsere eigene Stimme.

Diese technischen Kriterien der Ich-Stimme sind vor allem:

a) Das Kriterium der *maximalen Nähe* zum Mikrophon, die ein Gefühl der Vertrautheit mit dieser Stimme schafft, derart, dass keinerlei Distanz zwischen ihr und unserem Ohr wahrnehmbar ist. Diese Nähe wird durch (beinahe) untrügliche klangliche Anzeichen der *Präsenz* und der *Auflösung* der Stimme spürbar, die sogar in den schlechtesten Bedingungen des (Ab)Hörens und der Wiedergabe wahrnehmbar bleiben, selbst über den ›*low-fidelity*‹-Kanal des Telephons.

b) Damit korreliert als zweites Kriterium ihre Gedämpftheit, das Fehlen jeglichen Halls, der die Stimme umhüllen würde und den Eindruck eines Raumes erzeugen könnte, der sie einschließt. So als ob, damit die Ich-Stimme wie unsere

Michel Chion

152

eigene klingt, sie sich nicht in einen bestimmten und wahrnehmbaren Raum einfügen darf. Sie muss sich selbst Raum sein. Damit aus einer *einschließenden* und sich *einschleichenden*, also einer echten Ich-Stimme, eine *eingeschlossene* und *distanzierte* Stimme wird, genügt es, am Mischpult eine Ich-Stimme etwa dadurch zu manipulieren, dass man ihr ein künstliches Echo hinzufügt und sie weiter weg mischt. Sie ist dann nicht mehr länger ein Subjekt, mit dem sich der Zuschauer identifiziert, sie ist dann ein Stimm-Objekt, das man wie einen Körper im Raum wahrnimmt.

Genau darin besteht der Unterschied, über dessen Aufrechterhaltung Hitchcock in *PSYCHO* gewacht hat: Der Unterschied zwischen den ›inneren Stimmen‹, den Stimm-Objekten, die, so nimmt man an, Marion hört, als sie in ihrem Auto flieht – und der sogenannten ›inneren‹ Stimme der Mutter, die aber in Wirklichkeit Stimm-Subjekt, Ich-Stimme ist und die am Ende des Films über die Bilder des stummen und niedergeschlagenen Norman in seiner Zelle herrscht.

In der ersten Szene sitzt Marion (Janet Leigh) am Steuer ihres Autos und spult in ihrem Kopf einen ganzen Film darüber ab, welche Äußerungen diejenigen von sich geben, die ihr vertraut hatten: der Bankdirektor, einer seiner Kollegen und der Millionär, dem das gestohlene Geld gehörte. Deren zunächst beunruhigte, dann empörte akusmatischen Stimmen sind über den Bildern des (à la Hitchcock ›neutralen‹) Gesichts der autofahrenden Marion und der an ihr vorüberziehenden monotonen Straßenlandschaft zu hören. Woran hört man, dass diese Stimmen ›in ihrem Kopf‹ erklingen und dass es nicht im Gegenteil die über sie sprechenden Stimmen sind, welche die Bilder hervorrufen, in denen sie erscheint? Daran, dass sie gemäß den Konventionen technisch manipuliert sind, welche einen Klang, indem sie ihn ›irrealisieren‹, als subjektiv etablieren – was das genaue Gegenteil des Ich-Klangs ist, da eine ›subjektive‹ Wahrnehmung im Film sich dadurch definiert, als solche *objektiviert* zu werden. Diese Manipulation erfolgt in *PSYCHO* durch eine ausgesprochen starke Filterung, die diese Stimmen Telephonstimmen ähneln lässt, sowie durch einen künstlichen Hall, der sie in einen imaginären Ort einschließt, den Ort ihres Kopfes, ihrer Vorstellungskraft. Könnten wir noch einmal von den einzelnen Elementen der Tonmischung von *PSYCHO* vor dieser Manipulation ausgehen, und *über die selben Bilder die selben Stimmen* legen, aber ungefiltert und ohne Hall, nahe und gegenwärtig, so könnten wir darauf wetten, dass sich die Beziehung umkehren würde: Anstatt umschlossen zu werden, würden die Stimmen beginnen, das Bild einzuschließen und über es zu bestimmen: Sie erschienen nicht wie das von Marion Gehörte, vielmehr sähe man Marions Gesicht als von den Stimmen hervorgerufenes Bild.

Umgekehrt zeigt die zweite der fraglichen Szenen den niedergeschlagenen



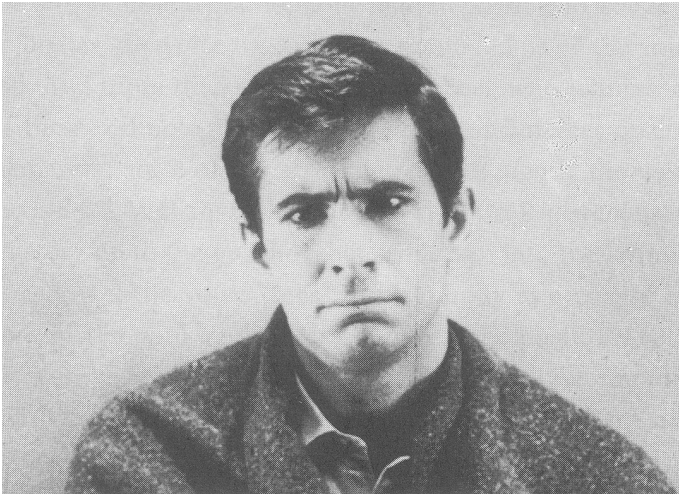


Abb. 6a/b

Janet Leigh und Anthony Perkins in  
PSYCHO (Hitchcock, 1960)



Norman (Anthony Perkins) in seiner Zelle mit einem ebenso »neutralen« Gesicht wie Marion, während die Stimme der Mutter einen paranoiden Monolog abspult. Innere Stimme Normans, von dem man uns gesagt hat, dass er sich völlig mit seiner Mutter identifiziert? Mehr als das: Sie ist nah, deutlich, einschmeichelnd, ohne Echo, es ist eine Ich-Stimme, die Normans Körper ebenso vampirgleich aussaugt wie das gesamte Bild und den Zuschauer selbst. Eine Stimme, *der das Bild innewohnt*.

Man bemerkt den Parallelismus zwischen diesen beiden Szenen, die gleichzeitig in entgegengesetzter Richtung funktionieren: die gleichen Großaufnahmen stummer, ausdrucksloser Gesichter, die beide mit akusmatischen Stimmen

Michel Chion

154

überzogen werden. Doch die inneren Stimmen, die Marion faszinieren, erklingen in ihrem Kopf – die umschließende Stimme, die sich über Normans Gesicht legt, erklingt in uns. Es ist eine Stimme, der es an einem Körper gebricht, eine Stimme, die weder in die ausgetrocknete Mumie zurückkehren kann, welche in einem Kellerraum gefunden wurde, noch in den für sie unpassenden Körper von Norman Bates einziehen kann: den lebendigen Körper des Sohnes, der von nun an von ihr besessen ist, da Norman es nicht vermag, sie zu seiner eigenen Stimme zu machen und ihr Grenzen zu setzen.

Die Extremform dieses Effekts der Besetzung des Zuschauers und seiner *körperlichen Verwicklung*, in der die Stimme ausschließlich dazu dient, in unserem Körper die Vibration des Körpers des anderen zu spüren, der Person, die als Identifikationsträger fungiert, findet man, wenn es weder Dialog noch Text noch Worte gibt, sondern nur ein sehr nahes Atmen, Röcheln oder Seufzen, von dem wir uns häufig um so schwerer distanzieren können, als Geschlecht, Alter, Identität desjenigen oder derjenigen, der oder die da atmet, stöhnt, seufzt nicht in der Stimme markiert sind. Es könnten wir, du, er, sie sein.

Dies gilt beispielsweise am Ende von 2001 für das Atemgeräusch von Dave, dem entkommenen Astronauten, das wir so stark und gegenwärtig wahrnehmen wie er es für sich in seinem Raumanzug hört – und dabei sehen wir ihn doch als winzige Marionette, verloren im interplanetaren Vakuum.

Doch bedarf es nur dieses Atmens, um aus dieser fernen und gesichtslosen Puppe, die im Vakuum oder inmitten von Maschinen umherschwebt, ein Subjekt zu machen, mit dem wir uns durch den Mimetismus unseres Hörens identifizieren.

Der gleiche Effekt der *körperlichen Verwicklung* findet in David Lynchs ELEPHANT MAN in der Szene Anwendung, als der Elefantenmensch ins Büro des Doktor Treeves geführt wird. Das Monster trägt noch seine Kapuze, und wir haben seine Gesichtszüge noch nicht gesehen. Wie gelähmt verharrt es vor dem Doktor, der ihn mit Fragen bedrängt. Aber wir hören, so gegenwärtig wie nur es selbst es hören kann, sein Atmen und sein mühevolleres Schlucken und wir verspüren seine Angst in unserem Körper. Es ist ein Beispiel für eine Szene, deren Standpunkt ganz durch den Klang geschaffen wird. Dies markiert den Extrempol der Ich-Stimme, an dem es nicht einmal mehr um die Stimme geht (der Elefantenmensch bleibt noch stumm), sondern um einen vorstimmlichen Ausdruck, bevor die Luftsäule überhaupt den Kehlkopf in Schwingung versetzt.

## 1. DER HAUSHERR

Um Besitz vom Zuschauer, vom Bild und sogar von den Personen zu ergreifen, ist es sich die Stimme schuldig, alles zu vermeiden, was auf sie als ein Objekt verweist, das erfasst werden könnte. Zu dieser Erfassung kommt es, sobald es dem Zuschauer nur gelingt, sich darüber klar zu werden, von ihr besetzt worden zu sein, sobald er sie mit ihren Umrissen, in ihrer Identität wahrnimmt. Diesen Effekt des ›Verlusts der Illusion‹ der Ich-Stimme, der Distanzierung von ihr, hat Pascal Bonitzer sehr gut analysiert:

Dem Körper der Stimme zu begegnen (Barthes sagt: dem ›Korn‹ der Stimme, dem, was vom Sinn abfällt) bedeutet, in dieser Teilung dem Subjekt dieser Stimme zu begegnen, dem in den Rang eines Objekts abgesunkenen und demaskierten Subjekt [...]. Diese Stimme hört man schließlich als Stimme.<sup>14</sup>

Um zu vermeiden, solchermaßen als Körper angetroffen zu werden, muss die Stimme sich wie gesagt in den Vordergrund schieben, ohne Resonanzphänomene zu erzeugen. Anders als die öffentliche Ansprache, die an dem Ort wiederhallen muss, an dem sich der Redner Gehör verschafft, um ihre Wirkung nicht zu verfehlen, darf sie nicht projiziert werden.

Warum wirken in Guitrys und Cocteaus Filmen die Eingriffe der Autoren aus dem *Off* derart merkwürdig? Weil diese Stimmen, obwohl sie doch die klassische Rolle der Erzähler- oder Ich-Stimme für sich in Anspruch nehmen, mit den Konventionen brechen und sich in ihrer Einzigartigkeit und als projizierte Stimmen exponieren. Anstatt neutral zu sprechen und so zu tun, als wüssten sie nicht, dass sie sich an einen Saal voller Zuschauer wenden, tragen diese ungewohnten *acousmètres*, Cocteau in *LES ENFANTS TERRIBLES – DIE KINDER DER NACHT* und Guitry in *LE ROMAN D'UN TRICHEUR – ROMAN EINES SCHWINDLERS*, in ihrer Diktion, ihrer Artikulation, ihrem Timbre klar und deutlich der *Distanz* Rechnung, die sie von uns trennt. Und auch wenn dieses *acousmètre* »ich« sagen kann, erlaubt es uns nicht, uns für dieses Ich zu halten. Cocteaus Stimme in Melvilles *LES ENFANTS TERRIBLES* klingt eher wie die Stimme eines Schriftstellers bei einem Vortrag denn wie die gewohnte Stimme des Erzählers. Gleiches lässt sich von Guitrys Stimme sagen, die auf uns geworfen wird, damit er *sich* vermittelt des durch seine Art des Deklamierens postulierten Zuhörers *beim Sprechen zuhören kann* (*s'écouter parler*). Sie lässt sich nicht als ›innere Stimme‹ oder als die Stimme aller assimilieren. Denn eine gewisse Neutralität des Timbres und der

Michel Chion

156

Aussprache wird, in Verbindung mit einer gewissen einschmeichelnden Diskretion, eben normalerweise von der Ich-Stimme erwartet. Damit jeder Zuschauer sie zu der seinen machen kann, muss sie sich etwas annähern, das ein *geschriebener Text* sein könnte, der im Innersten der Lektüre *spricht*.

Wenn wir hören, dass eine *Off*-Stimme sich beim Sprechen zuhört, schirmt uns das Bild eines Körpers, einer Person auf der Leinwand von der Möglichkeit ab, uns mit ihr zu identifizieren. Deutlich sichtbar nimmt sie ihren Platz zwischen dem Bild und uns ein, anstatt uns in es einzuführen und uns daran anhaften zu lassen. Daher rührt die unerhörte und faszinierende Wirkung, welche diese falschen Ich-Stimmen Cocteau und Guitrys im Kino entfalten können, die, während sie doch die Narration tragen, sie mit ihrer Präsenz, ihrer Korpulenz beladen. Man kommt nicht um sie herum, will man in die Erzählung eintreten, aber sie werden nicht von einem ablassen, wie ein aufdringlicher Hausherr, der darauf besteht, einen überallhin zu begleiten.

## 2. DER TAG, ALS DIE »ACOUSMÈTRES« ZWEIFELTEN

Der Tag, als die *acousmètres* zweifelten, als sie sich nicht mehr wie Stimmen verhielten, die alles wissen und alles sehen ... Kann man diesen Tag datieren? Lässt sich nicht sagen, dass es im Sprechfilm neben den Befehle erteilenden, einschüchternden und allsehenden *acousmètres* immer eine andere Spezies gegeben hat, die zweifelt, der kein vollständiges Wissen und Sehen zuteil wird? Sternbergs Stimme in der SAGA VON ANATAHAN, von der weiter oben die Rede war, gehört dieser Spezies an mit ihrer Art und Weise, ›wir‹ zu sagen und ihrem eingeschränkten Wissen im Verhältnis zu den Bildern, die sie begleitet. Auch heute haben wir uns noch nicht an diese Stimmen mit unklarem Status gewöhnt. Gleichwohl hat es den Anschein, als ließe sich an einer ganzen Reihe jüngerer Filme ein gewisser Zusammenbruch der *Off*-Stimme als Repräsentation des Wissens des Anderen, des Meisters, ablesen.

In Bertoluccis bereits zitiertem Film LA TRAGEDIA DI UN UOMO RIDICOLO – DIE TRAGÖDIE EINES LÄCHERLICHEN MANNES erzeugt die innere Stimme der Hauptperson Primo Effekte des Zweifels, welche die übliche Wirkung der *Off*-Stimme um so mehr ins Gegenteil verkehren, als diese Stimme vom Regisseur zu einem Gutteil erst nachträglich hinzugefügt worden ist, angeblich, um für größere Klarheit zu sorgen,<sup>15</sup> in Wirklichkeit aber, um den Film noch komplexer zu machen. Gewiss wird durch diese Stimme offensichtlicher, dass der Film *vom Standpunkt Primos aus* gedreht ist (weil die innere Stimme einer Person bei deren Auf-

tritt ausreicht, um diese Szene ihrem Blick zu unterstellen). Da sie jedoch über Bilder gleitet, denen der ›Erzähler‹ allem Anschein nach nicht beiwohnt, erzeugt sie eine noch verwirrendere und beunruhigendere Wirkung als bei Sternberg. Tatsächlich wissen wir in ANATHAN zumindest, worauf sich die wahrhaftige oder behauptete Unwissenheit der *Off*-Stimme bezieht. Bei Bertolucci kommt es zur völligen Verwischung der Grenzen dieses Wissens und selbst seines Gegenstands.

Es scheint so, als sei die blinde, nur mit einem ausschnitthaften Sehen ausgestattete Stimme diejenige des aus der Urszene ausgeschlossenen Dritten. Ausgeschlossen ist vermutlich nicht das richtige Wort, da die Szene ja *nur für ihn* existiert und er ihren Knoten schürzt. Hier muss man natürlich an Marguerite Duras denken, an LE RAVISSEMENT DE LOL V. STEIN – DIE VERZÜCKUNG DER LOL V. STEIN, einen Film, der zur Vorlage (*matrice*) einer ganzen Reihe literarischer und cinematographischer Werke geworden ist, in denen es blinde oder halbblinde Stimmen gibt, die nicht alles sehen, nicht alles wissen. Doch handelt es sich vor allem um Frauenstimmen, während (es wird Zeit, darauf hinzuweisen) *die Mehrheit der acousmètres dem männlichen Geschlecht angehören*. Die weiblichen *acousmètres* sind im klassischen Kino selten und außergewöhnlich (wie die Stimmen der Ehefrauen in Mankiewicz' HONEY POT – VENEDIG SEHEN ... UND ERBEN, die ihrerseits auch als dritte Personen in Bezug auf den Ehemann und die ›andere Frau‹ auftreten).

Hingegen verleihen jüngere ›Manipulationen‹ der *Off*-Stimme in so unterschiedlichen Filmen wie Terence Malicks THE DAYS OF HEAVEN – IN DER GLUT DES SÜDENS und Claude Lelouchs LES UNS ET LES AUTRES – EIN JEDLICHER WIRD SEINEN LOHN EMPFANGEN (selbst wenn ein sehr junges Mädchen mit einer der Stimmen betraut wird) hervorragend einem spezifisch männlichen Voratz Ausdruck: Dieser zeigt sich in der Art und Weise, die Herrschaft des akusmatischen Wesens und seine Fähigkeit, alles zu sehen, zum Zusammenbruch zu bringen und dadurch die Verkehrung ihrer Wirkung ins Gegenteil zu erreichen.

Im ersten dieser beiden Filme ist die *Off*-Stimme einmal mehr die des Dritten, der Dritten im Verhältnis zum Paar: Es handelt sich um die jüngere Schwester des Helden, und diese Stimme spielt ein sehr eigentümliches Versteckspiel, was ihr Wissen über die Sex- und Gewaltbeziehungen der Erwachsenen anbelangt (Malick hatte bereits in seinem ersten Film BADLANDS versucht, die *Off*-Stimme zu erneuern und ihr einen ungewohnten Charakter zu geben, der dazu geeignet war, die übliche Erzählung zu zerschlagen und den Standpunkt des Zuschauers aufzulösen). Bei Lelouch ist die Stimme, wenn man so sagen kann, auf unbefangener Weise ›schräg‹ in ihrer Beziehung zur Erzählung. Denn der Autofilmer hat mit Francis Huster einen seiner Darsteller gebeten, neben der von

Michel Chion

158

ihm gespielten Person dem erklärenden Kommentar des Films, der Verlesung der gesprochenen *Credits* und sogar der Simultanübersetzung von Fremdsprachensequenzen (das Vorlesen von Briefen) seine Stimme zu verleihen, ebenso wie einigen Lautsprecherdurchsagen in Szenen, die in einem Konzentrationslager spielen! Selten hat man eine Stimme gesehen, die so jedes Orts, jeder Position entkleidet war und als akusmatisches Faktotum diente, um es vorsichtig auszudrücken. Es scheint uns, dass dieses stimmliche Faktotum sehr symptomatisch ist für die Verwirrung zwischen den Standpunkten der einen und der anderen, die Lelouch herbeizuführen sucht.

Warum soll dieser Zusammenbruch, dieses Aus-der-Art-schlagen der Narration, stärker in der Stellung der *Off*-Stimme zum Ausdruck kommen als in jedem anderen Element der Erzählung? Eben weil die *Off*-Stimme über das Subjekt/Sujet der Geschichte Auskunft gibt, und dies in der zweifachen Hinsicht dessen, ›was geschieht‹, und desjenigen, dem es ›widerfährt‹, denn sie wirft die Frage nach dem Wissen und dem Begehren des Subjekts, nach seinem *Standpunkt* auf. Aus sehr unterschiedlichen Gründen sind in den Filmen Bertoluccis, Malicks und Lelouchs der Ort, von dem aus das *acousmètre* spricht, die Instanz oder das Begehren, die es verkörpert, in mehr oder weniger bewusster und manipulativer Weise ›aus der Bahn geworfen‹ und durcheinander gebracht worden, und dies ist kein Zufall: Zeichen der Zeit, einer Zeit, in der eine Geschichte zu erzählen den Erzähler wahrscheinlich sehr viel mehr exponiert als früher. Bei diesen drei Autoren könnte es tatsächlich einen mehr oder weniger gerissenen Versuch geben, »die Geschichte zu verstecken, die sie erzählen«, um eine ausgezeichnete Formulierung von Uriel Peres aufzunehmen.

All das gehört einer ›seltsamen‹ Periode des Kinos an, in der man einen unglaublich sprunghaften Anstieg der Zahl von zwischen zwei Stühlen sitzenden Filmen, Geschichten und Autoren hat beobachten können. Und dabei darf man diejenigen nicht vergessen, die zwischen diese zwei Stühle sozusagen ihren eigenen stellen und darauf arbeiten, wie es Raul Ruiz tut. In der Art, wie er den ›Hausherrenton‹ der Kommentarstimmen parodiert und zwei Stimmen und zwei Kenntnisstände in ihrem wechselseitigen Verhältnis gegeneinander ausspielt, baut ein Film wie *L'HYPNOSE DU TABLEAU VOLÉ* auf einem subtilen Spiel mit der traditionellen Stellung des *acousmètre* auf, das sich als solches zu erkennen gibt, und der Zuschauer wird alles andere als hinterrücks in dieses Spiel hineingezogen. Das *acousmètre* wird immer häufiger zu einem komplizierten und berechnenden Wesen: Denn das Kino jeder Epoche hat das *acousmètre*, das es verdient.

übersetzt von Michael Cuntz

- 1 Der vorliegende Text ist die Übersetzung des ersten Teils von Michel Chion: *La voix au cinéma*, Paris 1982. Es darf nach wie vor als das Referenzwerk zur Stellung der Stimme im Kino gelten. Dieser erste Teil setzt sich aus drei Kapiteln zusammen, die hier ungekürzt wiedergegeben werden. (Anm. d. Übers.)
- 2 Und so überraschend dies auch erscheinen mag, hat es bis zum zwanzigsten Jahrhundert und den Arbeiten von Pierre Schaeffer gedauert, bis man den Versuch der Benennung der Klangwahrnehmungen in ihrer konkreten Allgemeinheit erleben konnte. Siehe diesbezüglich Pierre Schaeffer: *Le traité des objets musicaux*, Paris 1966.
- 3 In Jacques Lacan: *Écrits*, Paris 1966, S. 318.
- 4 Der häufig verwendete Begriff *off* ist viel zu mehrdeutig und sein Sinn zu fließend, daher haben wir es vorgezogen, den Term »akusmatisch« wieder aufzunehmen.
- 5 Georges Sadoul: *Histoire du cinéma mondial*, Paris 1972, S. 234.
- 6 Pascal Bonitzer: *Le regard et la voix. Essais sur le cinéma*, Paris 1976, S. 32.
- 7 In der Einleitung zu seinem Buch führt Chion aus, dass die Bezeichnung Stummfilm erst nachträglich, bei der Einführung des Tonfilms geprägt wurde, eine Bezeichnung, die jenen Vokozentrismus illustriert, der den Zuschauer im Vergleich zur Stimme alle anderen Klänge als marginal wahrnehmen lässt. Gleichzeitig hält er die Bezeichnung für unzutreffend, denn ein Kino ohne sonore Begleitung – und häufig war dies die Stimme eines Kommentators oder Erzählers – hat es de facto nie gegeben. Ebenso wenig sei es stumm in dem Sinne gewesen, dass die Personen nicht gesprochen hätten – ganz im Gegenteil. Daher hält Chion es für adäquater, vom tauben Kino zu sprechen, denn dieses hatte kein Ohr, um die Stimmen und Geräusche der Handlung zu hören und weiterzugeben: »Wenn man es »taub« nennen kann, dann insofern, als uns das Kino von den realen Geräuschen der Handlung abschnitt, als es kein Ohr für das *hic et nunc* der Handlung besaß.« Chion: *La voix* (Anm. 1), S. 17. (Anm. d. Übers.)
- 8 Es handelt sich um den Film *DIE 1000 AUGEN DES DOKTOR MABUSE*.
- 9 In seiner Monographie über Fritz Lang. Luc Moullet: *Fritz Lang*, Paris 1963.
- 10 In der deutschen Fassung heißt es: »... Dave hör auf, ich bitte dich, hör auf, Dave, bitte lass es sein ... Dave mein Gedächtnis, mein Gedächtnis schwindet, ich spüre es, ich spüre es ...« (Anm. d. Übers.)
- 11 »Guten Tag, meine Herren, ich bin ein HAL 9000 Computer ...« (Anm. d. Übers.)
- 12 In der deutschen Fassung singt HAL *Hänschen klein*. (Anm. d. Übers.)
- 13 »Présence réelle que cette voix si proche – dans la séparation effective! Mais anticipation aussi d'une séparation éternelle! Bien souvent, écoutant de la sorte, sans voir celle que me parlait de loin, il me semblait que cette voix clamait des profondeurs d'où l'on ne remonte pas, et j'ai connu l'anxiété qui allait m'étreindre un jour, quand une voix reviendrait ainsi (seule, et ne tenant plus à un corps que je ne devais plus jamais revoir) murmurer à mon oreille des paroles que j'aurai voulu embrasser au passage sur des lèvres à jamais en poussière.« Marcel Proust: *Le côté de Guermantes*, in *ders.: À la recherche du temps perdu. Bd. 2*, Paris 1954, S. 132. (Übers. M. C.)
- 14 Pascal Bonitzer: *Le regard et la voix. Essais sur le cinéma*, Paris 1976, S. 42 f.
- 15 Siehe die *Cahiers du cinéma* Nr. 330/1981.

**Gereon Blaseio**

»GENDERED VOICES« IN DER FILMSYNCHRONISATION.

»FIRST BLOOD« VERSUS »RAMBO«

Während man nicht von einer weltweiten Dominanz der Filmsynchronisation sprechen kann – international kommen verschiedene Formen der Filmübersetzung zum Einsatz<sup>1</sup> –, ist doch zumindest der deutsche und österreichische Filmmarkt maßgeblich durch diese Methode geprägt. Von den 290 fremdsprachigen Filmen, die 1998 in Deutschland zur Uraufführung kamen, liefen 207 in synchronisierten Fassungen, 81 untertitelt und 2 in der ›Original‹sprache.<sup>2</sup> Bei keinem dieser 83 nicht synchronisierten Filme handelt es sich um eine *mainstream*-Produktion; zudem finden sich in der Liste der 100 einspielstärksten Kinofilme des Jahres 1998 ausschließlich Synchronfassungen.<sup>3</sup> Auch in anderen Ländern dominiert das Verfahren der Synchronisation, u. a. in Spanien, Italien und Frankreich, aber auch in Japan. Die jeweils präferierte Filmübersetzungsmethode ist dabei beim bzw. kurz nach dem Medienumbruch vom Stumm- zum Tonfilm ausgehandelt worden<sup>4</sup> und danach weitgehend stabil geblieben; in Deutschland kann die Filmsynchronisation ab Mitte der 30er Jahre als etabliert gelten.

Die Bezeichnung der Synchronisationsbranche als »schwarzes Gewerbe«<sup>5</sup> verweist zum einen auf die Produktionsbedingungen der Synchronisation in abgedunkelten Ateliers, zum anderen aber auch auf die Geringschätzung und das daraus resultierende geringe Interesse, das der Filmsynchronisation über lange Zeit zuteil geworden ist. Dies trifft aber insbesondere seit den 90er Jahren nicht mehr zu: Es lässt sich eine verstärkte Thematisierung der Synchronisation in den Printmedien, dem Fernsehen, aber auch in spezifischen Internetforen und auf zumeist die Stimmen nach SprecherInnen und Sprechern aufschlüsselnden Homepages feststellen. Und nicht zuletzt hat das neue Medium der DVD auch einem breiteren RezipientInnenkreis die Möglichkeit eröffnet, durch direkte Anwahl verschiedener Tonspuren die ›Original‹- mit Synchrontonspur(en) zu vergleichen.<sup>6</sup> Die verstärkte Aufmerksamkeit richtet sich vor allem auf die ›Stimmen‹ der Synchronisation, die (gar nicht so zahlreichen) SprecherInnen, die in Deutschland für die Filmsynchronisation eingesetzt werden. So erschien 2001 ein Lexikon, das die SynchronsprecherInnen aus über 2000 Filmen und Fernsehserien aufführt.<sup>7</sup> Seit Mitte der 90er Jahre nennt die deutschsprachige Dependance des Walt Disney-Filmvertriebs Buena Vista im Abspann der deutschen Verleihfassungen die beteiligten SynchronsprecherInnen; nicht nur bei den Animationsfilmen, sondern auch bei den (zumeist unter dem Label Touchstone oder Holly-



wood Pictures produzierten) *live-action*-Filmen.<sup>8</sup> Auffälliger für die RezipientInnen aktueller Kinofilme ist der in letzter Zeit verstärkte Einsatz von Stars als SynchronsprecherInnen, die zuvor wenig oder gar keine Erfahrung mit der Synchronisation hatten und deren Bekanntheit aus Auftritten als SchauspielerIn, SängerIn oder TV-ModeratorIn resultiert. Diese aus anderen medialen Zusammenhängen bekannten Stimmen werden in der Filmvermarktung beim *Mainstream*- wie beim Arthousefilm gleichermaßen als Werbefaktor eingesetzt.<sup>9</sup> Auch andere Medien wie das kommerzielle Hörspiel werden mit den deutschen Stimmen bekannter Hollywoodstars beworben.<sup>10</sup> Eine Stimmagentur weist auf ihre prominenten SprecherInnen (und deren Synchronitätigkeit) im Internet hin.<sup>11</sup> Seit 1997 verleiht der Bundesverband Kommunale Filmarbeit im Rahmen der Berlinale den *Liliput* genannten Preis für gelungene, aber vor allem auch für misslungene deutschsprachige Synchronisationsfassungen. Im Gegenzug veranstaltet die deutsche Synchronbranche seit 2001 eine *Gala der großen Stimmen*, in deren Rahmen der *Synchron*, ein Preis für herausragende Synchronisationsleistungen, verliehen wird.<sup>12</sup> All diesen unterschiedlichen Bemühungen ist gemeinsam, dass sie die Synchronstimme als Medium explizit thematisch werden lassen. Die vorgebliche Transparenz des Mediums Stimme wird nachdrücklich rückgängig gemacht, indem die Medienstimmen an Körper gekoppelt und mit Biographien, teilweise sogar Star-Images ausgestattet werden.

Aber auch die wissenschaftliche Forschung hat seit den 90er Jahren ein verstärktes Interesse an der Synchronisation gezeigt,<sup>13</sup> wobei die einzelnen Beiträge aus so unterschiedlichen Forschungszusammenhängen wie der Filmwissenschaft, der Linguistik, der Übersetzungswissenschaft, der Massenkommunikationsforschung und der Rhetorik stammen.<sup>14</sup> Dabei zeigt sich jedoch, dass die Stimmen der SynchronsprecherInnen, die im öffentlichen Diskurs im Mittelpunkt des Interesses stehen, von der Wissenschaft nur randständig zur Kenntnis genommen werden. So stellt Guido Marc Pruys seiner Dissertation *Die Rhetorik der Filmsynchronisation* eine Definition von Synchronisation voran, die neben dem üblicherweise als Synchronisation bezeichneten Verfahren des Austausches der gesprochenen Dialoge auf der Tonspur auch die Übersetzungsvariante ›Untertitelung‹ umfasst.<sup>15</sup> Diese – problematische – Begriffserweiterung erklärt sich durch die Arbeitsweise der traditionellen Synchronforschung: Sie vergleicht transkribierte Synchrondialoge mit Drehbuchdialogen in der Ausgangssprache. Dabei bleibt dasjenige am Rande der Aufmerksamkeit, was die Synchronisation von den anderen Filmübersetzungsverfahren abgrenzt. Von der Untertitelung etwa unterscheidet die Synchronisation ja gerade, dass deutschsprachige SprecherInnen den übersetzten Drehbuchdialog des Films in deutscher Sprache ein-

sprechen – die ›Original‹stimmen werden getilgt und gegen neue deutsche Stimmen ausgetauscht.<sup>16</sup> Die Darstellerkörper auf der Bildebene performieren dementsprechend in der deutsch synchronisierten Fassung ein unfreiwilliges *playback* zu den Synchronstimmen auf der Tonebene; in der Frühzeit der Synchronisation ist dies auch als Homunkulus-Effekt<sup>17</sup> beschrieben worden. Obwohl also gerade die neuen Stimmen der Synchronfassung im Mittelpunkt des populären Diskurses zur Film- und Fernsynchronisation stehen, wird der Stimmaustausch zwar in der Forschung konstatiert, in den Analysen aber so gut wie nicht berücksichtigt; vor allem der mediale Aspekt fand bisher keine Beachtung.<sup>18</sup> So geht z. B. der Anglist Thomas Herbst zwar auf den Stimmaustausch ein, jedoch ohne sein mediales Potenzial zu problematisieren:

[R]esearch has shown that there is absolutely no need for the voices of the original actor and the dubbing actor to be similar. Viewers tend to accept a particular voice as that of a given person, and apparently it is only in cases where one actor dubbing a particular character in a tv series has to be replaced by another for one reason or another that viewers complain that the new voice is unnatural or artificial. On the whole, if the casting of dubbing actors is done well, the dubbing voice of a particular character seems to appear as natural to a tv or cinema audience as the voice of the original actor to the audience seeing the film in the original. [...] voice casting in itself does not seem to present any great problems negatively affecting the quality of the translated film.<sup>19</sup>

Dass Herbst den Stimmaustausch als derart unproblematisch befindet, hängt auch mit dem sich im letzten Satz deutlich zeigenden, im Diskurs über die Filmsynchronisation vorherrschenden normativen Gestus zusammen. Dieser findet sich bereits in einem frühen polemischen Beitrag der Zeitschrift *Filmkritik*, der vom späteren Leiter des Filmmuseums München, Enno Patalas, verfasst wurde:

Immer noch herrschen weithin irriige Auffassungen von der Arbeit eines Synchronautors und -regisseurs. Immer noch dominiert die Ansicht, die Aufgabe eines solchen bestehe in nichts anderem als darin, die Dialoge des Originals so getreu wie möglich [...] in die eigene Sprache zu übertragen. Weit gefehlt! Die Arbeit eines Synchronisierers ist durchaus schöpferisch!<sup>20</sup>

Patalas ›meint‹, was er im Jahr 1963 sagt, durchaus ironisch: Er gehört zu jenen Filmkritikern und Cineasten, die Synchronisation als unerlaubten Eingriff in das künstlerische Werk des ›Filmschaffenden‹ sehen; sie verfälsche die vom *auteur* gewollte ›künstlerische Aussage‹ des Films. Wie aktuell diese Position auch heute noch ist, zeigte im Februar 2001 die Forderung des damaligen Staatsministers für Kultur und Medien Julian Nida-Rümelin nach der Ausstrahlung nicht-synchronisierter Filme im Fernsehen. Auch er führte das – topische – Argument ins Feld, Synchronisation beeinträchtige den Kunstgenuss; das deutsche Publikum müsse erzogen werden, ›Original‹filme zu sehen.<sup>21</sup> Andererseits spiegelt das Zitat von Patalas – insofern lässt sich die gesamte Bandbreite des Diskurses über die Filmsynchronisation an dieser Äußerung ablesen – aber auch das Selbstverständnis der in der Synchronbranche Beschäftigten wider: Die synchronisierte Filmfassung sei »im günstigsten Fall besser als das Original«,<sup>22</sup> oder zumindest, wenn sie gelinge, dem ›Original‹ ebenbürtig. Thomas Bräutigam formuliert:

Grundsätzlich kann es nicht darum gehen, den Originaldialog sklavisch zu adaptieren, entscheidend ist es vielmehr, eine ›innere Synchronität‹ herzustellen. Hierfür sind auch Änderungen erlaubt oder sogar zwingend erforderlich, wenn sie im Dienste des Originals stehen.<sup>23</sup>

Zugleich soll Synchronisation ›unauffällig sein‹, den Film also so übertragen, dass es nicht zu Unterbrechungen in der Narration, mithin einer ›Störung‹ der Rezeption kommt. Die Verständlichkeit, auf inhaltlicher wie akustischer Ebene, steht daher im Vordergrund. Es gehört zu den Standardverfahren der deutschen Synchronisation, dass der Synchrondialog lauter, bei Raumklang zudem mehr in den Vordergrund abgemischt wird und damit besser zu verstehen ist als dies in den ›ursprünglichen‹ Tonfassungen<sup>24</sup> der Fall ist. Diese Maxime der (›vorweggenommenen‹) Ent-Störung wird auch bei der Auswahl der SprecherInnen herangezogen: So wird das Stimmcasting bei noch nicht an StammsprecherInnen vergebenen SchauspielerInnen allein durch das Rollendesign der zu synchronisierenden Akteurinnen und Akteure bestimmt. Um nur ein prägnantes Beispiel zu zitieren: Dem Rollendesign eines als besonders viril markierten Protagonisten wie Thomas Magnum in der gleichnamigen Fernsehserie wird eine Stimme zugeordnet, die zur Vorstellung des Zuschauers vom zwei Meter großen, breitschultrigen Privatdetektiv Magnum passt. Diese Stimme ist die dunkle, brummige<sup>25</sup> Stimme des deutschen Synchronsprechers Norbert Langer, auch wenn sie in starkem Kontrast zur hellen, leicht piepsenden ›Original‹stimme des US-Schauspielers Tom Selleck steht. Damit wird die Figur des Thomas Magnum zudem in die Nähe einer an-

deren von Norbert Langer gesprochenen besonders virilen Stammrolle, Burt Reynolds, gerückt. Im Fall der Fernsehserie *MAGNUM, P. I.*<sup>26</sup> lässt sich nicht nur die Ausstattung des Schauspielerkörpers mit einer zum Rollendesign ›passenden‹ (ebenfalls besonders viril wirkenden) Stimme konstatieren. Auch eine andere Strategie von Synchronisation ist an der Erstsynchronisation der Serie abzulesen: Die Vorlage wird in den eigenen kulturellen Kontext transferiert und dabei dasjenige eliminiert, was den SynchronautorInnen als ›zu amerikanisch‹ und für das deutsche Publikum uninteressant zu sein schien – die Vietnamvergangenheit *Magnums*, die in der US-Serie keine geringe Rolle spielt.<sup>27</sup> Die Synchronisationsfassung wird also von dem ›befreit‹ oder ›entstört‹, was dem deutschen Publikum nach Einschätzung der SynchronautorInnen fremd anmuten könnte; sie wird an das deutsche Publikum umadressiert.<sup>28</sup>

Der Forschung ging es bislang nicht darum, solche kulturellen Verschiebungen und interkulturellen Verwerfungen zu beschreiben und in ihren Funktionsmechanismen zu analysieren. Synchronisationsforschung war bislang ebenso normativ ausgerichtet wie der populäre Diskurs: Man versuchte vor allem, Kriterien für eine ›gute‹ Synchronisation zu entwickeln, und man bemühte sich darum, einen systematischen Rahmen zu konzipieren, durch den Theorie und Praxis der Synchronisation ›verbessert‹ und vereinheitlicht werden sollten. Jedoch versteht, wie das Beispiel von Herbst gezeigt hat, eine derartig normative Ausrichtung, die ›gute‹ von ›schlechten‹ Synchronisationen unterscheiden will, den Blick auf die kulturelle Eigenständigkeit und Wirkmächtigkeit des synchronisierten Films. Daraus ergibt sich die Notwendigkeit, Synchronisation nicht aus normativer, sondern aus kulturvergleichend deskriptiver Perspektive zu betrachten.<sup>29</sup> Gerade der *cultural turn*, den auch die Übersetzungswissenschaft vollzogen hat, bietet Möglichkeiten, das Phänomen Synchronisation neu zu konzeptualisieren, die Synchronisation als Transposition, als hochkomplexen Vorgang zu begreifen, der die kulturelle Matrix eines Filmes verschiebt. So plädieren die ÜbersetzungswissenschaftlerInnen Bassnett und Lefevere für den Übergang von der textnahen Übersetzungskritik hin zur Untersuchung des kulturellen Bedingungsgeflechts der jeweiligen Übersetzung. Dabei wird jede Übersetzung als *re-writing*<sup>30</sup> aufgefasst, das immer auch Manipulation des Textes ist – eine Manipulation, die die Analyse des kulturellen Zusammenhangs erlaubt, in den (und aus dem) ›über-setzt‹ wird. Unter Einbezug dieser Konzepte der kulturwissenschaftlich informierten Übersetzungswissenschaft<sup>31</sup> kann Synchronisation als Transfer und Adaption kultureller Patterns, insbesondere auch von *gender*-Patterns, verstanden werden. Wichtig erscheint es daher, die durch den Stimmaustausch zwischen Ursprungsfassung und Synchronfassung differierenden medialen Effekte

in den Blick zu nehmen. Der Fokus der folgenden Ausführungen liegt dabei auf *gender*-Performanzen. Natürlich könnten die Verschiebungen qua Synchronisation auch an einem anderen Fokus untersucht werden; *gender* eignet sich für die Untersuchung aber deshalb besonders, weil gerade *gender* eine privilegierte Funktionsstelle darstellt, an der sich kulturelle Differenzen ablagern.<sup>32</sup>

Stimmpervormanz, das kann wohl als Befund angesehen werden, ist immer *gender*-spezifiziert. Neben der inhaltlichen Erfassung einer mündlichen Äußerung nimmt der Hörer bzw. die Hörerin immer auch eine Einordnung der Stimme nach ›männlich‹ und ›weiblich‹ vor. Die Stimme bietet demnach eine wichtige Markierungsinstanz für *gender*-Konnotationen. Stimmwahrnehmung ist immer auch Geschlechtswahrnehmung – und auch die Wahrnehmung des geäußerten Inhalts ist abhängig davon, ob man eine Frau oder einen Mann als Sprecherinstanz annimmt. Häufig beobachten lässt sich bei Synchronisationen, dass ›Original‹-stimmen, die in ihrer geschlechtlichen Semantisierung changieren (z. B. Jaye Davidson in *THE CRYING GAME*<sup>33</sup>), in der Synchronisation ›vereindeutigt‹, durch eine deutlich ›männlich‹ oder ›weiblich‹ konnotierte Stimme wiedergegeben werden. Die Synchronisation verfolgt offensichtlich die Tendenz, unklare *gender*-Semantisierungen, stimmliches *gender-crossing* zu ›begradigen‹.

Das hier vorgestellte Konzept der *gendered voices* erlaubt es, die Stimmpervormanz(en) dahingehend zu untersuchen, inwieweit und wie diese Konstruktion von *gender*-Zuweisungen via Sprechstimme erfolgt. *Gendered voice* bezeichnet demnach die Stimme als performativen Akt, da alle stimmlichen Äußerungen Träger von *gender*-spezifischen semantischen Effekten sind, unabhängig von der Frage, ob diese bewusst oder unbewusst erzeugt werden. Ebenso wie die Kategorie *gender* ist die immer geschlechtlich konnotierte Stimme nicht ›Natur‹, sondern ›Artefakt‹: Das ›Stimmmaterial‹ wird auf unterschiedlichste Weise ›modelliert‹ und so zu akustischer *gender*-Maskerade<sup>34</sup> eingesetzt. Filmwissenschaftlich kann in Bezug auf die zentrale Kategorie der Stimme an Überlegungen von Chion<sup>35</sup> angeschlossen werden, der sich mit den Implikationen von Funktions- und Wirkungsweise der ›Filmstimme‹ beschäftigt. Chion betrachtet das Verhältnis von Filmstimme und Filmkörper. Er fokussiert die Stimme als körperlose Entität in seinen Konzepten des *acousmètre* sowie des *anacousmètre*. *Acousmètre* ist dabei eine Stimme im Film, die nicht an einen sichtbaren Körper gebunden ist, aus dem *off* kommt. Mittels des *acousmètre* lässt sich die Synchronisation auch als eine Art Bauchrednertechnik auffassen, deren ›Handpuppen‹ die Schauspieler sind.<sup>36</sup> Dagegen bezeichnet der Begriff des *anacousmètre* die Verbindung von Körper und Stimme. Da die Stimme im Film aber nicht von der Person auf der Leinwand hervorgebracht wird, sondern vom zentralen Lautsprecher dahinter,

*sound track* und *image track* also keine Einheit bilden, ist jede Stimme im Film ein *acousmètre*. Dies ist von Gültigkeit, auch wenn moderne Tontechniken (*Surround*-Gestaltung etc.) zunehmend die Illusion des *anacousmètre* perfektionieren, dass die akustischen Effekte durch die SchauspielerInnen auf der Leinwand erzeugt würden. Insbesondere die Synchronfassung versucht, die Illusion des *anacousmètre* zu generieren, obgleich sie eine körperlose Stimme an einen (in diesem Fall sogar fremden) Körper koppelt, ohne dass diese Differenz intradiegetisch thematisiert wird. Dieser Zusammenschluss von Körper und nicht-körper-eigener Stimme stellt eine für eine derartige Untersuchung besonders relevante Konstellation dar. Bezieht man die Eigenschaft der Stimme als performativ-geschlechtsstiftende Instanz mit ein, können die Chionschen Konzepte dazu funktionalisiert werden, sowohl *gender*-Dekonstruktionen, *gender-crossings* als auch die Re-Konstruktionen hegemonialer *gender*-Konzepte qua *acousmètre* respektive *anacousmètre* beschreibbar zu machen.

Es geht also zusammengefasst darum, differenztheoretisch ›Original(- und Synchronfassung in den Blick zu nehmen und dabei die evozierten kulturellen Patterns mit dem Fokus auf *gender* einer genauen Analyse zu unterziehen. Eine solche Analyse ist – haben wir es doch beim Medium Film mit einem hybriden Medium zu tun – nur im Rahmen einer Gesamtbetrachtung von Bild- und Tonebene sinnvoll. Folgt man Michel Chions Definition von Film als ›Audio-Vision‹, also als untrennbare Einheit von Bild und Ton,<sup>37</sup> so führt der Austausch der Tonspur zur Eigenständigkeit des durch die Synchronisation entstandenen filmischen Artefakts; zugespitzt formuliert: Bei der Synchronisation eines Films entsteht durch den Austausch der Tonspur ein neuer Film.

Die folgenden Ausführungen zum amerikanischen Vietnamheimkehrerfilm *FIRST BLOOD*,<sup>38</sup> der in Deutschland unter dem Titel *RAMBO* zur Aufführung kam, sollen diese Überlegungen exemplifizieren. *RAMBO* ist eine der bekanntesten und erfolgreichsten Actionfilmreihen der 80er Jahre. Der Nachname des Protagonisten John J. Rambo schaffte es gar ins *Oxford English Dictionary* – und wird dort als Synonym für übertriebene, ideologisch fragwürdige Actionfilme und haudegenhaftes Benehmen geführt. Susan Jeffords hat in ihrer vielbeachteten Studie zum Actionkino der 80er Jahre<sup>39</sup> herausgearbeitet, welche Bedeutung dieser Figur, in der *gender*-Clichés und politische Konzepte konvergieren, im Kontext der damaligen Politik Reagans zuzumessen sei – so hat Reagan selbst mehrfach in seinen Reden auf die Filmfigur verwiesen. Nach der Befreiung amerikanischer Geiseln im Libanon 1985 äußerte er etwa: »Last night I saw *RAMBO* – Next time I'll know what to do.«<sup>40</sup> Mit *RAMBO* war dabei allerdings nicht der hier zu

besprechende Film *FIRST BLOOD* gemeint, sondern dessen Fortsetzung *RAMBO: FIRST BLOOD PART TWO*,<sup>41</sup> erst ab diesem zweiten Film darf Rambo auch im US-amerikanischen Raum den Titel schmücken. Jeffords differenziert nicht zwischen den drei Rambo-Filmen, sondern liest den ersten Film auf der Folie der Nachfolger. Dabei entgeht ihr aber, dass gerade dieser erste Teil signifikante Unterschiede zu seinen Fortsetzungen aufweist. Aus dem Blick gerät ihr die in diesem ersten Film besonders interessante Verhandlung des traumatisierten Vietnamveteranen. *FIRST BLOOD* wurde 1982 mit Sylvester Stallone in der Hauptrolle gedreht. Es handelt sich um einen Nachzügler jener Vietnamheimkehrerfilme, wie sie vor allem Anfang bis Mitte der 70er Jahre auf den amerikanischen Leinwänden zu sehen waren und in denen die zumeist psychotischen Kriegsheimkehrer als brutale Antagonisten zu *Bikern* und *Hippies* agierten. Im Unterschied zu diesen bemüht sich *FIRST BLOOD* aber um eine Psychologisierung des traumatisierten Protagonisten und nimmt damit Bezug auf die Ende der 70er Jahre entstandenen Vietnamfilme *COMING HOME*,<sup>42</sup> *THE DEER HUNTER*<sup>43</sup> und *APOCALYPSE NOW*.<sup>44</sup> *FIRST BLOOD* führt vor, wie der Vietnamkrieg in die USA transferiert wird, in einen kleinen Ort namens Hope im mittleren Westen. Der Vietnamveteran John Rambo wird unrechtmäßig als Landstreicher festgenommen und im Zellentrakt des Polizeireviere misshandelt. Dadurch werden in ihm traumatische Erinnerungen an Folterungen in vietnamesischen Gefangenenlagern wachgerufen. Rambo flieht; bei der anschließenden Verfolgungsjagd in einem Waldgebiet kann er die restlichen Polizisten ›kampfunfähig machen‹. Der Film reinszeniert nun zunehmend den Verlauf des Vietnamkriegs: Es kommt zu einer immer größeren Materialschlacht, bei der aber selbst die Nationalgarde und das eilig herbeigerufene Militär nichts gegen den Einzelkämpfer Rambo ausrichten können, der sich zunehmend der Guerilla-Kampftechniken der Vietcong bedient.

Der Vietnamfilm wird in den vielen filmwissenschaftlichen Studien, die zu diesem Genre im angloamerikanischen Raum veröffentlicht worden sind, als filmische Aufarbeitung der durch den Vietnamkrieg in der amerikanischen Gesellschaft verursachten Traumatisierung(en) verstanden. Aus diesem kulturellen Kontext heraus, die dem breiten deutschen Publikum der frühen 80er Jahre weitgehend unbekannt waren,<sup>45</sup> musste *FIRST BLOOD* nun – das Problem hatte die Synchronfassung zu lösen – in den deutschen Sprach- und Kulturraum verschoben werden; der Film sollte für ein deutsches Publikum ›verständlich‹ gemacht werden.

In der ersten Szene nach dem Vorspann (in dem wir den Protagonisten ein Stück seines langen Weges – *It's a long road* heißt denn auch der über den Ab-

spann gelegte Filmsong – literal bewältigen sehen) erkundigt sich der leicht verwahrlost aussehende Veteran John Rambo bei einer älteren schwarzen Frau nach einem Kriegskameraden.<sup>46</sup> Zwei Elemente der Synchronisation verweisen bereits in dieser ersten Szene, die symptomatisch für die starken inhaltlichen Eingriffe dieser Synchronfassung steht,<sup>47</sup> auf eine Entpolitisierung der Vorlage: Zum einen wird auf der inhaltlichen Ebene ein direkter Verweis auf das krebserregende Entlaubungsmittel Agent Orange und dessen Folgen eliminiert. Während in der ›Original‹ fassung der schwere (Krebs-)Tod des Vietnamveteranen Delmare eindrücklich geschildert und mit Agent Orange in Verbindung gebracht wird, lässt sich dieser Kontext dem deutlich weniger anklagend vorgebrachten »Er brachte den Krebs aus Vietnam mit« nur mit einigem Vorwissen entnehmen; das Sterben Delmares wird abgeschwächt und mit dem im ›Original‹ nicht vorhandenen Satz »Ich versuche, das alles zu vergessen« beendet. Dieser Satz erfüllt noch eine weitere Funktion, soll er doch dem deutschen Publikum eine Erklärung für das schroffe Benehmen der schwarzen Frau bieten. Deren Verhalten ergibt sich im ›Original‹ eher aus dem sozialen Konflikt zwischen den von Rambo und der afro-amerikanischen Frau vertretenen Ethnien und Schichten, die auch dementsprechend durch Soziolekte und dialektale Färbungen repräsentiert werden. Hier verweist der Film darauf, dass in Vietnam vor allem Soldaten aus den unteren sozialen Schichten, darunter überwiegend *african americans*, eingesetzt wurden. So spricht Rambo ein stockendes *working class*-Amerikanisch, das zusätzlich durch seine grammatischen wie Sprechfehler (Nuscheln, Hölzeln und leichtes Stottern) markiert ist. Der Sprachgebrauch der Frau ist hingegen deutlich *race*-konnotiert; sie wird über ihren Ethnolekt sowie ihre regional südstaatlerische Sprachfärbung den ärmsten schwarzen Schichten zugeordnet. Die Sprecher der deutschen Version hingegen sprechen beide akzentfreies, leicht monotones Hochdeutsch. Die Übertragung von Dialekten, Soziolekten und Ethnolekten kann als ein Grundproblem der Synchronisation beschrieben werden, da zumeist keine direkten kulturellen Äquivalente existieren. Sofern sie nicht für das Handlungsverständnis zwingend notwendig sind, werden sie in den Synchronfassungen häufig nicht berücksichtigt. Versucht wurde eine derartige Übertragung z. B. im Fall der deutschen Synchronisation von *MY FAIR LADY* (USA 1964, R.: George Cukor), wo der britische, stark *class*-akzentuierte Cockney-Akzent durch das eher regionale ›Berlinerisch‹ ersetzt wurde. Für die Adaptierung von ›Ghettoslang‹ wird in deutschen Synchronisationen die (supponierte) Sprechweise deutschsprachiger Türken eingesetzt. In beiden Fällen, bei der Übertragung ins Hochdeutsche (häufig unter Verwendung leichter ›Färbungen‹) oder bei der Suche nach Äquivalenten sind inhaltliche Verschiebungen unvermeidlich. Bei *First Blood* eliminiert die



deutsche Fassung dementsprechend die präzise soziale Verortung der handelnden Personen und die darin enthaltenen Verweise auf die vom Vietnamkrieg am meisten betroffenen Ethnien und Schichten.<sup>48</sup>

Diese Entpolitisierung wird in der deutschen Synchronfassung konsequent bis zum Schluss durchgeführt: Der nächste größere Dialogtext Rambos, seine den Film schließende flammende Rede, in der er vor allem der Friedensbewegung die Schuld am verlorenen Krieg gibt, wird ebenfalls entschärft. Hier wird die Emotionalität Rambos, der gerade einen Anfall männlicher Hysterie<sup>49</sup> durchlebt, in der deutschen Fassung so übersteigert, dass das, was Rambo letztlich sagt, kaum zu verstehen ist – aber auch in dem, was er sagt, werden direkte Verweise auf die Friedensbewegung entfernt. Durch derartige Auslassungen wird ein durchaus politischer (wenn auch rechtsideologischer), den Krieg oft thematisierender Film, von der in diesem Fall sehr deutlich Vorgaben der ›Original‹fassung verschiebenden Synchronisation in einen apolitischen Film umgewandelt. Der in der deutschen Fassung neu hinzugefügte Satz »Ich versuche, das alles zu vergessen« beschreibt demnach programmatisch, wie die Synchronisation vorgeht.

Unmittelbar verbunden mit dieser Strategie der Synchronisation, der Entpolitisierung, ist eine zweite Strategie, die ›Verheldung‹, die Heroisierung des (nur in der Synchronfassung titelgebenden) Protagonisten. Diese Heroisierung ermöglicht der deutschen Fassung des Films eine Maskulinisierung Rambos. Bereits der deutsche Titel RAMBO verschiebt die Emphase eindeutig auf die Titelfigur, während die ›Original‹tonfassung sich auch im Titel mit der Frage der Schuld an dem Massaker beschäftigt (*drawing first blood* entspricht dem deutschen Ausdruck des ›ersten Blutvergießens‹). Bemerkenswert ist die Besetzung der Stimme Rambos in der Synchronfassung. Während bis zu diesem Film der dem gänzlich anderen Schauspielstil des *method acting* verpflichtete, zu Improvisationen neigende Jürgen Prochnow die Rollen Stallones sprach, übernimmt mit RAMBO Thomas Danneberg die Synchronarbeit für Stallone. Der Rezipient bzw. die Rezipientin der Synchronfassung hört also eine Synchronstimme, die aus dem Actionfilm der frühen 80er bereits vertraut ist: fungierte Danneberg doch bereits als Synchronsprecher für Chuck Norris und Arnold Schwarzenegger. Obwohl Danneberg in dieser ersten Synchronarbeit Stallone etwas höher ansetzt, als er z. B. Schwarzenegger spricht (später wird er beide Schauspieler noch deutlich ähnlicher sprechen), liegt seine Stimme doch signifikant tiefer als die ›Original‹stimme.<sup>50</sup> Hier liegt also eine der eingangs beschriebenen Homogenisierungsstrategien vor: Synchronisationen versehen virile Körper bevorzugt mit viril wirkenden Stimmen (auch wenn die ›Original‹stimmen gerade nicht besonders ›viril‹ ausfallen). Allein durch diese Sprecherwahl rückt die Synchronfassung von FIRST BLOOD

also weiter in die Nähe der *hard bodies*, wie sie Jeffords vor allem für das Actiongenre der 80er Jahre beschreibt – und deutet auch bereits die weitere Entwicklung Sylvester Stallones zum Actionstar und Hauptvertreter dieses Genres an.

Die Wahl des Sprechers Danneberg für die Filmfigur John Rambo ließe sich als akustischer Kuleschow-Effekt<sup>51</sup> beschreiben. Wie schon herausgestellt entspricht die Stimme Dannebergs nicht Stallones ›Original‹-stimme: Der virile *hard body* Rambo wird in der ›Original‹-sprachfassung durch seine kieksenden Höhen und verschiedene Sprechfehler, die auch auf die Traumatisierung des Protagonisten und seine Charakterisierung als männlicher Hysteriker verweisen, deutlich konterkariert. Dannebergs Synchronstimme wirkt maskuliner, sonorer, sympathischer als Stallones Stimme – und erleichtert damit dem Zuschauer auch die Identifikation mit der Hauptfigur. Im Gegenzug wirkt der ermittelnde Sheriff im Deutschen rauer und unfreundlicher. Dass Rambo in der deutschen Fassung deutlich ›heroischer‹, ›vorbildhafter‹ daherkommt, dafür gibt es viele Belege. So bewundert in der Synchronfassung einer der Polizisten Rambo für seinen Mut und Einfallsreichtum, während er im ›Original‹ seine Furcht vor dem psychopathischen Vietnamheimkehrer ausdrückt. Die ›Verheldung‹ wird vollendet durch die veränderte Intonation und das differierende Sprachregister der Worte von Colonel Trautman: Statt vor Rambo im militärischen Sprachregister zu warnen, sein Trauma plastisch darzustellen (»His job was to get rid of enemy personnel, to kill them. Winning by attrition! And Rambo was the best!«), generiert im Deutschen die deutlich weniger militärisch und viel eher väterlich-stolz agierende Stimme Friedrich W. Bauschultes einen Heldengesang: »In Vietnam war es seine Aufgabe, Feinde umzubringen, zu töten, brutal! Sieg durch Zerstörung. Und Rambo war der Beste!« Diese in der deutschen Synchronfassung veränderte Anlage der Hauptfigur nimmt die amerikanische Weiterentwicklung der Rambofilme vorweg: Es handelt sich um einen interkulturellen Rückkoppelungseffekt. Dieser Rückkoppelungseffekt kann zumindest teilweise auch produktionsästhetisch begründet werden. *FIRST BLOOD* nimmt insofern filmhistorisch eine besondere Rolle ein, als es einer der ersten Filme war, die von nicht in Hollywood ansässigen Produzenten in den USA (Carolco) produziert wurde, wobei die Finanzierung durch Kreditverträge im Ausland und vor allem durch Vorverkäufe erzielt wurde. Der Erfolg des Films im Ausland war daher für die Produzenten genauso wichtig wie das Inlandseinspiel. *FIRST BLOOD* war nun in den Ländern (v. a. Deutschland und Italien), in denen der Name des Protagonisten titelgebend war, noch einmal deutlich erfolgreicher als in den übrigen Aufführungsorten. Diese Situation wird bei der Titelgebung und der inhaltlichen Gestaltung der Sequels sicher eine Rolle gespielt haben.

Die US-amerikanischen Rambo-Sequels, die auf FIRST BLOOD folgen, entwerfen keinen traumatisierten Protagonisten mehr, sondern übernehmen die Heroisierungstendenz der deutschen Synchronfassung von FIRST BLOOD. John Rambo wird in RAMBO: FIRST BLOOD PART TWO aus dem Jahr 1985 zum eindeutig positiven Protagonisten und zur Titelfigur, der den Vietnamkrieg nachträglich für sich entscheidet. Der dritte Teil (RAMBO III)<sup>52</sup> und die Zeichentrickserie für Kinder (RAMBO)<sup>53</sup> verzichten auf Reminiszenzen an den Titel des ersten Teils, sondern tragen dann nur noch den Namen des Helden, heißen nur noch RAMBO.

Zusammenfassend lassen sich also zwei Strategien ausmachen, wie der Film aus seinen US-amerikanischen Kontexten heraus für das deutsche Publikum adaptiert und umadressiert wird: Die deutsche Synchronisation entpolitisiert den Film, schwächt die (vor allem rechts-)ideologischen Bezüge auf den Vietnamkrieg ab oder eliminiert sie. Damit in engem Zusammenhang steht die die Männlichkeitsinszenierung verändernde, über die Synchronstimme operierende Strategie der Virilisierung, der Heroisierung des Protagonisten. Viele dieser Effekte werden durch die Stimmumbesetzung und daraus resultierende veränderte *gender*-Performanzen erzielt. Die Fokussierung auf die *gender*-Ebene erlaubt demnach einen genauen Einblick in Funktionsmechanismen des Stimmaustauschs in Synchronfassungen, wobei diese Funktionsmechanismen an jedem Einzelbeispiel erneut konturiert werden müssen.

Für Anregungen und Kritik zu diesem Beitrag bedanke ich mich bei Claudia Liebrand.

- 1 Neben der Filmsynchronisation werden überwiegend Untertitelung (z. B. in den englischsprachigen Ländern) und *voice over* eingesetzt.
- 2 Der Begriff des ›Originals‹ wird auch im Folgenden in Anführungszeichen gesetzt: Ausgegangen wird davon, dass erst eine Synchronisation die Filmtönung der Uraufführung im Ursprungsland zu einem ›Original‹ machen kann. Diese wird im Folgenden als ›Original‹ fassung bezeichnet. Wie problematisch es um den Begriff einer ›Original‹ tonspur bestellt ist, zeigt auch das Beispiel des italienischen Kinos: Dort werden nahezu alle Produktionen erst im Tonstudio mit einer Tonspur ausgestattet; dabei erhalten häufig die italienischen DarstellerInnen fremde Stimmen zugeordnet. Es ist aber dennoch wichtig, den Einsatz der Synchronisation als Filmübersetzungsverfahren zunächst abzugrenzen von dem Filmproduktionsverfahren des Stimmaustauschs in der *post production*. Beide Verfahren – darin zeigt sich ihre technische Verwandtschaft – werden im angloamerikanischen Raum als *dubbing* bezeichnet.
- 3 Auswertung nach Lothar Just (Hg.): Heyne Film-Jahrbuch 1999, München 1999.
- 4 In dieser Übergangsphase kam z. B. auch die Sprachversion zum Einsatz; vgl. Ginette Vincendeau: Hollywood Babel. The Coming of Sound and the Multiple-Language Version, in: Screen 29/2 (1988), S. 24–39.
- 5 Wie gebräuchlich diese Bezeichnung in Branchenkreisen ist, zeigt auch der Internetauftritt des Synchronstudios Johannistal Synchron (vgl. [www.johannisthalsynchron.de/deu\\_syn.htm](http://www.johannisthalsynchron.de/deu_syn.htm); letzte Abfrage am 28.02.2003).
- 6 Wie wichtig diese ›Beigabe‹ für die Etablierung der DVD als Konkurrenz zur herkömmlichen VHS-Videocassette ist, zeigt sich auch in den geringen Verkaufszahlen solcher DVDs, die keine zusätzliche ›Original‹ tonspur zur Verfügung stellen.

- 7 Thomas Bräutigam: Lexikon der Film- und Fernsynchronisation, Berlin 2001.
- 8 Seit den späten 60er Jahren ist es unüblich geworden, die Filmvorspanne fremdsprachiger Filme in der deutschen Fassung komplett neu zu gestalten; zumeist wird lediglich der deutsche Titel anstelle des oder zusätzlich zum ›Original‹-titel eingeblendet. Eine Ausnahme stellen jedoch für das deutsche öffentlich-rechtliche Fernsehen hergestellte Synchronisationen sowie DEFA-Synchronisationen dar, die jeweils den kompletten Vor- und Abspann austauschen und dabei die SynchronsprecherInnen aufführen.
- 9 Diese Verfahrensweise war zunächst im Bereich des Animationsfilms üblich, wird aber zunehmend auch in anderen Filmen, vor allem im Genre der Komödie, eingesetzt. So wurde bereits 1989 die deutsche Synchronisation des Babys in LOOK WHO'S TALKING (USA 1989, R.: Amy Heckerling) durch Thomas Gottschalk in der Filmvermarktung prominent herausgestellt. Während im Jahr 2002 der Film ICE AGE (USA 2002, R.: Carlos Saldanha, Chris Wedge) in Deutschland sehr stark über den Comedian Otto Waalkes vermarktet wurde, bewarb der deutsche Verleih des französischen Arthaus-Musicals HUIT FEMMES (F 2002, R.: François Ozon) die deutsche Kinoauswertung mit den sechzehn Schauspielerinnen der Synchronfassung – zur Synchronisation der acht französischen Stars (u. a. Catherine Deneuve und Emmanuelle Béart) wurden acht deutsche Stars (u. a. Katja Riemann und Senta Berger) engagiert. Im deutschen Kinovorspann wurden die Synchronsprecherinnen gleichrangig mit den französischen Schauspielerinnen benannt.
- 10 So wirbt der Verlag WortArt für seine Hörspielreihe »John Sinclair Edition 2000« damit, dass die Hauptfigur von ›der deutschen Stimme von James Bond‹ gesprochen werde. Auch bei Christian Brückners Hörbüchern fehlt selten der Verweis darauf, dass er die Synchronstimme Robert de Niro ist.
- 11 Die Agentur findet sich online unter [www.stimmgerecht.de](http://www.stimmgerecht.de) (letzte Abfrage am 28.02.2003).
- 12 Zu den bisherigen PreisträgerInnen zählen Christian Brückner und Corinna Harfouch; gerade in dieser Auswahl wird der Werbecharakter dieser Veranstaltung für die Synchronbranche besonders deutlich.
- 13 Das Anfang 2003 erschienene Themenheft der Filmzeitschrift »Der Schnitt« zur Filmsynchronisation (Der Schnitt, Heft 29, 2003) bietet einen interessanten Schnittpunkt zwischen wissenschaftlichem und populär geführtem Diskurs.
- 14 Die deutliche Zunahme der Synchronisationstätigkeit seit den 80er Jahren durch den stark vergrößerten Video- und Fernsehmarkt spielt hier sicher ebenso eine Rolle wie die wissenschaftstheoretisch breit wirksame Erweiterung des Textbegriffs durch die *Cultural Studies*.
- 15 Guido Marc Pruys: Die Rhetorik der Filmsynchronisation. Wie ausländische Spielfilme in Deutschland zensiert, verändert und gesehen werden, Tübingen 1998. Pruys bietet nicht die erste problematische Definition, die die Synchronisation in der Forschung bislang erfahrene hat. So geht die viel zitierte Definition von Fodor davon aus, dass neben der neu aufgenommenen Dialogtonspur eine isolierte Tonspur mit Musik und Geräuschen vorliegt: »This new recording of the sound, apart from the original background noises and the film music, has to be matched with the picture in a way that the new speech sounds coincide with the visible movements of the articulatory organs as perfectly as feasible.« (István Fodor: Film Dubbing. Phonetic, Semiotic, Esthetic and Psychological Aspects, Hamburg 1976, S. 9.) Gerade dieses IT-Band (*International Tape*) liegt jedoch häufig bei älteren Filmen, aber auch z. B. bei Dogma-Filmen nicht vor. Selbst bei vorhandenem IT-Band wird dieses jedoch nicht in jedem Fall vom Synchronstudio auch verwendet, da dort häufig eigene Präferenzen über die Tonabmischung entscheiden. In der Forschung wird darüber hinaus zu selten nach den Medien unterschieden, für die eine Synchronfassung erstellt wird. Film- und Fernsynchronisationen unterliegen differierenden Produktionsbedingungen, die in der Einzelanalyse zu berücksichtigen sind. Im vorliegenden Beitrag steht die Filmsynchronisation für den Kinoeinsatz im Mittelpunkt, es wird aber auch ein Beispiel aus der Fernsynchronisation zitiert.
- 16 Dabei wird auch die Tonabmischung, somit letztlich die gesamte Tonspur verändert.
- 17 Das sog. ›Homunkulus-Phänomen‹ bezieht sich auf den Zusammenschluss von Körper und nicht-körpereigener Stimme und stellt damit eine besonders relevante Konstellation für die Untersuchung von Synchronisationen dar.  
Vgl. Joseph Garncarz: Die bedrohte Internationalität des Films. Fremdsprachige Versionen deutscher Tonfilme, in: Sibylle M. Sturm/Arthur Wohlgemuth (Hg.): ›Hallo? Berlin? Ici Paris!‹ Deutsch-französische Filmbeziehungen 1918–1939, München 1996, S. 127–140 (hier: S. 132).
- 18 Diese Nichtbeachtung von Synchronisationseffekten, die geringe Beachtung des Tons überhaupt, hat ihre Wurzeln auch in der traditionell nahezu ausschließlich visuell orientierten Filmanalyse.

- 19 Thomas Herbst: Dubbing and the Dubbed Text. Style and Cohesion, in: Anna Trosborg (Hg.): Text Typology and Translation, Amsterdam/Philadelphia 1997, S. 291–308 (hier: S. 291 f.).
- 20 Enno Patalas: Schöpferische Synchronisation, in: Filmkritik 7/11 (1963), S. 510–512 (hier: S. 510). Auch Pruys verweist in seiner Einleitung auf diesen Beitrag.
- 21 Die Äußerung erfolgte in der TV-Sendung MAISCHBERGER des Senders n-tv am 26.02.2001.
- 22 Diese Auffassung verbreitete Christian Brückner im Rahmen einer von ihm moderierten und mitgestalteten Aufführung von Filmmusik in der Kölner Philharmonie im Oktober 2002. Er trug dabei die von ihm in der Synchronfassung von TAXI DRIVER (USA 1976, R.: Martin Scorsese) gesprochenen Synchrondialoge zur vom Orchester live eingespielten Filmmusik vor.
- 23 Bräutigam: Lexikon (Anm. 7), S. 33.
- 24 Dabei darf nicht vergessen werden, dass auch diese vermeintlich »ursprünglichen« Tonfassungen zumeist ihrerseits wieder (unter Verwendung des Aufnahmetons) im Tonstudio abgemischt worden sind. Hier liegen dementsprechend kulturell differente Tonabmischungspraxen vor.
- 25 Bei der Beschreibung und Analyse von Synchrononspur und »Original«-tonspur kann unter anderem auf Kategorien der Phoniatrie zurückgegriffen werden. Stimmen werden dort als Komposita aus Stimm- und Sprechqualitäten betrachtet. Qualitäten, die nicht eindeutig einer der beiden Kategorien zugeordnet werden können, werden unter dem Begriff der Prosodie behandelt. Dabei ist zu beachten, dass alle Stimm- und Sprechqualitäten erst im Zusammenspiel die (veränderliche) Stimme ausmachen. Die Verwendung phoniatischer Stimmbeschreibungen würde zudem das Erstellen von Tonfilmsequenzprotokollen ermöglichen, anhand derer sich die Unterschiede zwischen beiden Tonfassungen aufzeigen lassen. Die Beschreibungskategorien, die in diesem Beitrag für die Stimmbeschreibung eingesetzt werden, sind zum einen durch diese phoniatischen Begriffe bestimmt, müssen aber auch das Schauspiel der Original- und SynchronsprecherInnen beschreiben. So lässt sich die Stimmlage relativ unproblematisch bestimmen, sagt jedoch für sich alleine noch nicht viel darüber aus, wie die Stimme im Einzelfall eingesetzt wird. Eine hermeneutische Beschreibung der vermittelten Emotionen lässt sich daher nicht umgehen. Eine exakte Umschrift von Stimmen ist, wie jegliche Verschriftlichung von Ton, nicht möglich; es kann daher immer nur um eine Annäherung gehen, die im folgenden häufig dadurch erleichtert wird, dass zwei Stimmen miteinander verglichen werden, die auf das gleiche Bild referieren.
- 26 Dt. Titel MAGNUM, USA 1980–1988.
- 27 Zum Verhältnis zwischen der ersten Synchronfassung, die von der ARD in Auftrag gegeben wurde, und der von RTL als (gescheitertes) Prestigeobjekt konzipierten Neusynchronisation vgl. Katrin Wehn: Die deutschen Synchronisation(en) von Magnum, P. I., unter: [www.medienkomm.uni-halle.de/forschung/publikationen/halma2.pdf](http://www.medienkomm.uni-halle.de/forschung/publikationen/halma2.pdf) (letzte Abfrage am 28.02.2003).
- 28 Zum Begriff der Adressierung vgl. Stefan Andriopoulos/Gabriele Schabacher/Eckhard Schumacher (Hg.): Die Adresse des Mediums, Mediologie Bd. 2, Köln 2001.
- 29 Hier lässt sich teilweise an Überlegungen von Niemeier anschließen: »[...] it is especially the context conditions which are important in film synchronisation, because it is – apart from the influence of different sign vehicles – particularly the cultural context conditions of the sign receivers which may contribute intercultural misunderstandings in the reception of films.« (Susanne Niemeier: Intercultural Dimensions of Pragmatics in Film Synchronisation, in: Jan Blommaert/Jeff Verschueren (Hg.): The Pragmatics of Intercultural and International Communication, Amsterdam/Philadelphia 1991, S. 145–162 [hier: S. 147 f.].) Wichtig ist allerdings, durch differierende kulturelle Kontexte bedingte Veränderungen in der Rezeption von Filmen nicht als »misunderstandings« zu konzipieren, sondern als kulturell eigenständige Lesarten des Films.
- 30 Im Vorwort eines von ihnen herausgegebenen Sammelbandes, der sich bemüht, die Auswirkungen des *cultural turn* auf die Übersetzungswissenschaften nachzuzeichnen, heißt es programmatisch: »Translation is a re-writing of an original text.« (Susan Bassnett/André Lefevere: Preface, in: dies. (Hg.): Translation, History and Culture, London 1999, S. IX).
- 31 Auch in der Übersetzungswissenschaft verschiebt sich der Fokus vom Ideal einer exakten Übertragung hin zur Untersuchung kultureller Adaptionsmechanismen: »In most cases, equivalence can hardly be obtained in translation across cultures and languages, and it may not even be a desirable goal.« (Anna Trosborg: Introduction, in: dies. (Hg.): Text Typology and Translation, Amsterdam/Philadelphia 1997, S. VII).
- 32 Dieser Gedanke wird insbesondere ausgeführt in der Einleitung zu Claudia Liebrand: Gender-Topographien. Kulturwissenschaftliche Lektüren von Hollywoodfilmen der Jahrhundertwende, Köln 2003.

- 33 GB 1992, R.: Neil Jordan.
- 34 Angeschlossen werden kann hier an das insbesondere für die feministische Filmwissenschaft zentrale Rivieresche Konzept der Weiblichkeitsmaskerade, dass von Liebrand auch für die Untersuchung von Männlichkeitsrepräsentationen fruchtbar gemacht wurde; vgl. Claudia Liebrand: Prolegomena zu Cross-dressing und Maskerade. Zu Konzepten Joan Rivieres, Judith Butlers und Marjorie Garbers – mit einem Seitenblick auf David Cronenbergs Film *M. Butterfly*, in: *Freiburger Frauen-Studien* 5, Bd. *Cross-dressing und Maskerade* (1999), S. 17–31.
- 35 Michel Chion: *The Voice in Cinema* [1984], New York 1999; vgl. auch den Beitrag in diesem Band.
- 36 Dieses Konzept wurde bereits von Ascheid für das Synchronisationsverfahren fruchtbar gemacht; sie verweist dabei auf Altman. Vgl. Antje Ascheid: *Speaking Tongues. Voice Dubbing in the Cinema as Cultural Ventriloquism*, in: *The Velvet Light Trap* 40 (1997), S. 32–41, sowie Rick Altman: *Moving Lips. Cinema as Ventriloquism*, in: ders. (Hg.): *Cinema/Sound*, *French Yale Studies* 60 (1980), (Sonderausgabe), S. 67–79. Interessant ist hier die Parallele zur eingangs bereits genannten Beschreibung der Filmsynchronisation als Homunkuluseffekt.
- 37 »A film's aural elements are not received as an autonomous unit. They are immediately analysed and distributed in the spectator's perceptual apparatus according to the relation each bears to what the spectator sees at the time. [...] These distinctly different triages of sounds emitted from the single real source of the loudspeaker, triages based on the simple criterion of each sound's relation to each image at each moment, already testify sufficiently that there is no soundtrack, to put it provocatively.« (Chion: *Voice* [Anm. 35], S. 3).
- 38 USA 1982, R.: Ted Kotcheff. Für die Analyse wurde die am 31.10.2000 erschienene deutsche DVD der Firma Kinowelt verwendet, die die zeitgenössische deutsche, für die Kinoauswertung erstellte Synchronfassung aus dem Jahr 1982 enthält. Zur Kontrolle wurde die amerikanische Special Edition *RAMBO TRILOGY* von Artisan (erschienen am 28.05.2002) herangezogen; der Vergleich ergab keine Kürzungen bei der deutschen DVD.
- 39 Susan Jeffords: *Hard Bodies. Hollywood Masculinity in the Reagan Era*, New Brunswick 1994.
- 40 Zitiert nach der Dokumentation *GUTS AND GLORY*, die auf der amerikanischen DVD-Edition *RAMBO TRILOGY* enthalten ist.  
Wie wirkmächtig der Name Rambo auch 13 Jahre nach dem letzten Kinofilm noch ist, zeigt die Titelseite des Nachrichtenmagazins »Der Spiegel« vom 01.03.2003: Ein Artikel über die Special Operations Group des CIA wird beworben unter dem Titel »Operation ›Rambo‹ – Die geheimen Spezialtruppen der USA«. Auch im Artikel selbst wird des Öfteren Rekurs genommen auf die Filmfigur John Rambo.
- 41 Dt. Titel: *RAMBO II*, USA 1985, R.: George Pan Cosmatos.
- 42 Dt. Titel: *COMING HOME – SIE KEHREN HEIM*, USA 1978, R.: Hal Ashby.
- 43 Dt. Titel: *DIE DURCH DIE HÖLLE GEHEN*, USA 1978, R.: Michael Cimino.
- 44 USA 1979, R.: Francis Ford Coppola.
- 45 So erschien die erste umfassende geschichtliche Darstellung des Vietnamkriegs in deutscher Sprache erst im Jahr 2000 (Marc Frey: *Geschichte des Vietnamkriegs*, München 2000).
- 46 Es sei erwähnt, dass dies fast der einzige längere Dialogtext ist, den das Drehbuch für John Rambo vorsieht. Nach dieser einleitenden Sequenz wird er bis zum Finale überaus wortkarg in Szene gesetzt, eine Inszenierungsform, die in der feministischen Filmforschung häufig als Ausdruck männlicher Hysterie analysiert worden ist.
- 47 Die deutsche Synchronfassung von *FIRST BLOOD* ist damit ebenfalls ein gutes Beispiel gegen eine in der Forschung oftmals vorgenommene Historisierung, die die starken Eingriffe deutscher Synchronfassungen allein den 50er Jahren zuschreibt. Ein aktuelleres Beispiel für eine derartige Veränderung bietet Verhoevens *STARSHIP TROOPERS* (USA 1997, R.: Paul Verhoeven): Hier wurde u. a. der Verweis auf die Zerstörung Hiroshimas in eine fiktive ›Zerstörung Washingtons im ersten Bug-Krieg‹ geändert. Die Synchronisation versucht hier wie andernorts, den Angriffskrieg (dem eben kein erster Bug-Krieg vorangegangen ist) der ›Starship Troopers‹, die auf der Bildebene deutlich als faschistoid markiert sind, in einen Verteidigungskrieg umzuwandeln (entsprechend der bekannten nationalsozialistischen Parole, dass ›zurückgeschossen‹ würde). Die den Faschismus des Geschehens satirisch entlarvende (und damit die Bildspur konterkarierende) ›Original‹-tonspur verwandelt durch (wahrscheinlich durch *political correctness* motivierte) Veränderungen den Film in einen weitaus faschistoideren Film. Interessant ist die Frage, ob diese Verfahrensweise der Synchronisation die Indizierung des Films in Deutschland unterstützt hat.
- 48 Auch wird die Frau in der deutschen Fassung durch ihre deutlich ältere Stimme stärker in eine

- Mutterrolle gerückt, während man in der ›Original‹-fassung die genaue verwandtschaftliche Beziehung zu Delmare nicht bestimmen kann.
- 49 Das Konzept der männlichen Hysterie ist in der psychoanalytisch informierten Filmwissenschaft seit einigen Jahren zu einem Schlüsselbegriff bei der Analyse der Inszenierungsformen von Männlichkeit, vor allem innerhalb des Actiongenres, geworden. Zu den Inszenierungsformen männlicher Hysterie, wie sie u. a. von Tasker (Yvonne Tasker: *Spectacular Bodies. Gender, genre and the action cinema*, London/New York 1993) beschrieben worden sind, gehören Sprachlosigkeit und irrationales Handeln. Einen guten Überblick über diese Theorie-debatte bietet Pentti Haddington: ›This is not going to have a happy ending‹. Searching for New Representations of Hollywood Masculinities in David Fincher's *Seven*. Pro Graduate Thesis, University of Oulu 1998. Online unter: [cc.oulu.fi/~poh/thesis/se7en.pdf](http://cc.oulu.fi/~poh/thesis/se7en.pdf); letzte Abfrage am 22.01.2003.
- 50 Prochnows Stimme ähnelt in Stimmlage und Rhythmus deutlich mehr Stallones ›Original‹-stimme als Dannebergs, der zudem einer signifikant differenten, der deutschen Theatertradition verhafteten Schauspiel- (und damit auch Sprech-)tradition angehört.
- 51 Analog zum filmischen Montagebegriff des Kuleschow-Effekts, der besagt, dass vorhergehende und nachfolgende Einstellungen die Wirkung eines Bildes auf den Zuschauer mitbestimmen, kann man auch die Wirkung veränderten Tons als einen akustischen Kuleschow-Effekt definieren: Die Verbindung des neuen Tons mit dem Filmbild erzeugt neue semantische Effekte. Synchronisation kann im Extremfall das Produkt so verändern, dass selbst in der Vermarktung nicht mehr von demselben Film gesprochen wird. Bekanntes Beispiel hierfür ist der japanische Agentenfilm *KOKUSAI HIMITSU KEISATSU. KAGI NO KAGI* (Japan 1964, R: Senkichi Taniguchi), der für den US-amerikanischen Markt von Woody Allen synchronisiert werden sollte, was dieser dazu nutzte, ihn in eine Komödie zu verwandeln (*WHAT'S UP, TIGER LILY?* [USA 1966]). Die deutsche Synchronisation (also die Synchronisation der Synchronisation) wiederum bemüht sich später, den Film in die Reihe der erfolgreichen Woody-Allen-Komödien zu stellen.
- 52 USA 1987, R.: Peter MacDonald.
- 53 Dt. Titel: *RAMBO – DIE SERIE*, USA 1986.

Ines Steiner

THE VOICE BEHIND THE CURTAIN.

ZUR INSZENIERUNG DER STIMME IN »SINGIN' IN THE RAIN«

### 1. GENRE, GENDER, MEDIENREFLEXION

In dem 1952 von MGM produzierten Filmmusical SINGIN' IN THE RAIN wird die Konstruktion der Sprech- und Singstimme im Mittel eines gleichermaßen nostalgischen wie ironischen Rückblicks auf jene Zeit thematisiert, in der die bewegten Bilder erstmals singen, sprechen, (und damit auch) tanzen lernten. SINGIN' IN THE RAIN betreibt so filmisch die Geschichtsschreibung der für den Film entscheidenden Jahre 1927 bis 1929; das Musical erörtert selbstreflexiv den Medienumbruch vom elaborierten Kode des Speechless Film,<sup>1</sup> der die menschliche Stimme durch medienspezifische Surrogate auf der Bildebene substituiert hatte, zu den (technisch zunächst kruden) Verfahren des frühen Tonfilms, die nun unverhofft nicht allein das theatralische Mittel des Dialogs, sondern vor allem auch Musik und Gesang analog zu reproduzieren vermochten. Dieser Medienumbruch wird unter dem Stichwort der »revolution« verhandelt – schon deshalb mit Recht, weil SINGIN' IN THE RAIN hiermit auf die Bedingung seiner Möglichkeit verweist. Das Musical stellt – »All Talking! All Singing! All Dancing!« den einzigen vom Tonfilm initiierten Neuzugang unter den wichtigen Filmgenres dar: *Das* hatte es vorher in der Tat noch nicht gegeben.

SINGIN' IN THE RAIN, betreibt derart Medienarchäologie als Erforschung des eigenen Ursprungs. Entstanden ist dabei »[an] exuberant and malicious satire of Hollywood in the late twenties«, der es an kritischer Schärfe nicht fehlt.<sup>2</sup>

Das Autoren-Team Betty Comden und Adolph Green verpackt seine detaillierten Kenntnisse des Zeitmilieus in eine an pointierten Anspielungen und geschickten intertextuellen Verweisen reiche Behind-the-Screen-Story, welche die historisierende Zitation der berühmten Songs von Arthur Freed und Herb Nacio Brown in beeindruckenden Gesangs- und Tanznummern intradiegetisch motiviert.<sup>3</sup> Regie führte Stanley Donen in Kooperation mit Gene Kelly, der als Choreograph auch die tänzerischen Bewegungsabläufe in den *production numbers* entwarf und einstudierte. Jean Hagen stattet die Stummfilmdiva Lina Lamont mit einem attraktiven Körper und einer schauerlich schrillen Stimme aus; die damals siebzehnjährige Debby Reynolds verkörpert das aufstrebende Starlet, das zum Stimmdouble der Diva wird; Gene Kelly brilliert in der Rolle Don Lockwoods, der als Tänzer und Sänger den Übergang zum Tonfilm mühelos zu meistern ver-



steht; ihm steht als Supporting Actor Donald O'Connor als der Mood Musician Cosmo Brown zur Seite.

Aus dem Ensemble technischer, ästhetischer und kommerzieller Probleme, die mit der Umstellung der Branche auf den Tonfilm einhergingen, gewinnt SINGIN' IN THE RAIN zahlreiche komische Effekte – nicht allein bezüglich der Filmproduktion, sondern auch im Hinblick auf Distribution und vor allem die Fabrikation von Stars und die mit ihnen assoziierten Körperbilder. So wird der Topos von der »geraubten Stimme« (der deutlich an Märchenmotive anknüpft) nicht etwa Thema eines Melodrams,<sup>4</sup> sondern im Rahmen einer Musical Comedy mit brillanten Production Numbers, pointierten Dialogen und präzise konstruierten Gags verhandelt. Vielfach variierte Legenden Hollywoods über den Fall etablierter Stummfilmstars, die in Ermangelung einer adäquaten Sprech- und Singstimme die neuen Anforderungen an das Schauspiel vor dem Apparat nicht mehr bedienen konnten und damit den durch die filmtechnische Innovation bedingten veränderten Produktionsmethoden zum Opfer fielen, werden explizit zitiert.

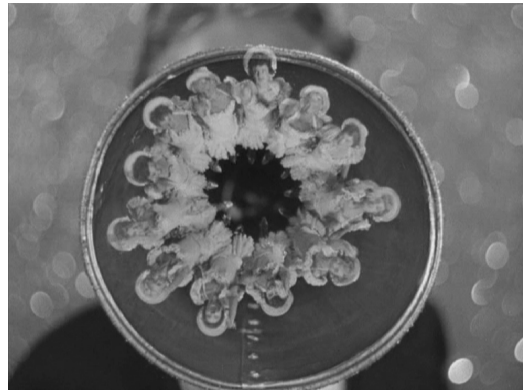
Open season had officially started, with no rules. The frenzy and clamor made it immensely convenient for most of the studios to plow ahead without looking. So many factors to heed, so little time, so many shortcuts and, as ought to be expected, such frightfully mixed results. Often as not, the final product made it clear that the new and unfamiliar demands led film to be conceived in panic and manufactured in ignorance. In the beginning, about all they had was enthusiasm and microphones, backed up a modicum of intuition and an avid yen for profits. Every other component, personnel and material alike, had to be built or bought or hauled in or finessed or fudged.<sup>5</sup>

Die hier von Rick Barrios beschriebene bewegte, chaotische Interimszeit zwischen dem unerwarteten Erfolg von Alan Crosslands THE JAZZ SINGER (1927) und der endgültigen Durchsetzung des Tonfilms 1929 komprimiert SINGIN' IN THE RAIN schlaglichtartig in Form einer rasanten, rhythmisch virtuos komponierten Montagesequenz, die den Effekt einer »Revolution in Hollywood« freisetzt. Rotierende Schlagzeilen, »Studios convert to talkies«, aus der *Variety* mit den wichtigsten Informationen über die filmhistorische Entwicklung werden als »burning news« ins Bild gesetzt, die den Blick auf einen »Flächenbrand«, in dem sich das neue Genre Musical ausbreitet, erzeugen. So eröffnen sich unter der verbrennenden Schrift palimpsestartig durch diverse Blendenformate, die den filmtechnischen Stil der Zeit imitieren, initiierte kurze Film-Sequenzen, Clips, die

den Musical Boom der Jahre als ›Explosion‹ ins Bild setzen, und zwar in Form einer rhythmischen Kombination ausschnittshafter Nachinszenierungen von bekannten tänzerischen Choreographien und populären Gesangsummern der Jahre 1929–30.<sup>6</sup>

Die Transformation des Mediums vollzieht sich, intradiegetisch, zwischen zwei Premieren von Filmen im Film. SINGIN' IN THE RAIN beginnt mit der spektakulären Premiere eines ›stummen‹ Mantel-und-Degen-Epos mit dem Titel THE ROYAL RASCAL im Kinopalast Grauman's Chinese Theatre, mit der die Feier des populären Darsteller-Paares, Don Lockwood und Lina Lamont, einhergeht. Der Film endet mit der Premiere des ersten Tonfilm-Musicals der Monumental Pictures, THE DANCING CAVALIER, in dem die gleichen Akteure dem Publikum nun auch als Dream-Team der tönenden Leinwand verkauft werden sollen. Die Story, die zwischen diese beiden Ereignisse eingehängt ist, referiert historisch auf die chaotische Zeit des Medienumbruchs<sup>7</sup> zwischen 1927 (dem die Branche schockierenden Erfolg der Warner Brothers-Produktion THE JAZZ SINGER) und der endgültigen Durchsetzung des Tons im Film um 1929/30. Übersetzt werden diese Ereignisse in eine Fiktion, deren Grundprinzip in vielfach variierten Figurationen der Substitution zu erkennen ist. Das bewährte System der Konstruktion von Ausdrucksbewegungen im Speechless Film kann nicht etwa im Nu durch ein überlegenes, weil realitätsnäheres Abbildungsverfahren abgelöst werden. So will es allenfalls das teleologische Missverständnis von der Notwendigkeit historischen Fortschritts. Vielmehr muss mit Einführung der Tonspur, und das heißt vor allem: mit der Möglichkeit einer analogen Repräsentation der menschlichen Stimme, ein ganz neues System der Konstruktion von Ausdruck erarbeitet, am Publikum erprobt, von diesem akzeptiert und schließlich in Produktion wie Rezeption konventionalisiert werden.

Diese Prozesse werden in SINGIN' IN THE RAIN gleichsam im Zeitraffer abgespult: Monumental Pictures, ein fiktives Hollywood-Studio, muss sich, um zu überleben, auf den Tonfilm einlassen. Doch der Übergang ist aus denselben Gründen schwieriger als angenommen, aus denen Hollywood den Verlockungen des Talkie (das technisch längst vor 1927 möglich gewesen wäre) so hartnäckig widerstanden hat. Die Filmindustrie hatte im Studiosystem immerhin eine ebenso profitable wie elaborierte Produktionsmaschinerie etabliert, deren Professionals nun plötzlich gezwungen waren, alles zu vergessen, was sie binnen zweier Jahrzehnte gelernt, standardisiert und ständig verfeinert hatten. Nicht umsonst waren viele der hier Beteiligten von dem Rückfall in eine mediale Steinzeit, der ihnen aufgenötigt wurde, geradezu angewidert, selbst wenn sie auch in der neuen Maschinerie reüssierten. Davon ist in SINGIN' IN THE RAIN wenig zu spüren.



Medienumbruch – ›Revolution in Hollywood‹, annonciert als ›burning news‹ in der Schlagzeile der *Variety*.

Das Megaphon verdeckt das Gesicht des *juvenile*-Tenors, dessen Stimme, so suggeriert es die Überblendung auf den Trichter, scheinbar die Imagination einer Rosettenformation aus Körpern von ›beautiful girls‹ im Busby-Berkeley-Stil hervorbringt.

Der Plot des Films trägt die Geschichte des Medienumbruchs in ein schlichtes, dafür aber um so effektiveres Schema von Oppositionen ein, die streng entlang der Gender-Kodierungen<sup>8</sup> verlaufen. Mit Hilfe dieses Schematismus, der ganz und gar auf ›Stimme‹ abgestellt ist, können film-technische Substitutionsverfahren wie etwa der technisch hergestellte Stimm-Wechsel im Dubbing, diegetische Evidenz gewinnen. Als überaus dankbar erweist sich zudem der Umstand, dass seither (wenn auch ungleichzeitig) eine weitere mediale Revolution stattgefunden hat, so dass die Wechsel der Register auch durch ein Hin und Her zwischen Technicolor und archaischem Schwarz-Weiß illustriert werden können. Unter Rückgriff auf Ergebnisse einer umfangreichen Forschungsliteratur konzentriert sich die folgende Analyse auf die Verfahren der in *SINGIN' IN THE RAIN* explizit zur Schau gestellten Inszenierungen der Sprech- und Singstimmen und versucht, einen kritischen Blick auf die Performanzen der medientechnisch fabrizierten Stimmen und die durch sie hervorgebrachten semantischen Effekte zu eröffnen. Dass dabei induktiv verfahren wird, muss kein Nachteil sein. Vielfalt, Bandbreite und Komplexität des in diesem Musical auf sämtlichen Darstellungsniveaus geführten Diskurses über die Stimme, ihre mediale Herstellung und die von ihr induzierten medialen Effekte lässt sich auf diese Weise vielleicht am besten bewahren.<sup>9</sup>

## 2. ENDE EINER ÄRA

### 2.1. MECHANISIERUNG ALLER GESTEN – INSZENIERUNGSPRAXIS IM STUMMFILM-STUDIO

Hollywood, 1927: the silent-film romantic team of Don Lockwood (Gene Kelly) and Lina Lamont (Jean Hagen) is the Toast of Tinseltown. While Lockwood and Lamont personify smoldering passions on screen, in real life the down-to-earth Lockwood can't stand the egotistical, brainless Lina. He prefers the company of aspiring actress Kathy Selden (Debbie Reynolds), whom he met while escaping his screaming fans. Watching these intrigues from the sidelines is Cosmo Brown (Donald O'Connor), Don's best pal and on-set pianist. Cosmo is promoted to musical director of Monumental Pictures by studio head R. F. Simpson (Millard Mitchell) when the talking-picture revolution commences. That's all right for Cosmo, but how will talkies affect the upcoming Lockwood-Lamont vehicle »The Dueling Cavalier«? Don, an accomplished song-and-dance man, should have no trouble adapting to the microphone. Lina, however, is another matter: put as charitably as possible, she has a voice that sounds like fingernails on the blackboard. The disastrous preview of the team's first talkie has the audience howling with derisive laughter.<sup>10</sup>

Der Referent will anscheinend nicht verraten, wie die Geschichte ausgeht – eine höchst überflüssige Zurückhaltung. Selbst wer den Film noch nicht gesehen hat, wird das Genre-Muster erkennen und voraussehen, dass das Gute, das in diesem Fall mit der guten Stimme identisch ist, sich durchsetzen wird – freilich erst nach dem allfälligen Umweg, der hier über die Entlehnung der Stimme des Starlet an den Filmkörper der Diva verläuft. Mit diesem Trick wird das Studio gerettet. Lina Lamont freilich will mehr. Sie ist verständlicherweise darauf aus, sich Kathys Leinwandstimme auf Dauer anzueignen. Das könnte sogar gelingen, wäre Lina nur in der Lage, im ›wirklichen Leben‹ den Mund zu halten. Da sie dazu nicht bereit ist, wird sie auf spektakuläre Weise auf offener (Kino-)Bühne enttarnt.

Anders als das im Film vorgeführte Publikum, welches das mit stummfilmischen Mitteln erzeugte Körperbild der Lina Lamont mit einer Imagination ihrer realen Person – dem *Star* – ineins setzt, wird der Zuschauer des Tonfilms *SINGIN' IN THE RAIN* von Beginn an mit der Differenz konfrontiert, die zwischen Leinwandbild, Star-Persona und dem Leben ›behind the screen‹ existiert, denn diese Differenz ist – verkörpert in Linas Stimme – die treibende Kraft des Plot. Von daher wird dem Zuschauer auch von Beginn an nahegelegt, konsequent auf die

enorme Künstlichkeit des (für ihn längst prähistorischen) ›stummen‹ Kinos achtzugeben. Erst in dem Moment jedoch, in dem die Zeit dieses Kinos verstrichen ist, wird ihm vorgeführt, mit welchen Mitteln es seine Geschichten erzählt und seine Diskurse hergestellt hat. Dieser genau in der Handlungsmitte des Films plazierte Blick auf den Arbeitsalltag im Studio zeigt prägnant den Konstruktionscharakter all dessen, was ein Silent Movie an Ausdruckswerten vermittelt. Es ist darum der Mühe wert, zunächst einmal diese Szene näher zu betrachten.

Don Lockwood, dessen Star Value durch den aktuellen Erfolg von *THE ROYAL RASCAL* einmal mehr bestätigt worden ist, betritt, sportlich gekleidet, die Studios der Monumental Pictures und durchquert diverse, nicht weiter von einander abgetrennte Sets, in denen simultan und unbekümmert lautstark an der Herstellung von Filmen gearbeitet wird. Dieser Gang durch das Studiogelände entwirft eine kleine Enzyklopädie der zeitgenössisch populären Genres und stellt verschiedene Verfahren der Konstruktion von Illusion aus. Im Set eines exotistischen Films führt Don einen Smalltalk mit einem als afrikanischer Kannibale verkleideten weißen Darsteller namens Maxie, der in der Drehpause ein Butterbrot verzehrt und *Variety* liest, sowie mit Cosmo; alles spricht über den unverhofften Erfolg des *JAZZ SINGER*, dem »all-time smash end of first week« (*Variety*). Der Kannibale meldet noch Zweifel an: »All-time flop end of the second.« So wird vergegenwärtigt, dass jene medienhistorische Wende, deren Vorboten sich bereits in der Exposition des Films angekündigt haben, nun endgültig eingetreten ist – ohne dass es die Profiteure des hoch elaborierten Produktionssystem ›Stummfilm‹ schon wahrhaben wollen.<sup>11</sup> Auch Don ist wie selbstverständlich zum Drehbeginn des neuen ›Lockwood & Lamont‹-Streifens angetreten, der – wie den Beteiligten bewusst<sup>12</sup> – unter dem Titel *THE DUELLING CAVALIER* und im Design einer »French Revolution story« nichts anderes als einen Neuaufguss von *THE ROYAL RASCAL* liefern soll.

Das Set – Rokoko à la Hollywood – steht bereit, die Arbeit beginnt, und sie verläuft wie am Schnürchen. Don schreitet aus dem Hintergrund auf Lina zu, die anmutig auf einer marmornen Bank sitzend das Schäferstündchen erwartet; als er sie erblickt, wirft er seinen Stab hinter sich und geht auf sie zu. Während der Probe führt das Paar einen simultanen Small Talk über aktuelle Tages-Ereignisse, der sich zum Streitgespräch entwickelt, als die Rede auf Kathy Selden kommt. Don erfährt, dass Kathy aufgrund einer gezielten Intrige Linas ihren Job verloren hat. Seine Wut verbergend, spielt er gleichwohl routiniert weiter den Galan. Die Anweisungen des Regisseurs, der den aktuell geführten Dialog nicht hören kann, geraten notwendig unpassend. Gerade in dem Moment, als Don von Linas Intrige

Ines Steiner

182

erfährt, verlangt er: »Okay Don, now you remember, you're madly in love with her ... And you have to overcome her shyness and timidity.« Auf das Kommando des Regisseurs, »Roll'em«, wird die zuvor choreographisch eingeübte und ausgeleuchtete Liebesszene dann abgedreht, Cosmo spielt ein Menuett aus Mozarts *Don Giovanni*, um das Paar ›in Stimmung‹ zu versetzen. Im Hintergrund des Sets verweist eine marmorne Amor-Statue mit Pfeil und Bogen auf den erotischen Diskurs. Durch den Doppeldiskurs von galanten Bewegungsabläufen und dem verbalen Schlagabtausch entsteht ein äußerst komischer Effekt.

Roscoe Dexter gibt Regieanweisungen für die galante Szene bei der Probe der Stummfilmversion von *THE DUELLING CAVALIER*



Don kommt von hinten ins Set, mit ›romantischem Blick‹ nach seiner Liebsten suchend, vollführt eine ekstatische Geste und wirft den Stab in die Kulissen. Er legt seine Hände auf Linas Augen, ihre Schultern zucken, ›freudig‹ erkennt sie ihn wieder. Sie blicken einander an »with the purest of love shining in their eyes.«<sup>13</sup>

Don (freudig lächelnd, aber mit grimmigem Tonfall): »Why, you rattle-snake, you. You got that poor kid fired.«

Lina (strahlend): »That's not all I'm gonna do if I ever get my hands on her.«

Don (lächelnd, aber mit verächtlichem Tonfall): »I never heard of anything so low.«

Regisseur (begeistert): »Fine. Fine. Looks great.«

Don (kniert schmachmend neben ihr nieder): »What did you do it for?«

Lina (tippt ihm mit dem Fächer, dem Instrument der ›stummen‹ Koketterie, auf die Schulter): »'Cause you liked her. I could tell.«

Don (kniend, nimmt ihre Hand und küsst sie ›leidenschaftlich‹): »So, that's it. Believe me, I don't like her half as much as I hate you. You reptile!«

Lina (lächelnd): »Sticks and stones may break my bones – «

Don (erhebt sich): »I'd like to break every bone in your body.«

Lina (erhebt sich): »You and who else, you big lummox!«

Regisseur (aufgeregt): »Now, kiss her Don.«

Don und Lina küssen einander ›leidenschaftlich‹

Dexter (begeistert): »That's it. More. Great. Cut.«

Der Kuss wird abrupt abgebrochen.

Lina (atemlos, kokett fragend, als habe sie die Distanz zur performierten Rolle verloren): »Oh, Donny. You couldn't kiss me like that and not mean it just a teensy-weensy bit.«

Don (ärgerlich, ironisch): »Meet the greatest actor in the world. I'd rather kiss a tarantula.«

Lina (lächelnd gespielt naiv): »Ahhh – you don't mean that.«

Don (ernst, drohend): »I don't? Hey Joe, bring me a tarantula.«

Regisseur (maliziös, die Situation als romantisches Zwiegespräch verkennend, drängt zur Eile): »Stop that chit-chat, you love birds. Let's get another take.«

Doch dieser nächste Take lässt für immer auf sich warten, denn genau in diesem Moment unterbricht der Produzent, der aufgeregt zum Set geeilt ist, die Dreharbeiten und ruft die Revolution aus: »We're going to make THE DUELLING CAVALIER into a talking picture.«

Diese Episode setzt (auf mit allein sprachlichen Mitteln nicht einholbare Weise) auf ein recht schlichtes Stilmittel: die Diskrepanz von Bild und Ton. In ihm ist gleichwohl die für den gesamten Film entscheidende Leitdifferenz enthalten: Bilder lügen – Stimmen sprechen die Wahrheit; der im Bild bewegte Körper führt eine täuschende Fassade vor – Rede und Gesang dagegen kommen von Herzen.<sup>14</sup> Hier, im Studio, wird diese Differenz gleichsam an ihrem Ursprungsort behaftet, denn durch den kontrastiven Einsatz von Bild und Ton wird die Schauspieltechnik im ›stummen‹ Film als ganz und gar veräußerlichte Leistung der Akteure vor der Kamera evident gemacht. Noch ein- und gewissermaßen zum letzten Mal wird vorgeführt, wie der ›stimmlose‹ Körper auf der Leinwand zum Sprechen gebracht wird: mit Hilfe eines vorkodierten und auf Kommando der Regie jederzeit abrufbaren Repertoires einstudierter Gesten, die als romantisch gemeinte Ausdrucksbewegungen misszuverstehen dem Publikum allmählich und mühsam beigebracht worden war. Der Audio-Video-Dissolve stellt als solcher natürlich kein neues oder originelles Kunstmittel dar, und er ist auch hier intradiegetisch eindeutig an die Vorgeschichte der unlösbar erscheinenden professionellen Liaison und persönlichen Aversion zwischen Don und Lina zurückgebun-

den; er hat zudem (wie die Exposition vorgeführt hat) seinen Sitz im Leben, denn für eine auf die Körperbilder ›ihrer‹ Stars fixierte Öffentlichkeit muss die Fassade auch Off-Screen stets aufrechterhalten werden.

Gleichwohl ist kaum einmal so prägnant wie hier in *Bild und Ton* gesetzt worden, was Filmtheoretiker wie Kuleschow, Eisenstein oder Benjamin als Eigentümlichkeit des frühen Filmschauspiels begrifflich zu fassen versucht haben. Und als schöne Pointe am Rande fällt ab, dass die von ›Monumental‹ projektierte Schmonzette zwar ums andere Mal auf ein Publikum zugeschnitten ist, das den Kode ›romantischer‹ Liebe internalisiert hat, dass die erzählte Story aber genau am Ende jener ›galanten‹ Ära situiert ist, in der, kulturgeschichtlich verstanden, die Konventionalisierung der Gesten und die Disziplinierung des Körpers oberstes Gebot gewesen ist. Die für den Film im Film intradiegetisch angekündigte Revolution findet derart extradiegetisch – im Musical – statt, und darum muss sich auch das projektierte Melodram als solches revolutionieren lassen: Es kann, ganz folgerichtig, nichts anderes aus ihm hervorgehen als – ein Film im Film – ein Musical. So weit sind wir jedoch noch lange nicht. Im Gegenteil: Noch herrscht die Illusion, man könne im Grunde so weitermachen wie gehabt. Irritiert bringen Techniker, Regisseur und Darsteller ihre Argumente gegen den Wechsel zum Tonfilm vor. Doch der Produzent bleibt hart. Nach dem sensationellen Erfolg des *JAZZ SINGER* kann sich kein Studio mehr dem Branchentrend entziehen: »Every studio is jumping on the band waggon [...]. All the theaters are putting in sound equipment. We don't want to be out of it.« Zwar ist die Skepsis der versammelten Professionals allgemein, doch die Zuversicht ihres geradezu unglaublich kollegial agierenden Vorgesetzten (»You do what you always did. You just add talking to it«) könnte durchaus ansteckend wirken – bis sich Lina in aller ihr vom Drehbuch vorgeschriebenen Unschuld und »in her coarest, flattest tones, cutting trough like a knife«<sup>15</sup> zu Wort meldet: »Well, of course we talk. Don't everybody?« – In diesem Moment wird unmissverständlich deutlich: Mit dem »just add talking to it« kann es nicht getan sein. Linas Stimme wird zur unüberhörbaren Störung in der schönen Illusion einer problemlosen Remodellierung des Mediums. Alle Beteiligten (außer der Sprecherin selbst) haben Ohren, es zu hören: Linas Star Value, von dem aller Existenz abhängt und der sich bislang durch geschickte Camouflage hat erhalten lassen, steht nun plötzlich auf dem Spiel – die Diva hat einen medial erzeugten Körper, der ihren Namen trägt, aber sie wird ihn nicht mit ihrer Stimme signieren können. Und die Monumental Pictures haben ein Problem.



## 2.2. VORSPIEL VOR DEM FILMTHEATER

Der frühe Film mag Bilder von Körpern auf die Leinwand gebracht haben, die man dort sprechen sah, jedoch nicht sprechen hörte. Es stellt jedoch eine geradezu fahrlässige Verkürzung der Komplexität des Mediums dar, wenn man annimmt, dieses Medium habe der Stimme vor 1927 nicht bedurft. Die breit angelegte und überaus kunstvoll konstruierte Exposition von SINGIN' IN THE RAIN scheint absichtsvoll darauf angelegt, solchen Vereinfachungen entgegenzutreten. Abend / Außen – eine Totale etabliert Zeit und Raum des Geschehens: Vor dem Eingang von Grauman's Chinese Theater hat sich anlässlich der Premiere von THE ROYAL RASCAL eine schaulustige Menge versammelt, die gebannt die Ankunft der Stars erwartet. Polizeiketten bilden einen Korridor und halten die Fans<sup>16</sup> vom roten Teppich zurück. Das hysterische Kreischen der Menge<sup>17</sup> bildet einen Wall of Sound, vor dessen Hintergrund sich das Spektakel gleichsam ›stumm‹ entwickelt. Vor dem Kinoeingang hat freilich die matronenhafte Film-Kolumnistin Dora Baily<sup>18</sup> Aufstellung genommen, die – ›highly exited; in an overecstatic gushy voice‹<sup>19</sup> – den Auftritt der Stars kommentierend bespricht. Diese über Lautsprecher vor Ort verstärkte und im Radio verbreitete *Stimme Hollywoods* bringt gleichsam fortlaufend den Paratext zum sichtbaren (oder vor dem Tonempfänger imaginierten) Star-Image hervor, und sie wird durch das Verhältnis von Sound Track und Image Track geradezu ›sichtbar‹ gemacht. In den ersten Einstellungen ist Baileys Stimme nur als Voice Over zu hören; mit einem Reißschwenk sucht die Kamera nun erst nach der Quelle des Tons und setzt schließlich den korpulenten Körper der Sprecherin vor einem kopfgroßen (›old fashioned‹) Mikrophon – dem Symbol ihrer Diskursmacht – stehend ins Bild. So wird die Zweideutigkeit des ›miracle of sound‹ bereits erkennbar, bevor das Medium Film selbst direkt vom Ton infiziert wird. Tatsächlich war es längst auf die Macht der Stimme(n) und der mit ihr assoziierten Medienverbände angewiesen.

Die Interaktion zwischen der Rednerin und dem anwesenden Publikum ist wie ein Dialog gestaltet. Während die Klatsch-Reporterin in rascher Folge ›every star in Hollywood's heaven‹ ankündigt, beginnt die angespannte Menge anlässlich jeder vorfahrenden Limousine aufgeregt zu kreischen. Es sind die zur Kennlichkeit karikierten Stars der Silent Screen der Zwanziger Jahre, die hier durch die maliziösen Kommentare vorgestellt werden: Das rothaarige ›Zip Girl‹ Zelda Zanders (Rita Moreno), ›the darling of the flapper set‹ (nach dem Vorbild des ›It Girls‹ Clara Bow) oder der typische Vamp Olga Mara (Judy Landon), der auf das reale Vorbild Nita Naldi rekurriert. Beim Anblick des Dream Teams Don Lockwood und Lina Lamont, ›the stars of tonight's picture, those romantic lovers of

Ines Steiner

186



»The miricale of sound«: Die noch »unsichtbare« Stimme Hollywoods, verstärkt durch den Lautsprecher, kommentiert den Einzug der Stars in der Exposition.

Die Klatsch-Kolumnistin Dora Bailey (Madge Blake) adressiert durch das Mikrophon das Fan-Publikum vor Ort und an den Radioapparaten.

the screen«, überschlägt sich die Stimme von Dora Bailey nahezu: »Oh, oh folks, this is it! This is it!« Don Lockwood, in einen weißen Kamelhaarmantel und weißen Borsalino gekleidet, trägt ein »million-watt white toothed smile« im Gesicht und wirft Küsse in die kreischende Menge. Ein weiblicher Fan fällt bei seinem Anblick in Ohnmacht. Don Lockwood ist, das wird sich auch im weiteren Verlauf des Films erweisen, in *SINGIN' IN THE RAIN* die eigentliche Diva. An seiner Seite schreitet geziert lächelnd und mit den Augenlidern klimpernd Lina Lamont, eine Blondine<sup>20</sup> in eleganter zartgrüner Abendrobe und gleichfarbigem Pelz. Sie verläßt sich ganz auf die lang eingelernten Ausdrucksbewegungen ihres Körpers und ihres Gesichtes und wird aus gutem Grund in den ersten Minuten des Films daran gehindert, auch nur ein einziges Wort zu verlieren.<sup>21</sup>

Denn, ganz klar: This is not it – hier ist nichts, wie es sich nach außen darstellt. Auch wenn die Reporterin verheißungsvoll von Hochzeitsglocken spricht, sind Don und Lina im Privatleben alles andere als ein Traumpaar, und die bestens präparierte, dem Illustriertenklischee entsprechende Geschichte des Making of a Star, die Don auf Aufforderung – »Oh Dora, not in front of all these people« – vor dem Mikrophon abspult, erweist sich als gründliche Entstellung der Wahrheit.<sup>22</sup> Einmal mehr wird der Audio-Video Dissolve schlicht, aber wirkungsvoll eingesetzt: Die offizielle, vom Publicity Department vorgefabrizierte Erfolgs-Story, die Lockwood unter dem Motto »dignity, always dignity!« vorträgt, wird auf der Bildebene durch eine Art slapstickhafter Nummernrevue als Geschichte fortlaufender Entwürdigungen konterkariert. Don Lockwood – wie wir erfahren: Ex-Vaude-



Don Lockwood, Lina Lamont, Dora Baily und Cosmo Brown in einem four shot vor dem Mikrofon. Während der folgenden Rede an »sein« Publikum wird Don ausschließlich in Großaufnahme gezeigt, die ihn umgebende Gruppe, und mit ihr die Diva, wird somit aus dem Kader verdrängt.

villeclown, Ex-Mood-Musician, Ex-Stuntman und schließlich ein Star – bleibt freilich souverän. Er bedient als Profi auch mit seiner Stimme jenen Betrieb, der seinen Körper als Ikone brauchen kann, aber er geht nicht in ihm auf, sondern hält so viel ironische Distanz, dass er nicht einmal die Unwahrheit zu sprechen braucht.

Gerade weil er auf die Erfahrung der Entwürdigung durch das Show Business zurückblickt, erhält er sich – dies legt die Rückblende nahe – eine Würde, die den eigentlichen Geschöpfen des Systems abgeht. Diese sind, Lina Lamont vor allen anderen, ständig versucht, sich selbst mit dem Imaginären ihrer Persona zu verwechseln. Auch sie tritt in Dons Erinnerungen auf: als bereits etablierte Leading Lady, die den unbedeutenden Stuntman wie Luft behandelt, nur um sich ihm an den Hals zu werfen, sobald er als kommender Star zu erkennen ist. An dieser Stelle ist die Stimme des Stars noch immer nicht zu hören, sondern nur Dons Doubletalk: »Lina was and is always an inspiration to me: warm and helpful – a real lady«. Aber schon hier wird erkennbar, dass SINGIN' IN THE RAIN diese Figur auf konsequente, aber nicht unproblematische Weise in einem doppelten Register besetzt. Zum einen hat Lina Lamont, ganz einfach, einen schlechten Charakter oder sie ist, noch schlichter, dumm, intrigant und böse. Zum anderen aber ist sie, so wie sie in der Diegese des Films als Person erscheint, sehr deutlich als Produkt einer kollektiven Imagination ausgestellt, durch deren Projektionen sie ein realitätsbezogenes Selbstverständnis substituiert hat. In diesem Diskurs geht es nicht um moralische Wertungen, sondern um Einstellungen zum Medium: Don Lockwood verleiht seinen Körper fallweise an den Film, weil das sein Business ist; Lina Lamont verleiht den Körper ihrer Star Persona fallweise an die Wirklichkeit, weil ihr diese Wirklichkeit nurmehr als Extension des Filmbusiness erscheint.<sup>23</sup>

Die damit eröffnete Alternative ist höchst konsequent gestellt, und sie gewinnt um so mehr an semantischem Reichtum, als SINGIN' IN THE RAIN auch Dons Zuversicht, Sein und Schein unterscheiden zu können, mehr als einmal in Frage stellt. Sie geht jedoch, seltsamerweise, gerade im Hinblick auf Lina Lamont nicht durchweg auf, weil sie durch jenen moralischen Dualismus von Gut und Böse unterlaufen wird, der unmittelbar an das Skandalon von Linas ›unmöglicher‹ Stimme gekoppelt wird.<sup>24</sup> Lina Lamont wird strikt kontrastiv zu ihrem Leinwandpartner charakterisiert. Sie wird eingeführt als eine ihre darstellerischen Fähigkeiten und ihre körperliche Attraktivität überschätzende Dumb Blonde, die ihre privilegierte Stellung im Studiosystem mit allen Mitteln zu behaupten sucht. Gleichsam als Erkennungsformel ist ihr die rekurrente rhetorische Frage: »Do you think I'm dumb or somethin'?« mitgegeben. Im Gegensatz zu Don scheint Lina zudem, bezeichnenderweise, stets versucht zu glauben, was die Presse kolportiert. Als sie ihren Partner nach der Premiere als »Donnie« und ihren Verlobten anspricht, kontert dieser ironisch: »My fiancée! Lina, you've been reading those fan magazines again. You actually believe that *banana oil* that Dora Bailey and all the columnists dish out. There is nothing between us! There has never been anything between us – just air!«

### 2.3. STIMME ALS WITZ – WITZ ALS STIMME

Im Innenraum von Grauman's Chinese Theatre wird schließlich das Mantel- und Degen-Epos THE ROYAL RASCAL mit großer Orchester-Begleitung dem Premierenpublikum vorgeführt. Der Ausschnitt, den das Musical aus diesem Kostüm- und Abenteuerfilm zeigt, zitiert die Verfahren der populären Swashbuckler-Filme eines Douglas Fairbanks; er wird als stummer Schwarz-Weiß-Film auf der von einem goldfarbenen Brokatvorhang gerahmten Leinwand im farbig und tönend repräsentierten historisch rekonstruierten Kinoraum gezeigt. Dieser Ausschnitt erzeugt einen komischen Effekt, indem er vorführt, wie das Hören im Stummfilm inszeniert wird:

The Royal Rascal macht sich über Stummfilmkonventionen lustig, wenn ›Philippe‹ [Gene Kelly] eine große pantomimische Lauschgeste macht und ein Zwischentitel verkündet: ›I think I hear a footstep‹, worauf seine Geliebte [Jean Hagen] ebenfalls die Hand ans Ohr legt, obwohl nur Begleitmusik zu hören ist.<sup>25</sup>

Die folgende Duellsszene zwischen Philippe und dem Villain im Treppenhaus stellt dann ein unvermitteltes filmhistorisches Selbst-Zitat dar, denn hier wurde einfach eine schwarz-weiß umkopierte Kelly-Sequenz aus dem Farbfilm *THE THREE MUSKETEERS* (1948) eingeschnitten.<sup>26</sup> Zumindest dem informierten Zuschauer wird damit deutlich, dass *SINGIN' IN THE RAIN* ganz explizit mit der Neukombination verfügbaren Materials operiert.

Wie bei Filmpremieren in den Zwanziger Jahren üblich, treten Don Lockwood und Lina Lamont nach Ende der Vorführung vor den Vorhang, um dem applaudierenden Publikum zu danken. Don hindert Lina mehrfach geschickt daran, das Wort zu ergreifen, ja er zerrt sie geradezu von der Bühne mit den Worten: »We screen actors aren't much good at speaking in public, and so we had better just act out our thanks!« Umringt vom Produzenten R. F. Simpson (Millard Mitchell) und dem aufgeregten Presseagenten Rod (King Donovan) beschwert sich Lina hinter dem Vorhang empört über diese Verdrängung aus dem öffentlichen Raum – und zwar mit schriller Stimme: »F' heaven's sake, what's the idea – can't a girl get a word in edgewise? They're my public, too!«

Wenn hier zum erstenmal die Stimme Lina Lamonts hörbar wird, so ist die schockierende Wirkung beabsichtigt: »We hear her voice for the first time. It is shrill, flat, and coarse, and a terrific shock coming out of that beautiful face«,<sup>27</sup> kommentiert die Regieanweisung.

Wenige Minuten zuvor wurde der im *Louis XVI*-Gewand kostümierte, mit dem Ausdruck von ›Anmut‹ bewegte Diven-Körper auf der stummen Leinwand von einem weiblichen Fan im Publikum als Inbegriff ›weiblicher‹ Perfektion wahrgenommen: »She's so refined. I think I'll kill myself.« Die Wahrheit hinter dem Vorhang sieht anders aus, und zwar in jeder Hinsicht: Nicht allein ist Lina von der Natur, was das ›Voice Department‹ angeht, gerade nicht begünstigt, sondern sie artikuliert sich zudem in einem ›vulgären‹, dialektal eingefärbten Soziolekt, der auf mangelnde Bildung – und! – auf einen Mangel an Ausdrucksvermögen hindeutet. Dabei hat sie mit ihrem Einwand ja vollkommen recht – da draußen sitzt *ihr* Publikum – doch ihre affektiv unbeherrschte Redeweise setzt sie ins Unrecht und zudem dem Verdacht aus, der ›unkultivierte‹ oder, weniger wertend, ›undomestizierte‹ Zug an ihr sei auch mit einem moralischen Defizit assoziiert. Die Inkompatibilität zwischen dem auf den Effekt glamouröser ›Schönheit‹ zugerichteten Körperbild und einer den Effekt ›Hässlichkeit‹ erzeugenden Stimme führt im Mittel des »comic shock« zu einer desillusionierenden Überraschung: Die lächerliche Wirkung zerstört jene Integrität, die dem Stummfilm-Star seine sichtbare und daher unmittelbar scheinende Attraktivität verliehen hat. Die Illusion der ›Intaktheit‹ des perfekten Stars der ›stummen Leinwand‹, das

Image Lina Lamonts, kann daher nur aufrecht erhalten werden, indem Studioboss und Presseagent der Diva unmissverständlich, aber mit der gebotenen Vorsicht nahelegen, in der Öffentlichkeit besser nicht zu sprechen.

Simpson (beschwichtigend): »But Lina, the publicity department – Rod here – decided it would be better if Don made all the speeches for the team.«

Lina (verständnislos, empört): »Why?«

Rod: »Lina, you're a beautiful woman – audience thinks you've got a voice to match. Studio has got to keep their stars from looking ridiculous, at any cost.«

Cosmo (sarkastisch): »No one's got that much money.«

Lina (empört): »What's wrong with the way I talk? What's a big idea – am I dumb or somethin'?«

Die Männer blicken einander entsetzt über soviel Mangel an kritischer Selbstwahrnehmung an.

Rod (zögernd, erklärend, wie zu einem Kind): »No – it's just that Don's had so much more experience, and – «

Lina (fordernd): »Next time, write me out a speech. I could memorize it!«

Cosmo (ironisch auf Linas mangelnde Schulbildung anspielend): »Sure, why don't you go out now and recite the Gettysburg Address?«

Lina (wütend): »What do you know about it, you – you piano player! Are you anybody?!«

Damit ist die ›eigene‹ Stimme des weiblichen Stars Lina Lamont, die sich Gehör verschaffen will, als Problem etabliert, das die Männer der Branche – Produzent, Regisseur, Tontechniker, Presseagent, Koakteur – zu kontrollieren und zu beherrschen haben. Und um die Lösung dieses Problems dreht sich der gesamte Plot. So weist der Film *Stimme* auf verschiedenen Ebenen als Artefakt aus. Er zeigt die fremdbestimmte Herstellung der ›adäquateren‹ Stimme der Diva, die auf der Enteignung ihrer ›eigenen‹ Stimme und deren Ersetzung durch eine ebenfalls enteignete, ›fremde‹, ›geliehene‹ Stimme basiert. Indem *SINGIN' IN THE RAIN* den Prozess der Konstruktion von Stimme performativ ausstellt, wird Stimme überhaupt, hier aber vor allem die als ›weiblich‹ semantisierte Stimme,<sup>28</sup> als ein komplexes, von hegemonialen kulturellen Vorstellungen geprägtes, auf spezifische semantische Effekte zugerichtetes Produkt offensichtliche.

Jean Hagen, die in *SINGIN' IN THE RAIN* die Stummfilmdiva Lina Lamont verkörpert und ihr die signifikant schrille, schauerliche Stimme verleiht, verfügte

selbst über eine äußerst wohlklingende Stimme und eine distinguierte, ausgebildete Sprechweise. In einigen gender-spezifizierten Lektüren des Films wird die Weigerung Lina Lamonts, auf ihre eigene, als »wild« semantisierte Stimme zu verzichten, als Akt des Widerstandes gegen die kultivierende Domestikation und medientechnische Zurichtung auf den Apparat interpretiert. Damit wird auch das Ende des Films, wenn sie erneut darauf insistiert, mit ihrer eigenen Stimme vor das Publikum zu treten, als Akt der Selbstbehauptung gegen das Studio-System interpretiert.

Lina Lamont has only one real problem – she is untalented. This sin of performance becomes identified with her voice. Two male-critics explain this with relish: She has ›a Brooklyn accent [that] ... could stop a beer truck at fifty yards‹, says John Mariani, while according to Clive Hirschhorn, she ›cannot string two syllables together without sounding like a grinding gear.‹ Lina suffers from this kind of exclusion twice: she doesn't talk like a lady and she doesn't sing like a star. Her voice promises no ecstasy; it cannot be desired like a fetish. In fact, not only does Lina have a voice unacceptable for a musical – it is not even human. In an early version of the screenplay, the script makers allusion to Lina's future career, that she is currently appearing IN JUNGLE PRINCESS, in which she doesn't say a word – just grunts.<sup>29</sup>

Dieser Einspruch ist insofern vollkommen berechtigt, als er auf die unbestreitbare Tendenz des Musicals verweist, Lina nicht nur als Fossil einer vergangenen Ära auszustellen, sondern ihr berechtigtes Interesse am Fortbestehen ihrer Karriere auch noch moralisch zu denunzieren. Er geht aber gleichwohl in die Irre, weil er auf einer teleologischen Lesart von Filmgeschichte beruht, die SINGIN' IN THE RAIN keineswegs unmissverständlich vorgibt. Im Gegenteil wäre zu behaupten, dass Lina ein Talent besitzt, das eine Kathy Selden nie erreichen wird: einen technisch vollkommen disziplinierten, beliebig ›verfilmbaren‹, zur Projektion einer Star Persona tauglichen Körper. *Not even human*: dieses Prädikat trifft auf sie doppelt zu. Ihre Stimme mag sich ›un‹-, gar ›untermenschlich‹ ausnehmen, und sie musste ja in der Tat künstlich hergestellt werden. Aber das gleiche gilt, in der anderen Ära des Mediums, für die Projektionsfläche ihres ›stumm‹ gestellten Körpers, der den Bedürfnissen des Mediums und seiner Konsumenten entsprechend in eine ›un‹- und ›übermenschliche‹ Maschine verwandelt worden war.<sup>30</sup>

Für Don Lockwood gilt entsprechendes. Und er kann der so gestellten Falle nur deshalb entkommen, weil ein Mangel an unmittelbarer Anerkennung ihn ge-

zwungen hat, sich zusätzliche Kompetenzen zu erwerben. Don hat mehr als seinen Körper; er ist *auch* Musiker, Sänger und Tänzer. Gleichwohl ist er, wie ihm Kathy Selden anlässlich beider erster Begegnung verdeutlicht, als stummfilmisch produzierte Persona niemals mehr gewesen als ein »shadow on film – a shadow – [...] not flesh and blood«. Kathy, die einstweilen noch zur Erheiterung eines Partypublikums aus der sprichwörtlichen Torte zu springen hat, träumt davon, auf dem Theater Shakespeare und Ibsen zu spielen und deren »glorious words« durch Einsatz ihrer Stimme aktuelle Bedeutung zu verleihen. Bildungsbürgerlich auf die Tradition des Phonozentrismus verpflichtet, kann sie das hoch komplexe Kommunikationsgefüge der Silent Movies ohne weiteres als bloßen Sonderfall von ›pantomime‹ abtun und verfügen: »The personalities on the screen just don't impress me. I mean, they don't talk – they ... don't act – they just make a lot of dumb show.«

Don ist indigniert – und dies nicht ohne Grund. Er steht zwar nicht in der Gefahr, sich selbst mit seiner medial produzierten Persona zu verwechseln, aber eben deshalb darf er darauf beharren, dass sein auf der Leinwand produziertes Körperbild Produkt einer künstlich-künstlerischen Übersetzungsleistung ist und nicht etwa nur das analoge Abbild quasi-natürlicher Bewegungsfolgen. Auf diese Weise inszeniert *SINGIN' IN THE RAIN* die erste Begegnung der zum Liebespaar bestimmten Akteure, sehr souverän, nach dem Genremuster der 1952 schon halb vergessenen *Screwball Comedy*, in der sich – in den frühen 1930er Jahren – die Film-Stimme um des Witzes willen konsequent vom Film-Körper emanzipiert hatte. Und seltsamerweise hat sich die feministische Filmkritik nie des wahren Skandalons des Musicals angenommen: nämlich, dass die vielleicht allzu hochnäsige, aber immerhin geistreiche Kathy Selden des Initialauftritts noch sehr viel hemmungsloser verraten wird als Lina Lamont. Don braucht sie nur im ›System Tonfilm‹ wiederzuentdecken und sie einmal singend anzuschmachten, indem er seine ›innere Stimme‹ unter dem Einsatz aller verfügbaren illusionistischen Technik in einer leeren Studiohalle in der Gesangs- und Tanz-Nummer *You were meant for me* sprechen lässt, und schon ist sie ihm bedingungslos verfallen und sogar dazu bereit ihre Stimme an Lina zu überschreiben. Der Don Lockwood, den uns die Exposition vorstellt, hätte ein wenig mehr Widerstand erwartet. Und deshalb mag man ihm auch nicht recht abnehmen, dass die Beziehung zu Kathy mehr darstellt als die wesentlich asexuelle Figuration einer ›Kameradschaftsehe‹ – so schön die an den Abschiedskuss angeschlossene Production Number *Singin' in the Rain* auch ausfällt. Und Kathy selbst hätte erst recht mehr von sich verlangen dürfen, als ihr für den Rest des Films zu tun erlaubt wird. Schließlich sind Shakespeare, Ibsen und Shaw [!] durch die Emergenz des Tonfilms nicht einfach erledigt.



### 3. ENDE EINES STARS

#### 3.1. HERZTÖNE – LIEBE IM TON(FILM)STUDIO

Um die Emergenz des Tonfilms ist es freilich eine eigene Sache. Wie allgemein bekannt, hat mit dem *JAZZ SINGER* keineswegs ein Tonfilm im heutigen Verständnis einen Sensationserfolg erzielt, sondern es lag da im Grunde ein Stummfilm vollkommen gängiger Machart vor, der lediglich – um als Star-Vehikel für den Sänger Al Jolson dienen zu können – einige mit der analog aufgezeichneten ›Originalstimme‹ unterlegte Gesangsnummern enthielt. Als Modell für weitere Blockbusters konnte dieses hybride Machwerk kaum dienen. Und dementsprechend zeigen sich auch die in *SINGIN' IN THE RAIN* präsentierten Professionals zunächst einmal grundlegend verwirrt. Jener Kode von Konventionen, der bis dahin die Filmherstellung organisiert hatte, bietet ihnen nicht einmal den Ansatz für die Lösung des Problems an, wie ein ›echter‹ Tonfilm zu machen wäre. Darum entschließt man sich, wider besseres Wissen, aber vollkommen einleuchtend, zunächst einmal dazu weiterzumachen, als sei im Grunde nichts geschehen – logischerweise genau dort, wo man mit der Verfertigung des Silent Movie aufgehört hat. Wo stummes Spiel war, muss nun das Talkie werden, und auch die so angelegten Parallelszenen fallen, recht besehen, höchst instruktiv und sehr komisch aus. Man kann allerdings kaum ignorieren, dass diese beinahe zur Überlänge gezogene Gagbahn (Ton im Tonfilm) ihre Effekte ganz allein auf Kosten der schon fast entthronten Diva gewinnt, und dass sie von daher deutlich hinter die Komplexität der vorigen diegetischen Konstellation (Stummfilm im Tonfilm) zurückfällt. Und man muss zugestehen, dass Lina Lamont im Zuge dieses Arrangements auf nachgerade denunziatorische Weise vorgeführt wird, weil nämlich die komischen Effekte nicht etwa aus der ›Natur‹ ihrer Stimme, sondern aus einer ihr unterstellten Unfähigkeit zur Anpassung ihres Schauspiels an die neuen medialen Gegebenheiten bezogen werden.

Das ist unfair insofern, als Don und Lina in dieser Hinsicht vor einer identischen Herausforderung stehen, und die gendersemantische ›Tendenz‹ des Films zeigt sich nirgendwo unverstellter als hier, wo dem Mann der Übergang zum Sprechen vor der Kamera wie von selbst gelingt, während die Frau ebenso vollständig und endgültig versagt. Unterdrückt und verschwiegen bleibt dabei, was der Ton- und Musikfilm *SINGIN' IN THE RAIN* sehr wohl im intrinsischen Gedächtnis behalten hat, aber nicht laut sagen darf: Dass die Stars des Silent Movie nicht etwa bessere Kleiderständer gewesen sind, die ein Produzent ihres kamera-tauglichen Aussehens wegen ins Studio gezerrt hat, sondern dass sie, um zu reüs-

sieren, über ein vielleicht banales, aber doch differenziertes Repertoire von Spielmöglichkeiten verfügen mussten. SINGIN' IN THE RAIN lässt sich, um dieses Eingeständnis zu vermeiden, auf die problematische Alternative ein, der kultur-beflissenen Kathy Selden im Nachhinein rechtzugeben und die stumme Pantomime ex post als ein Possenreißen und bloßes Simulakrum ›echten‹ Schauspiels abzutun. Doch mit der Trias des »triple-threat« (»she can't act, she can't dance, she can't sing«) stimmt zumindest so viel nicht, dass Lina sich das zum Starstatus gehörige Körperbild nicht hätte erwerben können, wenn ihr nicht gelungen wäre, den Habitus der Perfektion in ihrem *doing gender* auf der Leinwand so schön abzubilden, dass man ihr nachsehen möchte, dass sie diesen Habitus ständig mit ihrem wirklichen ›Ich‹ verwechselt.

SINGIN' IN THE RAIN bleibt der einmal eingeführten Leitdifferenz treu: Stumme Körper können lügen – sprechende Körper nicht. Tritt man nur einen Schritt zurück, so wird spätestens hier vollends deutlich, dass dieser Film auf einer mehr als nur gewagten Inversion kulturell tief verwurzelter Topiken beruht, denn traditionellerweise – beginnend bei der Schlange, die Eva im Paradies verführt – ist es die Stimme, die zu unendlicher Verstellung fähig ist, während der Körper, schaut man nur recht hin, die Wahrheit spricht.<sup>31</sup> Jene sogenannte Körpersprache, auf deren Verlässlichkeit sich Lügendetektoren bis heute verlassen, hat jedoch bereits der stumme Film der Illusion von Eigentlichkeit entkleidet. Im Übergang zum Tonfilm kann es darum nur darauf ankommen, das eine arbiträre Ausdruckssystem durch ein anderes, nicht weniger willkürliches zu substituieren. Und genau dies wird in SINGIN' IN THE RAIN tatsächlich vorgeführt – aber freilich so, als ob dabei nur die vom Medienwechsel induzierten Kinderkrankheiten des Tonfilms zur Debatte stünden. Gleichwohl macht der Film gerade darum deutlich, dass auch das neue, auf die technischen Innovationen eingestellte Schauspiel der medialen Konstruktion von Ausdruck dient, und dass es auch weiterhin auf einer veräußerlichten Darstellungstechnik beruht. Was nunmehr zählt, ist nicht die im von der Kamera erfassten Raum frei platzierbare Geste. Der Bewegungsspielraum des Körpers wird vielmehr extrem eingeschränkt und ganz auf das in der Requisite und den Kostümen versteckte Mikrophon ausgerichtet. Nur weil sie sprechen, können die Darsteller im um den Ton erweiterten Ausdruckssystem keineswegs einer wie auch immer gearteten ›inneren Stimme‹ hörbaren Ausdruck verleihen – eher im Gegenteil, müsste man sogar sagen, oder mit Lina Lamont: »I can't make love to a bush«.

Denn wie uns liebevoll gezeigt wird, ist die neue Technik so schwer- wie störanfällig. Nacheinander werden die technischen Katastrophen vorgeführt, die sich am Set eines frühen Tonfilms ereignen können, wobei zahlreiche Fehlleis-

tungen, die aus der Zeit des Medienumbruchs anekdotisch überliefert sind, geschickt zitiert werden. Die Aussage des Produzenten, »You just add talking to it«, erweist sich als trügerische Wunschvorstellung, denn die medientechnische, adäquate Darstellung und Herstellung der menschlichen Stimme im Film ist mit zahllosen Schwierigkeiten konfrontiert. Dazu gehört vor allem die strikte Kontrolle des Klangraumes, ohne die sich ›authentische‹ Effekte von Sound on Screen nicht erzeugen lassen. Das wird schon auf den ersten Blick anschaulich. Wieder blickt man auf die Rokokokulisse des *DUELLING CAVALIER*, die nun aber von ganzen Scharen von Fachleuten bevölkert ist: Tontechniker, Elektriker, zwei Regieassistenten. Der Boden ist bedeckt mit einer Vielzahl von Kabeln. Der Regisseur Roscoe Dexter, der unsicher und angespannt wirkt, gibt das neue Kommando »Quiet!«, erst dann befiehlt er, wie gewohnt, »Roll 'Em!« Sichtbar wird so bereits, dass das System der Filmproduktion hierarchischer geworden ist und mehr Arbeitsteilung verlangt, während der Spielraum für Spontaneität und Improvisation am Set geschwunden ist. Dexter zieht sich rasch in die schalldichte Aufnahmekabine zurück, um dort, neben dem Tontechniker sitzend, die Tonaufnahme zu kontrollieren; das gleichzeitige Schauspiel am Set, in das er nicht mehr ad hoc mit seinen Kommentaren eingreifen kann, muss er nun aus Distanz durch die Glasscheibe betrachten. Neben der Kostümbildnerin und Friseurin ist auch die Sprecherzieherin, Miss Phoebe Dinsmore, anwesend, die Lina Lamont wiederholt Instruktionen zur Artikulation des einfachen Satzes gibt, den sie zu sprechen hat: »Oh, Pierre, you shouldn't have come!«

Don und Lina führen dieselben Gänge und Gesten wie in der Stummfilmversion aus, doch die Szene muss am tönenden Set immer aufs neue wiederholt werden. Denn statt jener Vollprofis, die ihr Ausdrucksrepertoire en passant abspulen können, sind nun Anfänger am Werk, die erst lernen müssen, die Beweglichkeit ihres Körpers im von der Kamera erfassten Raum um der Tonaufnahme willen einzuschränken. Lina erweist sich als dazu vollständig unfähig, so oft Roscoe ihr die Sache auch erklärt: »Lina, look, Lina, don't you remember? I told you. There's a microphone right there – in the bush. [...] You have to talk into it. [...] The sound goes through the cable to a box. A man records it on a big record in wax, but you have to talk into the mike first – in the bush. Now try it again.« Da Lina ihre Rede nicht auf einen Punkt auszurichten vermag und den Oberkörper wie gewohnt in ausgreifenden Bewegungen hin- und her bewegt, geht jedes zweite Wort verloren. So wird das Mikrophon endlich ans Dekolletée der Diva geheftet und mit einem übergroßen Blumenschmuck kaschiert, während das Kabel, das Mikrophon und Aufnahmekabine verbindet, unter ihren Reifröcken versteckt wird.<sup>32</sup> So wird der Körper der Diva zwar auf die Technik zugerichtet, doch auch

dieser Versuch endet mit einer Störung, da nun das rhythmische Klopfen ihrer Herztöne den Dialog skandiert. Mit diesem asemantischen Pochen macht sich die kreatürliche ›Naturstimme‹ und gewissermaßen auch die (sonst durchgängig fehlende) ›innere Stimme‹ Linas ganz real bemerkbar – eine äußerst komische Verabschiedung jeglichen Versuches, das Schauspiel vor dem (neuen) Apparat als Verinnerlichung eines Veräußerlichten zu psychologisieren.

Der Regisseur, nun der Verzweiflung nahe, lässt das Mikrofon an Linas Schulter befestigen, doch auch diesen Ort kann Lina nicht als Adressaten ihrer Rede fixieren; zudem verändert sich nurmehr die Laut-Leise-Relation der von den Requisiten und Kostümen verursachten Geräusche, so dass es später wie ein Unwetter klingt, wenn Lina in bewährter Manier mit der Perlenkette klimpert, und jede Bewegung Don Lockwoods wird durch von seinem Seidenkostüm herverbrachte Knattertöne kommentiert.<sup>33</sup> Bei der Preview des Films, der Testleistung vor einem ausgewählten Publikum, werden solche Effekte als Inkongruenzen von Bild und Ton wahrgenommen, die jeden realistischen Raumeffekt dementieren und das Unternehmen Tonfilm als ganzes ins Lächerliche ziehen. Und schließlich betritt auch noch der Produzent das Set, stolpert entnervt über ›Linas‹ Kabel, zerrt daran – und lässt seinen Star mit schöner Slapstickwirkung auf den Hintern plumpsen. Die Initialszene im Tonfilmatelier endet so mit einem hysterischen Schrei:<sup>34</sup> Der ›stumme‹ Star wird überlaut zu Fall gebracht.

Man kann kaum glauben, dass unter solchen Mühen schließlich doch ein *100% Talking Picture* entstanden sein soll, das einem Testpublikum vorgeführt werden kann. Und in der Tat zwingt das Drehbuch um der Komik willen das Dream Team der Monumental zu früh aus der Deckung. Zum dritten Mal, nun schwarz-weiß auf der Leinwand, läuft die inzwischen hinreichend bekannte Szene ab – nun vollends zur Karikatur entstellt. Linas Stimme wirkt dank unzulänglicher Tontechnik im Kino noch schockierender als in der Wirklichkeit, und es kommt zum Aufstand der keineswegs stummen Gegenstände und Accessoires. Don schmachtet, wie ihm vorgeschrieben: »This Cupid himself that called me here, and I, smitten by his arrow, must come charging to your side, despite the threats of Madame la Guillotine.« Lina repliziert: »But the night is full of enemies.« und schlägt ihrem Verehrer mit dem Fächer zärtlich auf die Schulter, was einen lauten Knall verursacht: Die Szene verursacht schallendes Gelächter im Zuschauerraum, und ein Junge kommentiert laut und passenderweise im gleichen Soziolekt wie die tönende Diva: »Hey, Lina, wat'cha hittin' him with – a blackjack.« Doch auch Don kommt nicht ungeschoren davon: Durch den Schwulst des Drehbuchs irritiert hatte er, mit Zustimmung des nicht weniger unerfahrenen Regisseurs, beschlossen, sein Herz in ›Normalsprache‹ auszuschütten. Doch bei



»I can't make love to a bush!« – Krise bei der Inszenierung der galante Szene aus dem *DUELLING CAVALIER* im Tonfilmstudio.

der Preview erweist sich die Meinung, dass es auf den Text der Rede weniger ankomme als auf deren einwandfreie Reproduktion, als irrig. Als Don kniend Linas Hand ergreift und sich mit einem immer lauter werdenden »I love you, I love you, I love you« bis zu dem an ihrer Schulter versteckten Mikrophon zuletzt nach oben küsst, bleibt wiederum nur Gelächter – und der Kommentar aus dem Publikum: »Did someone get paid for writing that dialogue?«<sup>35</sup> Schließlich kommt es während der Preview auch noch zu einer Abspielpanne und dadurch zur Asynchronie von Bild- und Tonspur. Lina spricht in einer dramatischen Szene plötzlich mit der dunklen Männerstimme des Villains, der sie bedroht, »Yes, yes, yes«, während der Bösewicht, der sie entführen will, mit ihrer schrillen Frauenstimme ein »No, no, no« von sich gibt. So werden die vermeintlich authentischen Stimmen als ›Bauchrednerstimmen‹ offenbar, und damit ist der Konstruktionscharakter der Stimme im Medium Film vollends offensichtlich geworden: Zwar stellen Mikrophon und Kamera analoge Aufzeichnungssysteme dar, und man wird bald lernen, sie zu koordinieren. Aber hier wie dort genügt es nicht, den Apparat einfach ›draufzuhalten‹. Mediale Wirklichkeiten gibt es nicht; sie müssen, das ist eine ganz banale Erfahrung, immer gemacht werden, und sie müssen gut gemacht sein, damit sie ihre Adressaten erreichen.

### 3.2. DAS GENRE DER ERSETZUNGEN: »MAKE A MUSICAL!«

Auch in dieser Hinsicht erzählt *SINGING' IN THE RAIN*, wie es ›wirklich‹ war: Mit Einführung des Tons war die Zeit des Melodrams erst einmal abgelaufen. Das neu konfigurierte Medium brauchte den Ton zunächst nicht als Mittel der semantischen Bereicherung, sondern sehr viel eher als ›Totschläger‹ in einer Strategie der Überwältigung, das heißt: als Attraktion. Und möglichst viele tönende und

durch Ton ermöglichte Attraktionen konnte nur ein neues Genre bündeln: das Musical. Da es zudem darauf ankam, den Boom abzuschöpfen, bevor er zu Ende gehen konnte, war es absolut übliche Praxis, Fertig- und Halbfertigprodukte des Silent Cinema, die nun wie Blei in den Regalen lagen, ad hoc in Vehikel für Tanz- und Musikeinlagen zu verwandeln. Das ist auch der Weg, den die Monumental Pictures gehen:

THE DUELLING CAVALIER wird rücksichtslos in THE DANCING CAVALIER verwandelt, ein Musical, das sich mit moderner Rahmen- und historischer Binnenhandlung das Beste aller Welten anzueignen sucht. Die Idee dazu kommt von Kathy, aber es bedarf des Einfallsreichtum eines Cosmo Brown, um das Rezept für das Team Lockwood/Lamont brauchbar zu machen: Man braucht einfach nur die Störung der Ton-Bild-Synchronie zu jenem Normalfall zu machen, den sie technisch ohnehin darstellt. »Watch my mouth«: In Erinnerung an die Abspielpanne während der Preview verbirgt Cosmo Kathy hinter seinem Körper und lässt sie singen, während er vor Don nur stumm die Lippen bewegt. Mit diesem »positiven Einsatz des ›falschen‹ Tons«<sup>36</sup> wird aus der Katastrophe selbst das rettende Mittel hervorgerufen: Wenn Linas Stimme nicht zu brauchen ist, muss sie ganz einfach ersetzt werden, und anders als im wirklichen Leben ist dies auf der Leinwand ohne weiteres möglich. Bemerkenswert ist dabei nicht zuletzt, dass das Dubbing hier zum zweiten Mal nicht als illusionierender Tausch der Stimmen, sondern als »widernatürlich« komisches Verfahren der Gendersubstitution eingeführt wird.

Das selbstreflexive Spiel mit den verschiedenen Niveaus intradiegetischer Künstlichkeit, das für SINGIN' IN THE RAIN insgesamt charakteristisch ist, erreicht in den folgenden Einstellung, die Dubbing als technisches Verfahren in komprimierter Form ausstellen, einen Höhepunkt. Virtuos werden verschiedene Räume, Tonerzeuger und Aufnahmeapparate und schließlich auch Formate (Farbe und schwarz-weiß) ineinander geschaltet, um die aufwändige Prozedur zu zeigen, die nötig wird, damit die Leihstimme ins Körperbild des Stars eintreten und einen »natürlichen« Effekt hervorbringen kann – der aber für den Filmbetrachter schon deshalb nicht illusionierend wirken kann, weil er das Endprodukt, realistischere Weise, nur im distanzierenden schwarz-weiß auf der Leinwand im Vorführraum zu sehen bekommt. Noch einmal also naht sich Lockwood seiner Filmgeliebten, und endlich scheint aus ihr zu sprechen, was man hören will: eine Stimme, die im Einklang mit dem Körperbild ein Inneres (Gefühl) nach außen bringt.<sup>37</sup>

Hinter der Artifizialität des so gezeigten technischen Prozesses verbergen sich freilich, produktionsgeschichtlich, noch weitere Komplexitäten: Denn auch die Stimme der Lina Lamont ist eben kein Natur-, sondern ein Kunstprodukt, von Jean Hagen durch Entstellung ihrer »eigenen« Stimme hervorgebracht. Und Re-



Insenierung von Kathys geliebter Sing-Stimme bei der zur Nachvertontung dienenden Aufnahme des Song *Would You?*

Lina übt vor der play back machine die synchronen Lippenbewegungen zum Song *Would You?*

gisseur Stanley Donen berichtet über die »dubbing scenes« in *SINGIN' IN THE RAIN*: »We used Jean Hagen, dubbing Debbie [Reynolds], dubbing Jean [Hagen].«<sup>38</sup> Nicht Debbie Reynolds, die innerhalb der Narration des Films als Kathy Selden die Stimme Lina Lamonts synchronisiert, ersetzte also die »schreckliche« Stimme des Stummfilmstars. Vielmehr synchronisierte Jean Hagen selbst mit der ihr »eigenen« wohlklingenden, distinguiert artikulierenden Stimme in jenen Dubbing-Szenen, die im Synchronisationsstudio der Monumental Pictures in Szene gesetzt werden, die hier diegetisch von Kathy Selden nachvertonten gesprochenen Text-Passagen der Lina Lamont in der Tonfilmproduktion *THE DANCING CAVALIER*. Die Stimme Lina Lamonts in den Gesangsnummern, z. B. die im Film gezeigte Nachvertontung des Songs *Would you?*, wurde mit der Singstimme der Sängerin Betty Royce unterlegt.<sup>39</sup>

Nüchtern betrachtet, hat allerdings auch die durch diese Prozesse erzeugte Kompositgestalt einer sprechenden und singenden Lina Lamont keine wirkliche Überlebenschance im rekonfigurierten Medium. Linas Begabung besteht ausschließlich in der Fähigkeit, ihren Körper nach Maßgabe der Kamera so bewegen zu können, dass die Konventionen des Mediums eingehalten werden. Dieses Talent jedoch (und damit auch ihre Leinwandpräsenz) wäre auch dann veraltet, wenn sie eine wohlklingende Stimme ins Spiel zu bringen hätte, denn unter der Ägide des neuen Genres Musical sind zusätzliche Kompetenzen und Qualifikationen gefragt, die gleichsam von außen in den Film importiert werden. Der neue Star muss nicht nur sprechen, womöglich auch singen, er muss ganz anders spielen können, und dazu muss er, buchstäblich, einen anderen Körper einzusetzen

Ines Steiner

200

Scheiternde Zurichtung der ›undomestizierbaren‹ Stimme Linas auf den Apparat durch Sprech-Training: »Rrrround Tones! Rrrround Tones! Watch out for those dentalized ›d's‹ and ›t's‹ and those flat ›a's‹ ...«



haben. Man könnte meinen, dieser Aspekt werde in *SINGIN' IN THE RAIN* nur ex negativo, durch Vorführung der antiquierten Mechanik stummfilmisch kodierter Gesten thematisiert, und er verbleibe ansonsten insofern im Modus des Postulats, als implizit klar ist: Von nun an wird es nicht mehr genügen, vor der Kamera, mit Kathys polemischer Wendung, »a lot of dumb show« zu veranstalten. Von jetzt an muss man im Film nicht nur sprechen, sondern *spielen* können. Und auf den ersten Blick scheint es seltsam, dass der Film aus dieser Option, mit der die Love Story zwischen Kathy und Don plausibel hätte motiviert werden können, fast gar nichts macht. Dieser Eindruck ist jedoch nicht nur falsch, sondern er verfehlt zudem die spezifische Pointe des Musicals. Falsch ist er darum, weil – unbestreitbar – der gesamte Film gar nicht anders kann, als die seit 1927 im Medium erarbeitete Fähigkeit zum illusionierenden ›Acting‹ beständig vorzuführen: Was hier intradiegetisch als Wirklichkeitsebene denotiert wird, ist Produkt jenes Zusammenspiels von Verfahren, das in *THE DANCING CAVALIER* noch in den Kinderschuhen steckt. Das heißt aber auch, dass *SINGIN' IN THE RAIN* die Differenz zwischen diesem Medieneffekt und ›wirklicher Wirklichkeit‹ gar nicht unmittelbar artikulieren, sondern nur auf sie verweisen kann. Und genau hierin liegt die schönste Pointe der Idee, die Erinnerung an die Emergenz des Musicals selbst als Musical zu inszenieren.

Denn *SINGIN' IN THE RAIN* ist gerade nicht darauf aus, eine Geschichte von der Ankunft des Filmschauspiels im Gebiet der ernsthaften und ernstzunehmenden, weil: realistischen Künste zu erzählen. Die neuen Könner im Film sind nicht Shakespeare-Mimen, sondern einmal mehr Figuren von zweifelhafter Dignität – kurz: jene Vaudeville-Clowns und Artisten, auf deren Körperbeherrschung bereits der Stummfilm in seiner konstitutiven Phase gesetzt hat. Symbolisiert wird die neue Zeit des Mediums darum nicht durch die Fähigkeit zum gro-



ßen Monolog, in dem die Seele, sondern durch den Tanz, in dem der Körper spricht. Don Lockwood, der auf dem Set mit Lina stets nur seine (gutbezahlte) Arbeit gemacht hat – »well, it's a living« – kommt im Film eigentlich erst an, als ihm das Medium die Möglichkeit eröffnet, das zu zeigen, was er (anders als Lina) wirklich kann. Dieser Tatbestand wird in einer virtuoson Production Number des Films verbildlicht: dem gemeinsam mit Cosmo Brown aufgeführten logopädischen Slapstick. »Hollywood learns to talk«, verheißt *Variety*, und ergänzt: »Big Bonanza for Diction Coaches«. Das mag für Fälle wie die der auch in dieser Hinsicht unbelehrbaren Lina richtig sein. Und doch müsste es richtiger heißen: »Hollywood learns to dance.«<sup>40</sup>

Die Praxis des professoralen Sprachlehrers (Bobby Watson), der Don unterrichten soll, ist liebevoll mit Schaubildern ausgestattet, welche die korrekte Lippenstellung für die Artikulation der fünf Vokale illustrieren.<sup>41</sup> Doch Don vermag von vornherein nicht nur seine Filmzeilen, sondern auch die komplizierten Tongue Twisters nachzusprechen, mit denen ihn der Pädagoge traktiert, ganz gleich ob es sich um »Around the rocks the ruggeded rascal ran« oder um »Sinful Ceasar sipped his snifter, seized his knees and sneezed« handelt. So kann – Auftritt Cosmo – ohne weiteres auf die metadiegetische Ebene gewechselt werden, auf der die Production Number platziert ist.<sup>42</sup> Der der Nonsenspoesie ohnehin nahestehende Zungenbrecher »Moses supposes his toeses are roses, but Moses supposes erroneously – Moses he knowses his toeses aren't roses as Moses supposes his toeses to be« stellt den Übergang her. Durch Dons streng rhythmische und Cosmos synkopische Paraphrase wird er vollends entsemantisiert und zur Übersetzung in Gesang und den Rhythmus von Steptanzbewegungen tauglich: Was bleibt, ist eine Performance à la Gertrude Stein: »A rose is a rose, a Mose is a Mose, a toes is a toes«. Der Meister der korrekten (Aus-)Sprache wird zum Teil des Intérieurs degradiert und durch die Tänzer zum Lampenständer reduziert; gleichzei-



»Moses supposes ...«: Tongue Twister als Nonsens-Steptanz-Nummer in rhythmische Bewegung übersetzt.

tig verschwindet der Zungenbrecher in der Repetition des Namens, Moses, und löst sich schließlich in einen einzigen Buchstaben auf,<sup>43</sup> den laut in den Raum geschmetterten Vokal »AAAAAAH«, der zudem, in Form der entsprechenden Bildtafel, den Kopf des Sprachlehrers ersetzt. Das ist perfekt inszenierter Nonsens und gleichwohl mehr. Denn der biblische Moses, dessen Zunge so schwer war, dass ihm sein Bruder Aaron die öffentliche Stimme ersetzen musste, hat sich natürlich nicht zufällig ins Register logopädisch brauchbarer Tongue Twisters verirrt, und er wird auch nicht zufällig durch die perfekt synchronen Steptanz-Improvisationen der beiden Freunde wieder aus ihm befreit.<sup>44</sup> Nicht die Stimme zählt, sondern was aus ihr zu machen ist – nicht die Mitteilung, sondern der Rhythmus des Gesprochenen. A rose is a rose, a film is a film – und, vor allem: a musical is a musical.<sup>45</sup>

### 3.3. THE VOICE BEHIND THE CURTAIN – EIN FINALE UND KEIN ENDE

Auf die Kreisstruktur der durch die beiden Premierenevents gerahmten Handlung des Films wurde am Anfang bereits verwiesen. SINGIN' IN THE RAIN endet mit der spektakulären Premiere des Musicals THE DANCING CAVALIER, wiederum in Grauman's Chinese Theater, und mit dem Erfolg der Vorführung wiederholt sich auch, back stage, der Disput zwischen der vollends unkontrollierbar gewordenen Diva und den männlichen Protagonisten. Doch nun halten Cosmo, Don und R. F. Simpson die intrigante Leading Lady, die Kathy für alle Zeiten ihrer Stimme berauben möchte, nicht mehr davor zurück, mit ihrer »eigenen« Stimme vor das Publikum zu treten. Lina läuft willig ins Messer und produziert sich vor einem irritierten Publikum, dem der Wechsel von Stimmlage und Intonation natürlich auffällt. Aufgefordert, nicht länger zu reden, sondern ein Lied zu singen, schaufelt sich Lina vollends das Grab ihrer Karriere. Kathy wird genötigt, den falschen Live-Auftritt live, hinter dem Vorhang zu synchronisieren.<sup>46</sup> Doch mitten in der Darbietung des Songs *Singin' in the Rain* ziehen die drei Männer, vergnügt mitsingend und ihre Schadenfreude genüsslich auskostend, den Vorhang hoch, was einen überwältigend komischen Effekt erzeugt. Das Geheimnis um die »gestohlene« weibliche Stimme ist damit gelüftet, die Kompositgestalt der tönenden Lina Lamont auf der Bühne enttarnt. Kathy begreift die Peinlichkeit der Situation, singt aber tapfer weiter; Lina hingegen merkt zunächst nicht, was vor sich geht, obwohl das Publikum mit tosendem Gelächter reagiert. Erst als Cosmo Kathys Stelle hinter dem Mikrophon einnimmt,<sup>47</sup> läuft die düpierte Diva von der Bühne.

Der Star wird durch diese Zerstörung der zuvor im Film medientechnisch konstruierten Einheit von Körper und Stimme als ›unauthentisch‹ diskreditiert und seine Karriere damit beendet.

Linas synthetisierte Leinwand-Persona wird als Konstruktion enttarnt – aber was heißt das? Gewiss wird mit der Inszenierung der Stimmen in *SINGIN' IN THE RAIN* alles andere als gender-indifferent verfahren, und in der Tat erscheinen die Stimmen der weiblichen Darstellerinnen beliebig austauschbar, während der Leading Actor als ein ›geborener‹ Song- and-Dance-Man inszeniert wird, der souverän über eine, seine eigene, Sing- und Sprechstimme verfügen kann. Auch kann Kathy, mit dem Song *You are my lucky star*, nur dadurch als der ›wahre‹ Star des *DANCING CAVALIER* ins Rampenlicht gerückt werden, dass der Mann ihr diese Rolle zuschreibt. Und der Verdacht liegt nahe, der männliche Star habe den Medienumbruch auch dahingehende perfekt gemanaged, dass er die obsolet ge-

Stimmwechsel: ›The voice behind the curtain‹ wird sichtbar und hörbar gemacht – der Stimmenraub decouviert und vom Gelächter des Publikum kommentiert.



Ines Steiner

204

Neu kombiniertes Dream Team der tönenden Leinwand blickt auf seine auf dem Filmplakat eingefrorenen Star Images – aus dem Off singt die Stimme des Chors *You are my lucky star*.



wordene Leinwandpartnerin durch eine geeigneterere ersetzt. Auch dafür steht die Annäherung der beiden Stimmen im Gesang, in dem Don Kathy mit den Worten adressiert »You are my lucky star«, und Kathys von Glyzerintränen überströmtes Gesicht das Echo von sich gibt, »I saw you from afar«, mit der einem gerührten Publikum im Zuschauerraum die private und berufliche Vereinigung von Don und Kathy bekannt gemacht wird. Doch dieses erste Schlussbild wird sofort zum Filmplakat eingefroren. Mit der Überblendung der Großaufnahme von Dons Gesicht in sein gemaltes Starporträt findet ein Übergang vom Innenraum des Kinos in den Außenraum auf den Hügel über Hollywood statt. Die ›Geschichte‹, die das neue Traumpaar zusammengeführt hat, gerinnt zum Titel eines neuen Films namens SINGIN' IN THE RAIN. Und nun wird auch Kathys ins gemalte Starporträt fixierte Gesicht sichtbar und erhebt derart die neue Leinwandpartnerin Don Lockwoods zum neuen Star am Himmel Hollywoods. Einen Moment später steht das Paar in Alltagskleidung dann selbst vor dem von ihm entworfenen monumentalen Filmplakat mit den eigenen Starporträts, die für SINGIN' IN THE RAIN Reklame machen, um sich zu küssen. Mit dieser »mirror-mirror«-Konfiguration ist der Kreis zwar geschlossen. Es ist aber auch unmissverständlich signalisiert, dass zwischen Don und Kathy hier und Lina dort keine qualitative Differenz besteht, sondern nur ein Unterschied im Hinblick auf ihre aktuelle mediale Brauchbarkeit.

1 Vom ›Speechless Film‹ ist hier und in den folgenden Ausführungen nicht aus einer defizittheoretischen, sondern vielmehr einer differenztheoretischen Perspektive die Rede. Zudem ist zu berücksichtigen, dass der sogenannte ›Stummfilm‹ hinsichtlich der Präsentationsstrategien im Early und Silent Cinema niemals stumm war. Zu Beginn kommentierte ein Kinoerzähler die bewegten Bilder, später ersetzten Zwischentitel den Dialog. Die Filmmusik, Begleitung durch Klavier oder Wurlitzerorgel bzw. Orchester, nahm eine kommentierende Funktionsstelle ein. Der frühe Film setzte zudem, wie in SINGIN' IN THE RAIN explizit ausgestellt, während der Inszenierung vor dem Apparat sog. Mood Music ein, welche die Akteure beim Abruf ihrer Darstellungsleistungen vor der Kamera in ›Stimmung‹ versetzen sollte.

- 2 Rick Altman weist nach, dass die Tendenz zur selbstreflexiven Thematisierung der Genregeschichte in den Fünfziger Jahren für das Subgenre des Show Musicals bestimmend wird, und er bezieht sich dabei neben *THE BAND WAGON* (Vincente Minelli, 1953) explizit auf *SINGIN' IN THE RAIN* (1952). Diese beiden »MGM films in particular are often remembered for their role in this reflexive revival of the musical genre« (Rick Altman: *The American Film Musical*, Bloomington/Indianapolis 1987, S. 245). – Die Tendenz zur Selbstreferentialität des Musicals als Genre im Allgemeinen und die filmhistorische Selbstreflexivität von *SINGIN' IN THE RAIN* im Besonderen wird in der Forschung durchgängig hervorgehoben. Jane Feuer: *The self-reflective musical and the myth of entertainment*, in: Rick Altman (Hg.): *Genre. The Musical*, London 1981, S. 147–158; Gerald Mast: *Can't help Singin'*. *The American Musical on Stage and Screen*, Woodstock 1987; Beth Elliot Genné: *The Film Musicals of Vincente Minelli and the Team of Gene Kelly and Stanley Donen: 1944–1952*. Diss., University of Michigan 1984.
- 3 Zu den Tanz und Musiknummern des Films, die hier nicht einmal annähernd analysiert werden können, vgl. Stephan Prock: *Music Gender and Politics of Performance in SINGIN' IN THE RAIN*, in: *Colby Quarterly* 4 (2000), S. 295–318; Peter Wollen: *Singin' in the Rain* (BFI Classics), London 1992; Trixie Maraile Flügel: *Das Musical im Rahmen des klassischen Hollywood-Kinos*, Alfeld/Leine 1997 [zu *Singin' in the Rain*: S. 103–129].
- 4 Vgl. Sigrid Weigel: *Die geraubte Stimme und die Wiederkehr der Geister und Phantome*. Film- und Theoriegeschichtliches zur Stimme als Pathosformel, in: *Kunst- und Ausstellungshalle der Bundesrepublik Deutschland* (Hg.): *Der Sinn der Sinne*, Göttingen 1998, S. 190–206.
- 5 Rick Barrios: *A Song in the Dark. The Birth of the Musical Film*, New York/Oxford 1995, S. 105.
- 6 Verwendet wurden Arthur Freed/Herb Nacio Brown Songs wie *I've got a Feeling*, *You're a Foolin'*, *The Wedding of a painted doll* und *Should I?*. Zur Gestaltung der Sequenz vgl. auch: Flügel: *Das Musical im Rahmen des klassischen Hollywood-Kinos* (Anm. 3), S. 113. Diese Montagesequenz fungiert in *SINGIN' IN THE RAIN* auch als Einleitung zu der für das frühe Tonfilm musical charakteristischen Production Number *Beautiful Girl*, die am Set der Monumental Pictures von Sid Phillips (Tommy Farrell), einem jungen Regisseur, als dessen historisches Vorbild wohl Busby Berkeley diente, in Szene gesetzt wird. Vgl. dazu: Christoph Brecht/Ines Steiner: »Dames Are Necessary To(ols of) Show Buisness«. *Busby Berkeleys Production Numbers in der Multimedialität des Film-Musicals*, in: Claudia Liebrand/Irmela Schneider (Hg.): *Medien in Medien*, Köln 2002, S. 218–250.
- 7 Ausführlich: Barrios: *A Song in the Dark* (Anm 5); informativ auch der historische Quellen versammelnde Materialienband von Miles Kreuger (Hg.): *The Movie Musical from »Vitaphone« to »42<sup>nd</sup> street« as reported in a great Fan Magazine*, New York 1972.
- 8 Mary Anne Doane: *The voice in the cinema. The articulation of body and space*, in: *Yale French Studies* 60 (1980), S. 33–50; Kaja Silverman: *The Acoustic Mirror. The Female Voice in Psychoanalysis and Cinema*, Bloomington/Indiana 1988, S. 42–71; Kaja Silverman: *Dis-Embodiment The Female Voice*, in: Patricia Erens (Hg.): *Issues in Feminist Film Criticism*, Bloomington/Indianapolis 1990, S. 309–327; Martin Roth: *Women in Hollywood Musical. Pulling the Plug on Lina Lamont*, in: *Jump Cut* 35 (1990), S. 59–65. Spezifischer als die allgemeine feministische Kritik am Apparat setzt sich die jüngere Forschungsliteratur mit den Gender-Semantiken der Inszenierung der Stimme in *SINGIN' IN THE RAIN* auseinander; zur Sing- und Sprechstimme in diesem Film aus gender-spezifischer Perspektive vgl. vor allem: Prock: *Music Gender and Politics of Performance in SINGIN' IN THE RAIN* (Anm. 3), S. 295–318.
- 9 Die Auseinandersetzung mit medientheoretischen Aspekten der Synchronisation wird im vorliegenden Band an anderer Stelle geführt; vgl. die Beiträge von Michel Chion und Gereon Blaseio.
- 10 Hal Erickson: *SINGIN' IN THE RAIN*. Plot Synopsis, in: *All Movie Guide*, unter: <http://www.allmovie.com> (30.04.03).
- 11 Diese skeptische und ablehnende Haltung der Branche gegenüber dem Verfahren des Tonfilms wird vom Film selbst in einer Szene zum Thema gemacht. Während der Premierenparty präsentiert der Produzent in seinem Heimkino einen kurzen Streifen, einen frühen Tonfilm, als kuriose medientechnische Attraktion. Zu sehen ist nichts anderes als das innovative Verfahren, das hier selbst ausgestellt wird. Simpson verspricht den Gästen, »this is going to hand you a lot of laughs«. Die Leinwand öffnet sich, und die Zuschauer sehen den Erfinder selbst, der dem komischen Typus des Mad Professor entspricht, ein »madman«, der den Produzenten seit Monaten bedrängt. Das Rascheln mit Notizblättern in seinen Händen klingt »like a thunderstorm«. Das Verhältnis von Gesicht, das in einer unvorteilhaft nahen, perspektivisch aus der Untersicht verzerrenden Großaufnahme repräsentiert wird, und Stimme erzeugt einen grotesken Effekt. In langsamer Sprechweise mit einer übertrieben genauen Diktion und lauter Stimme, als würde er zu tauben Personen sprechen,

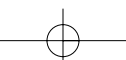
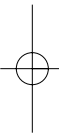
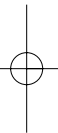
- adressiert er sein (für ihn: imaginäres) Publikum. Besetzt wurde der ›Erfinder‹ mit dem aus Deutschland stammenden Darsteller Julius Tannen, daher spricht er mit einem typengerecht fremdlichen Akzent; die Repräsentation seiner Stimme ist technisch noch unzureichend, denn die Wiedergabe produziert Echos. Im Zuschauerraum wird die Präsentation simultan kommentiert, es wird gefragt, wer da hinter der Leinwand spreche. Niemand nimmt die technische Innovation, die mit den Worten des Professors präsentiert wird, ernst. »This is a demonstration of a talking picture. Notice – it is a picture of me – and I am talking! Note how my lips and the sound issuing from them are synchronized together in perfect unison.« Die versammelten Repräsentanten der Branche mokieren sich über diesen Tonfilm als eine technische Kuriosität, bezeichnen es als »nur ein Spielzeug« oder als »vulgär«. Niemand nimmt das neue technische Verfahren ernst, erkennt eine Chance oder eine Bedrohung darin, und diese hier reinszenierte Unterschätzung des Tonfilms entspricht exakt der historischen Situation. Der JAZZ SINGER wird zu diesem Zeitpunkt von Warner Brothers gerade erst produziert und man prognostiziert »they'll loose their shirts«.
- 12 »If you have seen one, you have seen them all« kommentiert Cosmo, und Don gibt sich abgebrühter, als er im Grunde ist: »Well, it's a living«.
- 13 Betty Comden/Adolph Green: SINGIN' IN THE RAIN. Story and screenplay by Betty Comden and Adolph Green, with an introduction by the authors, London <sup>3</sup>1997, S. 28.
- 14 Nur dem ersten Anschein nach gilt diese Regel für die von Don in der Exposition erzählte, von Bildern konterkarierte, Success Story nicht. Das Bild, das er von sich ›nach außen‹ projiziert, ist und bleibt tatsächlich falsch. Für den Zuschauer des Films wird jedoch seine Rede als uneigentlich erkennbar, das heißt: Don lügt nicht, er spricht lediglich (selbst-)ironisch, und seine Erinnerungsbilder sind ›natürlich echt«.
- 15 Comden/Green: SINGIN' IN THE RAIN (Anm. 13), S. 30.
- 16 Durch Nahaufnahmen von der wartenden Menge werden die Medien der zeitgenössischen Fan-Kultur ausgestellt. So hält eine junge, aufgeregt kaugummi-kauende Frau ein Fan-Magazin hoch, »Screen Digest – 25 cents«, auf dessen Cover das Starporträt des zur Premiere erwarteten Dream-Teams Don Lockwood und Lina Lamont zu sehen ist. Die darüber plazierte spekulative Schlagzeile, »Lockwood and Lamont – Reel Life or Re-al Life Romance?«, nimmt die später folgenden, indiskreten Fragen nach dem Privatleben des romantischen Paares der ›stummen‹ Leinwand vorweg. Am Ende des Films werden die Gesichter des neuen Dream-Teams der tönenden Leinwand, Don Lockwood und Kathy Selden, wieder zu Starporträts auf einem Kinoplatat eingefroren, schon hier sei auf die zirkuläre Struktur der Narration von SINGIN' IN THE RAIN verwiesen.
- 17 Die Menge der das Spektakel betrachtenden Fans entäußert sich ausschließlich im ›Schrei‹. Auch beide Frauenfiguren, insbesondere Lina Lamont, artikulieren sich kreischend. Dadurch, will man durchaus nicht unproblematischen Thesen Kaja Silvermanns zu gender-spezifischen Semantiken der stimmlichen Artikulation folgen, entstehe ein effeminierender Effekt. Doch der Film zeigt, dass der Schrei auch männlichen Figuren zugeschrieben wird: Eine Großaufnahme zeigt einen jüngeren Mann im Publikum, der seinem Begehren nach dem Zip Girl, das in Shimmie-Schritten neben einem greisenhaften Sugar Daddy ins Kino schreitet, schreiend Ausdruck gibt: »Zelda! Ohhhhh! Zelda!« Man müsste die Zuschreibungen folgendermaßen modifizieren: Durch solche auf den Namen des angerufenen Stars und asemantisches Gestammel restringierte Schreie von Männern und Frauen wird die Menge der Fans insgesamt als »hysterisch« semantisiert, wodurch ein effeminierender Effekt in der Charakterisierung des ›Massenpublikums‹ freigesetzt wird. Mit der Anrufung der adorierten Stars aus der Ferne wird bereits in der Exposition der erst am Ende des Films präsentierte Song, »You are my lucky star / I saw you from afar ...« anzitiert. In anderen Szenen, etwa bei den Premieren und der Preview, erweist sich das Publikum in seinen pointierten, kritischen Kommentaren gegenüber den auf der Leinwand präsentierten Filmen als durchaus intelligent und eloquent.
- 18 Diese Figur ist nach dem Vorbild von Louella Parsons modelliert, die gemeinsam mit ihrer Konkurrentin Hedda Hopper in ihren von der Branche gefürchteten, aber viel gelesenen Kolumnen in Tagespresse und Fan-Magazinen eine nicht zu unterschätzende Macht über ganz Hollywood ausübte.
- 19 Joseph A. Casper: Stanley Donen, Metuchen, NJ 1983, S. 45.
- 20 Lina Lamont wird im Kontext der zeitgenössischen Blondinen-Semantik als Dumb Blonde charakterisiert und der dunkelhaarigen, jüngeren, unscheinbareren aber auf ›authentischere‹ Effekte hin konstruierten Kathy Selden konfrontiert.
- 21 »From now on we see that Don does all the public-appearance talking for the two of them, while

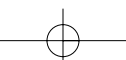
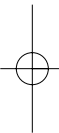
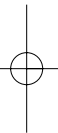
- Lina just smiles, with a certain false graciousness, covering her annoyance.« Comden/Green: SINGIN' IN THE RAIN (Anm. 13), S. 4.
- 22 Flügel beschreibt das Verfahren der Rahmung der ›verlogenen‹ Rede Dons nach dem Motto »dignity always dignity« präzise und ausführlich. Vgl.: Flügel: Das Musical im Rahmen des klassischen Hollywood-Kinos (Anm. 3), S. 106–107.
- 23 Der feministisch inspirierte Einspruch, demzufolge Don seine Karriere in gewisser Weise Lina verdanke, die dann zugunsten einer tonfilmisch besser passenden ›Zweitfrau‹ weggeworfen werde, greift zu kurz, weil er diese doppelte Kodierung der Figuren ignoriert. Vgl.: Roth: Women in Hollywood Musical (Anm. 8), S. 59–65.
- 24 Man mag wohl glauben, dass Lina sich als integrale Verkörperung weiblicher Schönheit imaginiert. Und man kann auch akzeptieren, dass eine klassische ›dumb blonde‹ auf einmal intrigantes Potential entwickelt. Aber man tut sich schwer damit, dies beides in einem Plot zu akzeptieren. Hier opfert der Film seine Subtilität zugunsten der Deutlichkeit und eines wenig überraschenden Überraschungseffekts.
- 25 Flügel: Das Musical im Rahmen des klassischen Hollywood-Kinos (Anm. 3), S. 108.
- 26 »Contemporary audiences often fail to recognize that the footage shown from *The Royal Rascal* in which ›Don Lockwood‹ leaps and duels in the manner of Douglas Fairbanks – in yet another illusion – is actually from Kelly's *THE THREE MUSKETEERS* (1947) originally a color film but here in a black and white print into which newly made shots with Kelly and Jean Hagen have been intercut.« James Card: ›More than meets the Eye‹ in: *Singin' in the Rain* and *Day For Night*, in: *Literature-Film-Quarterly* 12/2 (1984), S. 87–95 (hier: S. 89).
- 27 Comden/Green: SINGIN' IN THE RAIN (Anm. 13), S. 11.
- 28 Stephan Prock hat die Inszenierung der Stimme unter diesen Aspekten in SINGIN' IN THE RAIN jüngst recht präzise analysiert. Vgl.: Prock: Music Gender and Politics of Performance in SINGIN' IN THE RAIN (Anm. 3), S. 295–318.
- 29 Roth: Women in Hollywood Musical (Anm. 8), S. 60.
- 30 Linas vollends instrumentelles Verhältnis zu ihrem Körper wird deutlich, wenn sie ihren Mund, immerhin das Organ, mit dem die Stimme artikuliert wird, als einen »kisser« bezeichnet.
- 31 Don empört sich über die irritanten Manöver der Stummfilmdiva mehrfach dadurch, dass er Lina als ›falsche Schlange‹ (Rattlesnake, Boa Constrictor) bezeichnet.
- 32 Der durch Kostüm und Maske ohnehin schon als ›porzellanpuppenhaft‹ konstruierte und semantisierte Körper der Darstellerin Lina Lamont wird hier offensichtlich durch das Kabel, den ›Faden‹ eines tontechnischen ›Marionettenspielers‹, mit dem Apparat vernetzt; das Mikrofon dringt geradezu penetrierend in den als Artefakt für das Schauspiel vor der Kamera hergestellten Körper ein, wenn es ihr zwischen die Brüste gesteckt wird.
- 33 Zu den filmhistorischen Vorbildern der bei der Preview der Tonfilmfassung von *THE DUELLING CAVALIER* reinszenierten Pannen, siehe die Ausführungen von Flügel in: Flügel: Das Musical im Rahmen des klassischen Hollywood-Kinos (Anm. 3), S. 118.
- 34 Die Analogie des slapstickhaften Szenenabschlusses zur Szene, in der Kathy mit den Worten »this is what I learned from moving pictures« bei der Premierenfeier Lina eine Torte ins Gesicht wirft, wodurch das Gesicht der kreischenden Diva zur schrillbunten Fläche entstellt wird, ist unverkennbar.
- 35 Mit dieser verulkenden Rezeption des wiederholten »I love you« durch das Publikum haben die Drehbuchautoren eine filmhistorische Legende vom Scheitern eines populären Stars der stummen Leinwand, namentlich John Gilbert eingearbeitet (der mit Greta Garbo einst ein Traumpaar bildete). Vgl. dazu Flügel: Das Musical im Rahmen des klassischen Hollywood-Kinos (Anm. 3), S. 118.
- 36 Vgl. Flügel, die hier einen Befund von Altman zitiert, in: Flügel: Das Musical im Rahmen des klassischen Hollywood-Kinos (Anm. 3), S. 121.
- 37 Flügel hat die Szenen, in denen der filmtechnische Transfer minutiös vorgeführt wird, ausführlich beschrieben und kommentiert. Der Film stellt aus, wie Kathys Singstimme die Gesangspassage Linas in dem Song *Would You?* während der remodellierenden Produktion des Musicals *THE DANCING CAVALIER* nachvertont. Durch eine Überblendung wird sicht- und hörbar, dass Lina nun ihrerseits vor der Playback-Maschine, die den Gesang Kathys mechanisch reproduziert, schrill singend die lippensynchronen Bewegungen zum Song einübt. Dann wird diese Szene am Set mit der zuvor aufgenommenen und aus dem Off eingespielten Musik und Singstimme nochmals aufgenommen, wobei die kostümierte Lina ›stumm‹ die entsprechenden Körperbewegungen ausführt, und schließlich wird die gelungene Vereinigung von Linas Körper und Kathys Stimme als fertiges Konstrukt auf der

- schwarz weißen Leinwand im Vorführraum der betrachtenden Begutachtung ausgesetzt. Vgl. Flügel: Das Musical im Rahmen des klassischen Hollywood-Kinos (Anm. 3), S. 121–125.
- 38 Stanley Donen zum Making of *SINGIN' IN THE RAIN*, zitiert nach: Hugh Fordin: *The World of Entertainment*, New York 1975, S. 358. Um über die »tricky dubbing scenes« zu berichten, hebt James Card im Rekurs auf Hugh Fordin exemplarisch zwei Szenen hervor, in denen die komplizierte Ersetzung auch hörbar wird: »When Debbie Reynolds is shown dubbing ›Would you?‹ for Jean Hagen, the voice issuing from her mouth does not really sound like that of the girl who sang ›Good Morning‹ a few minutes earlier and in fact it is not. Here, her voice was dubbed by Betty Royce. Also, when Kathy is shown dubbing dialogue – the line is ›Our love will last till the stars grow cold‹ – we are allowed another demystification: we see the original image with Lina's voice, then watch as it is replaced by Kathy's voice. But again, the voice synchronized with the lip movements does not sound like the voice of Debbie Reynolds. ›We used Jean Hagen, dubbing Debbie, dubbing Jean‹, Stanley Donen has explained, the actress using first her ›Lina voice‹, then her own cultured speech, a perfect case of the dubber being dubbed, like the stunt man having someone perform his stunt or simply another illusion in a film of and about film illusions, like the color *The Three Musketeers* masquerading in black and white as the *Royal Rascal*.«, in: James Card: ›More than meets the Eye‹ (Anm. 26), S. 92.
- 39 Kathy Selden, die aufgrund einer Intrige Lina Lamonts in den Credits des produzierten Musicals *THE DANCING CAVALIER* nicht genannt wird, teilt dieses Schicksal mit zahlreichen Sängerinnen in der Geschichte des amerikanischen Film musicals, die Gesangsszenen unbegabter Darstellerinnen nachvertont haben, deren geliehene Singstimme aber »unsichtbar« gemacht wurde, indem sie in den Credits verschwiegen wurde, was auch ökonomische Konsequenzen hinsichtlich der Zweitverwertung des Soundtracks hatte. Vgl. dazu ausführlich: Marsha Seifert: *Image/Music/Voice. Song Dubbing in Hollywood Musicals*, in: *Journal of Communication* 45/2 (1995), S. 44–63.
- 40 In der Kontrast erzeugenden Parallelmontage der Sprecherziehungsszenen, in der Linas Stimme sich als »unkultivierbar« erweist, während Don denselben Satz, »And I can't stand him«, den sie trotz allen Sprechtrainings nicht in elaborientem Englisch zu artikulieren vermag, mühelos und ohne Übung spricht, wird nebenbei auseinandergelegt, was in der Sprecherziehungs-Szene von George Bernhard Shaws *Pygmalion* ineins gesetzt war: Das Naturkind Don hat es immer schon begriffen; das Naturkind Lina dagegen lernt es nimmermehr.
- 41 Im Set Design wird auch durch eine Farbtafel mit larynologischen Grafiken die Stimme »sichtbar« ausgestellt.
- 42 Die stummen Grimassen, mit denen Cosmo den Logopäden persiflierend synchronisiert, präfigurieren bereits die Einführung des Dubbing – ganz ähnlich wird Cosmo Kathys Song mit seinen Lippen begleiten.
- 43 Und zwar in den ersten Buchstaben des Alphabetes.
- 44 Die Tänzer hüllen sich für einen Moment in die »hebräisch« gestreiften weiß-braunen Fenstervorhänge und ähneln in diesem improvisierten Kostüm Akteuren in einem Sandalenfilm.
- 45 Die performative Übersetzung des Rhythmus der gesprochenen Sprache in den Rhythmus tänzerischer Bewegungsabläufe ist ein bekanntes Verfahren, zu denken ist in diesem Kontext auch an die Sprecherziehungs-Szene im Musical *MY FAIR LADY*. Dieses auf Shaws *Pygmalion* beruhende Musical übersetzt den Zungenbrecher »The rain in Spain ...« in einen Flamenco.
- 46 Für aufmerksame Zuschauer, die zuvor derart ausführlich mit den Verfahren des dubbing vertraut gemacht wurden, könnte durchaus offensichtlich werden, dass auch Debbie Reynolds als hinter dem Vorhang synchron singende Kathy Selden nicht mit ihrer »eigenen« Stimme singt, sondern auch nur dubbt.
- 47 Hinsichtlich der Gender-Konfiguration des Films erscheint auch dieser letzte Austausch bezeichnend: Cosmos Stimme verdrängt Kathy vom Mikrofon, die ihrerseits Linas Stimme ersetzt und damit Dons Erfolg gesichert hat – aber Cosmo ist und bleibt, trotz der Finalnummer, Dons eigentlicher Wunschpartner. Die Paarkonstellation ist von Beginn an im Topos des Male Bonding rekonfiguriert; Cosmo nimmt in dem so beschworenen Männerbund jene Position der treuen Gefährtin auf dem Weg zum Erfolg ein, die Dora Bailey Lina und die Story Kathy anträgt.



### III. STIMM-ART





**Daniel Gethmann**  
**DAS ZITTERN DER LUFT.**  
**DIE ERFINDUNG DER MECHANISCHEN STIMME**

Eine Geschichte der Stimme gibt es nicht. Stimme und Klang, den sie prägt, ist gemeinsam, dass sie nur anwesend sein können, wenn sie verschwinden. Beide verstoßen gegen die Konventionen der Diachronie, wie Jean Paul beobachtet hat:

Wenn die Töne sprechen, können wir nicht unterscheiden, ob sie unsere Vergangenheit oder unsere Zukunft aussprechen. Wir hören ferne Tage, weggegangne und herkommende; denn beide sind fern, und wir müssen zugleich uns erinnern und uns sehnen. Denn kein Ton hat Gegenwart und steht und ist; sein Stehen ist nur ein bloßes Umrinnen im Kreise, nur das Wogen einer Woge.<sup>1</sup>

Stimmen klingen, kommunizieren, werden reproduziert, mechanisch erzeugt, synthetisiert und vieles mehr. Man kann also in Untersuchungen über die Stimme und den Klang, den sie hervorbringt, eine Geschichte der Techniken schreiben, die zur Kommunikation, Reproduktion und Synthetisierung von Stimmen eingesetzt werden, und wird feststellen, dass es die eine Geschichte der Stimme nicht gibt. Statt dessen wurden unterschiedliche Konzeptionen der Stimme und ihrer Werkzeuge entworfen, die Techniken der Stimmreproduktion und Stimmerzeugung außerhalb des menschlichen Körpers imaginierten.<sup>2</sup> Sobald diese Techniken Apparate generierten, basierten sie jedoch auf Studien, die die Sprache und deshalb auch die Stimme am Menschen selbst erforschten und nicht außerhalb seines Körpers. Bei der Analyse der menschlichen Sprache und Stimme wurden die Organe der Stimmerzeugung untersucht,<sup>3</sup> wobei dann bekanntlich die Messgeräte zur Aufzeichnung und Speicherung der Augen und Ohren unsere heutigen technischen Medien präfigurierten. Die Erkenntnisse der Physiologie waren die Voraussetzung für die weitere Entwicklung, sie brachten sogar Sprechmaschinen zur Simulation der menschlichen Stimme hervor.

Allgemein lassen sich zwei unterschiedliche Verfahren der Konstruktion von künstlichen Stimmen unterscheiden, die historisch aufeinander folgten: das eine suchte nach einer zunächst mechanischen Nachbildung des menschlichen Sprachapparats, das andere nach einer Erzeugung seiner lautlichen Eigenschaften. Den mit beiden Verfahren verbundenen Schrecken, den dieses akustische Phänomen einer Trennung der Stimme vom Körper auslöste, deutete François Rabelais

durch seine imaginäre Erfindung der Klangspeicherung in *Gargantua und Pantagruel* (1532) an. Während einer Schiffsfahrt auf offener See hörte die Besatzung nämlich plötzlich Stimmen:

Und wirklich, je länger wir hinhorchten, desto deutlicher vernahmen wir Stimmen, bis wir zuletzt sogar ganze Wörter heraushören konnten. Wir waren darüber sehr erschrocken und hatten wohl Grund dazu; denn wir sahen niemand und hörten doch die verschiedensten Laute, die von Männern, Frauen, Kindern und Pferden herzurühren schienen.<sup>4</sup>

Das Schiff war in eine eisige Gegend gekommen, in der ein großes Gefecht stattgefunden hatte, »der ganze Lärm der Schlacht gefror damals.«<sup>5</sup> Das einsetzende Tauwetter enteiste den Klang des Krieges und führte darüber hinaus die ehemals »gefrorenen Wörter« als Lautpoesie in die Literatur ein. Denn ohne akustische Aufzeichnungs- und Wiedergabemedien blieben die aufgetauten Laute weiterhin an den Buchstaben und die alphabetische Ordnung gefesselt und lasen sich so: »hin, hin, hin, hin, his, ticque torche, lorgne, brededin, brededac, frr, frrr, frrr, bou, bou, bou, bou, bou, bou, bou, bou, tracc, trac, trr, trr, trrr, trrrrr, On, on, on, on ououououou.«<sup>6</sup> Der Schrecken, den diese Laute verbreiteten, wurde bereits von Rabelais auf die Erkenntnis zurückgeführt, dass die Stimme nicht länger ihrer sterblichen Hülle bedurfte, dass man also ›Männer, Frauen, Kinder und Pferde‹ hörte, ohne auch deren Körper zu sehen.

Außer der hier angesprochenen Imagination der Klangspeicherung sollte eine Trennung des Klangs von seiner Schallquelle noch durch eine weitere Technik realisiert werden: die Idee der Klangsynthetisierung. So befanden sich in den Forschungsstätten des Hauses Salomon in Francis Bacons: *Nova Atlantis* (1638) »akustische Werkstätten, wo wir alle Töne und ihre Erzeugungsarten untersuchen und erforschen. [...] Wir erzeugen alle artikulierte Laute und Buchstaben und ahmen sie künstlich nach, ebenso alle Stimmen und Laute der Säugetiere und Vögel.«<sup>7</sup>

Die Konzepte, Stimme aufzuzeichnen oder künstlich zu erzeugen, sind bis zum Phonographen Edisons, der häufig als »Talking Machine« bezeichnet wurde, diskursiv nicht streng unterschieden worden, durchaus im Sinne ihrer gemeinsamen Realisierung einer Ablösung des Schalls von seiner Schallquelle. Um Sprechmaschinen zu konstruieren,<sup>8</sup> musste zunächst die Fähigkeit zum Sprechen auch einer Maschine zugebilligt werden, was einen Paradigmenwechsel darstellte, denn die Stimme war lange Zeit gleichbedeutend mit der Sprache und insofern das bedeutsame Unterscheidungskriterium vom Tier. Das sprechende Wesen als

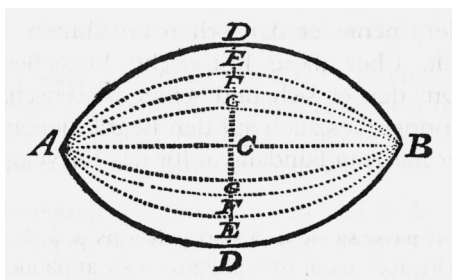
grundlegende Charakteristik des Humanen in der Aufklärung blieb mit der Erfindung der Sprechmaschinen jedoch keine eindeutige Zuschreibung mehr. Wenn Maschinen sprechen konnten, dann ließ sich die Stimme zum empirischen Modell der Studien des Lebens machen, worin dann allerdings auch Maschinen ihren Platz beanspruchen konnten. Doch auch der Körper war bereits getreu des mechanistischen Weltbildes der Physik unter das Zeichen der Maschinisierung geraten und galt ebenfalls seit Descartes als ein ausgeklügelter Mechanismus. Die Theorie des Menschen als Automaten<sup>9</sup> stellte explizit der Militärarzt Julien Offray de La Mettrie im Jahre 1748 auf. Der Leib erschien wie jede andere Materie als eine Maschine, die große Ähnlichkeit mit den damals bekannten Maschinen aufwies: »Le corps n'est qu'une horloge.«<sup>10</sup> Wenn der Körper eine Uhr war, konnte auch der Mensch nichts anderes sein: »que l'homme n'est qu'un animal, ou un assemblage de ressorts, qui tous se mentent les uns par les autres.«<sup>11</sup> La Mettrie berief sich bei seiner Auffassung vom Menschen als einer Maschinerie von Triebfedern, die sich gegenseitig spannen, ausdrücklich auf den berühmten Automatenkonstrukteur Jacques de Vaucanson und seine mechanischen Wunder aus den 30er Jahren des 18. Jahrhunderts: den Flötespieler, den Bläser und Trommler und die berühmte Ente, die Nahrung aufnehmen und wieder ausscheiden konnte. Schließlich folgte La Mettrie aus seinen Überlegungen, Vaucanson »eût dû en employer encore davantage pour faire un parleur: machine qui ne peut plus être regardée comme impossible.«<sup>12</sup>

Dieser vorausgesagte Nachbau der menschlichen Sprechfähigkeit konnte nur unter zwei Voraussetzungen geschehen: einmal musste die Stimme als ein mit anatomischen Modellen nachzubildender Gegenstand angesehen werden und zum anderen eine Änderung der herrschenden Schulmeinung darüber eintreten, was das Sprechen hervorbrachte und wodurch die Stimme in der Lage war, Sprachlaute zu bilden. Bis zum Jahre 1700 war man der Auffassung, die der römische Arzt Galenos aus Pergamon, genannt Galen um 165 n. Chr. ausdrückte, dass die Stimme nämlich als »Schlag der Luft« entstehe und mit einem Aulos (antiken Blasinstrument) vergleichbar sei. Statt nun die Vibrationen des Rohrblatts innerhalb dieses Instruments in Analogie mit denen der Stimmlippen zu setzen, beharrte Galen auf Schlägen der Luft auf Kehlkopf (Larynx) und Luftröhre (Trachea).<sup>13</sup> Das Medium des Hörens war die Luft. Man stellte sich vor, dass sie als zusammenhängende und einheitliche Masse bewegt werde; dazu bedurfte es in der damaligen Vorstellungswelt eines größeren Schlags statt leichter Vibrationen. Trotz seiner in vielen Vivisektionen gewonnenen zutreffenden Erkenntnisse über den Stimmapparat kam Galen daher zu der bereits von Aristoteles vorgezeichneten Auffassung, dass ein Schlag der Ursprung der Stimme sein müsse. Aristoteles schrieb:

Dass aber überhaupt Töne und Geräusche entstehen, ergibt sich dadurch, dass entweder Körper oder Luft auf Körper auftreffen, nicht indem die Luft geformt wird, wie manche glauben, sondern dadurch, dass sie in (den Körpern) entsprechender Weise bewegt wird: sich zusammenzieht, ausdehnt und angehalten wird, sowie ferner zusammenstößt aufgrund der Schläge, die bei der Atemluft und bei den Saiten entstehen.<sup>14</sup>

Abb. 1

Öffnungen der Glottis nach Dodart



Als Stimmorgane benannte Aristoteles die Lungen, die Luftröhre und den Mund. Um das Jahr 1700 findet eine folgenreiche Veränderung in der Ansicht darüber statt, welches Körperteil für die Erzeugung der Stimme verantwortlich sei. Der Physiologe Denis Dodart widersprach nämlich der bis dahin geltenden Schulmeinung und stellte die Theorie auf, dass die Glottis (auch Stimmritze, das aus den beiden Stimmlippen bestehende Stimmorgan im Kehlkopf) für die Produktion sämtlicher Töne des Menschen verantwortlich sei.<sup>15</sup> Die Glottis war jedoch schwierig zu untersuchen, da sie der Beobachtung nicht zugänglich war und auch eine Leichensezierung zwar ihre physiologische Struktur erkennen ließ, aber nicht ihre Funktion einsichtig machte. Dodart vertrat im Widerspruch zu Galen die für die Erbauer von Sprechmaschinen wesentliche Auffassung, dass kein Blasinstrument die Stimmerzeugung repräsentieren könne, sondern diese nur von der jeweiligen Spannung der Stimmlippen und kaum messbaren Veränderungen in der Geometrie der Stimmritze abhängen. Damit erschien die Glottis als eine Art Ventil, oder »Blasinstrument, an dessen variabler Öffnung schwingende Lippen eine Luftsäule zum Tönen brächten.«<sup>16</sup> Also ein Blas- mit Funktionen eines Saiteninstrumentes; die Trachea erhielt zu dieser Zeit in der Stimmphysiologie nur mehr die Funktion der Luftzufuhr.<sup>17</sup> Eine abweichende Auffassung darüber, welches Instrument der Stimme analog sei,<sup>18</sup> vertrat der Anatom Antoine Ferrein, dessen Experimente 1741 veröffentlicht wurden und ebenfalls die wesentliche Bedeutung der Glottis nachwiesen.<sup>19</sup> Er hatte in seinen Untersuchungen zunächst verschiedenen Tieren entnommene Kehlköpfe angeblasen und später als Marinearzt diese Experimente an Leichenteilen fortgesetzt: »Le larinx du cadavre

répondit par un éclat qui étonna les assistants, & c'est, je pense, la première fois qu'on ait vû pareil phénomène.<sup>20</sup> Seiner Auffassung nach erzeugten die Stimmbänder nämlich durch Vibration den Klang der Stimme, wobei die ausströmende Atemluft die Funktion des Bogens bei einem Streichinstrument (Stimmband) habe, wie er sagte.<sup>21</sup> Die Stimmbänder gaben in dieser neuen Auffassung von der Stimme den Ton, sie wurden demnach analog zu Geigensaiten konzeptualisiert, gleichzeitig erfasste Ferrein das menschliche Stimmorgan so genau, dass es eigentlich keiner musikalischen Analogie mehr bedurfte. In der ersten Bauanleitung einer Sprechmaschine überhaupt (1741), die direkt aus seinen anatomischen Studien von Leichenteilen und Tierkadavern, hier insbesondere seinen Studien des Kehlkopfes hervorging, notierte Ferrein:

Um die Kehle zum Tönen zu bringen, muß man die Stellknorpel zwischen Daumen und Zeigefinger gegeneinanderdrücken und von unten in die Luftröhre mittels eines Rohrs von mindestens 9–11mm Durchmesser hineinblasen. [...] Um die meisten der im Mémoire beschriebenen Experimente bequemer durchführen zu können, fertige man eine sehr einfache Maschine an, die aus einem kleinen Holzbrett besteht; auf ihm sind drei Stäbe von 8 Zoll Länge senkrecht befestigt, durch die bewegliche Wirbel ähnlich denen von Violen gesteckt sind. Ich hänge die Kehle mit Hilfe dreier starker Fäden auf, die an jeweils einem Ende auf die Wirbel gewickelt sind. [...] Diese Fäden dienen dazu, die Stimmbänder auseinanderzuziehen, sie gleichmäßig oder ungleichmäßig zu spannen und sie unter dem jeweiligen Maß an Spannung zu halten. [...] Wenn ich die Schwingungen anhalten und den Ton eines der Stimmbänder oder eines seiner Teile zum Verstummen bringen will, berühre ich diesen Teil mit kleinen Pincetten aus Weißblech oder aus Holz (sie verletzen weniger leicht als die anderen) und auf diese Weise bediene ich mich der Kehlen von Menschen, Hunden und Schweinen.<sup>22</sup>

Allerdings trennte Ferrein den Kehlkopf vom Ansatzrohr, also dem Raum zwischen Stimmritze, Nasenraum und Lippenöffnung, daher konnte auch die Entdeckung einer gemeinsamen Funktion von Kehlkopf und Ansatzrohr für die Stimmerzeugung von ihm mit dieser Maschine nicht gemacht werden. Den Stand des Wissens von dem Zusammenhang zwischen Körper und Stimme fasste der Philosoph und Mathematiker Christian Freiherr von Wolff im Jahre 1725 so zusammen:

Unter die Bewegungen im Leibe, welche dem Willen der Seele unterworfen sind und die mit ihren Verrichtungen übereinstimmen, gehöret auch insonderheit die Formirung der Stimme und der Sprache: wozu verschiedene Theile des Leibes dienen. Die Materie der Stimme und der Sprache ist die Lufft, welche aus der Lungen herausgestossen wird, und also dienen darzu die Lungen und die Lufftröhre. Insonderheit aber ist der Kopff (Larynx) hauptsächlich umb der Stimme willen vorhanden. Von den beyden gießkannenförmigen Knorpeln [...] wird der Ritz (glottis, rima) formiret, damit durch den engen Ausgang die Lufft geschwinde herausfähret, weil sonst keine Stimme und Sprache stattfinden könte. Und weil sich die Stimme ändert, nachdem der Ritz weit oder enge ist; so sind auch besondere Mäuslein vorhanden, welche ihn weiter und enger machen, nachdem es die Nothdurfft erfordert.<sup>23</sup>

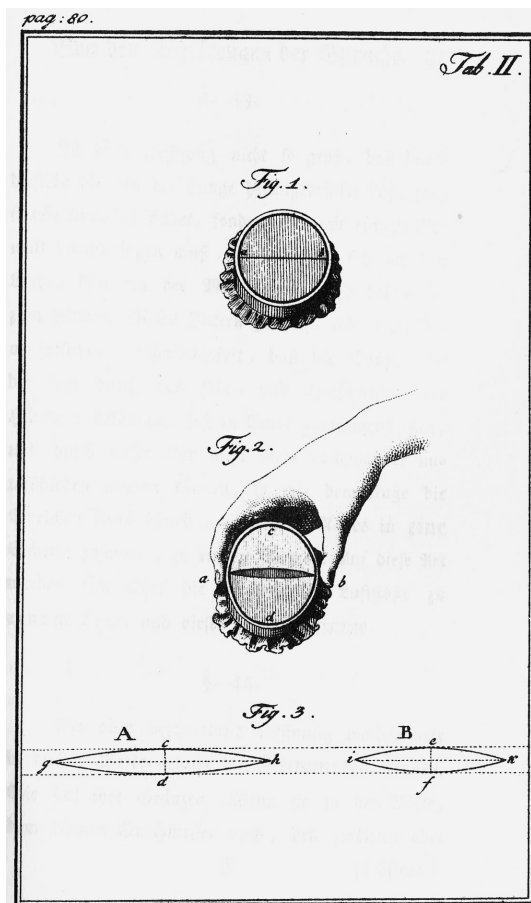


Abb. 2

Kempelens geteilte Membrane  
als Glottis



Die wesentliche Bedeutung der Glottis ist hier bereits Lehrmeinung, die einen wiederum zweifachen Zugang zur menschlichen Stimmerzeugung generierte. Einerseits untersuchten Physiologen den menschlichen Körper, um zu einer genaueren Kenntnis der menschlichen Sprechwerkzeuge zu gelangen, was dann nahe legte, diese Erkenntnisse auch an einem Modell überprüfen zu wollen.<sup>24</sup>

Mittels der neuen Erkenntnisse der Physiologie schien aber plötzlich auch ein altes Ziel vieler Konstrukteure und Erfinder erreichbar, das in einer künstlichen Erzeugung der stimmlichen Laute des Menschen bestand. Allerdings waren bislang nur Apparaturen entstanden, die mechanische Stimmen lediglich vortäuschten. So konstruierte ein gewisser Doktor Müller in Nürnberg zwei sprechende Apparate in menschlicher Lebensgröße und führte sie 1788 in Erlangen vor, die allerdings rasch in den Verdacht einer Täuschung gerieten.<sup>25</sup>

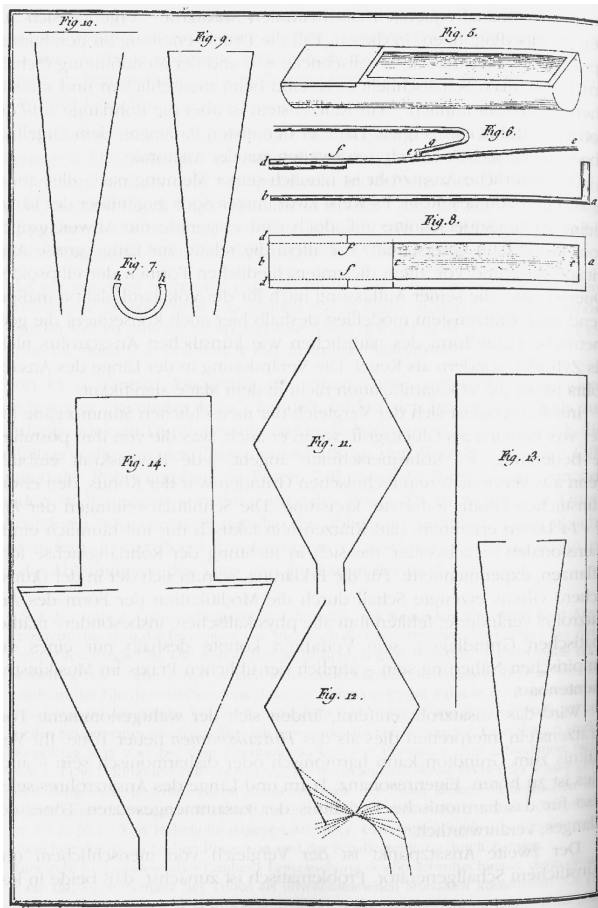
Neben solchen Hasardeuren traten um 1780 mindestens vier Apparatebauer nahezu gleichzeitig auf, die teils in jahrzehntelanger Arbeit Sprechmaschinen konstruierten, dabei jedoch mit unterschiedlichen Methoden vorgehen und zum Teil während ihrer Arbeiten die Konstruktionen der anderen gar nicht kannten. Eine erste Sprechmaschine wurde vorgestellt, nachdem der Mathematiker Leonhard Euler 1761 in der Nachfolge von La Mettrie die Konstruktion einer Maschine, »die alle Töne unsrer Wörter mit allen Artikulationen aussprechen könnte« ebenfalls für »nicht unmöglich«<sup>26</sup> erklärt hatte. Daraufhin setzte die Kaiserliche Akademie von St. Petersburg, der Euler angehörte, einen Preis für die Erklärung der physiologischen Unterschiede beim Sprechen der fünf Vokale und für den Bau eines Apparats aus, der sie erzeugen konnte. Euler ging bei seinen eigenen Forschungen zum Thema sprechender Maschinen von den akustischen Schallvorgängen der Artikulation und nicht ihren physiologischen Voraussetzungen aus.<sup>27</sup> Der Mediziner und Physiker Christian Gottlieb Kratzenstein, ebenfalls Mitglied der Akademie, gewann den 1779 ausgeschriebenen Preis, als er eine an der Physiologie der menschlichen Stimmbänder orientierte Konstruktion vorstellte, die mit Druckluft und Resonanzröhren betrieben wurde.<sup>28</sup> Dieser erste Apparat zur Erzeugung der mechanischen Stimme war noch gemäß den Forderungen der Ausschreibung nach dem Prinzip einer Orgelpfeife gebaut, der so genannten »Vox humana« Zungenpfeife.<sup>29</sup> Auf die für jeden Vokal einzeln mit Druckluft betriebenen Orgelpfeifen setzte Kratzenstein entsprechend gestaltete Resonanzröhren, die den Kehlkopf und das Ansatzrohr bei der Vokalartikulation vereinfacht nachahmen sollten.<sup>30</sup> Er reichte neben seiner Abhandlung auch eine kleine sprechende Orgel bei der Akademie ein, über deren weiteren Verbleib nichts bekannt geworden ist. In seinem Traktat revidierte Kratzenstein die Vorstellung von den Stimmbändern als Geigensaiten, statt dessen schlug er eher einen aerodynamischen Prozeß der

Daniel Gethmann

218

## Abb. 3

Kratzensteins Modelle zur künstlichen Vokalerzeugung



Stimmerzeugung vor.<sup>31</sup> Er stützte sich bei seinen Forschungen zum Bau von Sprechmaschinen neben Dodart, Ferrein und Euler auch auf den Naturforscher Albrecht von Haller und sein im Jahre 1766 auf Deutsch erschienenes Werk über die Stimme.<sup>32</sup> Für Haller trat die Resonanz in den Vordergrund der physiologischen Arbeiten, denn der ganze Körper werde zum Resonanzkörper, die Stimme entstehe durch »Zitterungen«. Damit setzte er sich deutlich von Dodart und Ferrein ab, die mit einer Stimmtheorie basierend auf Funktionen von Körperteilen analog zu musikalischen Instrumenten gearbeitet hatten. Auf Haller geht ebenfalls die Auffassung vom Ansatzrohr als Resonator zurück, denn man darf »wohl annehmen, dass in der Spalte der Zunge am Gaumen, im Widerschalle der Nasenhöhlen, und in der Verschiedenheit des Gaumens, die Ursachen liegen, welche den aus der Spalte heraufgetriebenen Klang auf allerlei Weise temperieren.«<sup>33</sup> Christian Gottlieb Kratzenstein legte in den folgenden Jahren keine weitere

Sprechmaschine vor, die über die Vokalartikulation hinausgekommen wäre, sondern beschäftigte sich als »Experimentalphysiker« mit Vorschlägen zur Luftschiff-Fahrt, in denen zum ersten Mal das Prinzip der Luftschraube ausgesprochen wurde und hatte eine medizinische Professur an der Universität Kopenhagen inne.

Einen weiteren Apparat zur Erzeugung von künstlichen Stimmen führte Abbé Mical im Jahre 1778 und erneut im Jahre 1783 vor, der unter dem Namen: *Têtes Parlantes* bekannt wurde. Abbé Micals *Sprechende Köpfe* basierten ebenfalls auf dem Prinzip einer Orgel, deren von einem Blasebalg gelieferte Luft in diverse Kanäle geleitet wurde, wo Membranöffnungen mit im Luftstrom schwingenden Zungen angebracht waren. Dies war das Bauprinzip der Orgelpfeife »Vox humana«, »doch erst durch die verschieden geformten Schallräume (boîtes sonores), in denen sich Membran und Zunge befanden, wurde der erzeugte Ton moduliert.«<sup>34</sup> Eine Kommission der Académie des Sciences unter Leitung des französischen Physiologen und Stimmforschers Félix Vicq d'Azyr legte am 3. September 1783 ein Gutachten zur Maschine vor, in dem die Kommission die Maschine als »interessant« bezeichnete und empfahl, Micals Arbeit weiterhin nach Kräften zu unterstützen.<sup>35</sup> Abbé Micals *Sprechende Köpfe* sprachen sogar in ganzen Sätzen. Die in der Abbildung überlieferten Aussagen der Maschine legen den Schluss nahe, dass der Apparat speziell für die Erschaffung einer ersten Kollektivstimme<sup>36</sup> und deren Ausführung konkreter Sprechakte konzipiert war, denn die *Sprechenden Köpfe* übernahmen wortreich die Aufgabe der Huldigung des Königs.<sup>37</sup> Eine frankophone Sprechfertigkeit ließ sich nunmehr mittels der Maschine für alle Zeiten festschreiben, die damit der Schrift in der Aufgabe der Standardisierung von Sprache und ihrem Klang mindestens ebenbürtig war. Dieser Vorläuferapparat der Tonaufzeichnungen inspirierte zu national- und universalsprachlichen Visionen, denn auf Tastendruck eines Symbols erklang der Sprachlaut, beide schießen zum ersten Mal seit Gutenberg wieder vereint:

Mical befestigt zwei Tastenbretter an seinen ›Sprechenden Köpfen‹, wovon eines nur eine beschränkte Anzahl von Sätzen ermöglicht, jedoch mit korrekten Wortintervallen und richtiger Betonung. Das andere hat weit mehr Tasten und enthält alle Klänge und Laute der französischen Sprache, die der Erfinder dank seiner ausgeklügelten Methode auf ein Minimum zu beschränken wusste. Mit etwas Übung und Geschicklichkeit wird man bald mit den Fingern so gut wie mit der Zunge reden können. Die Köpfe werden fließend und ruhig reden, ihrer Sprache wird man die ganze Eigenart verleihen können, die unbewegt ist von jeder Leidenschaft.<sup>38</sup>

Daniel Gethmann

220

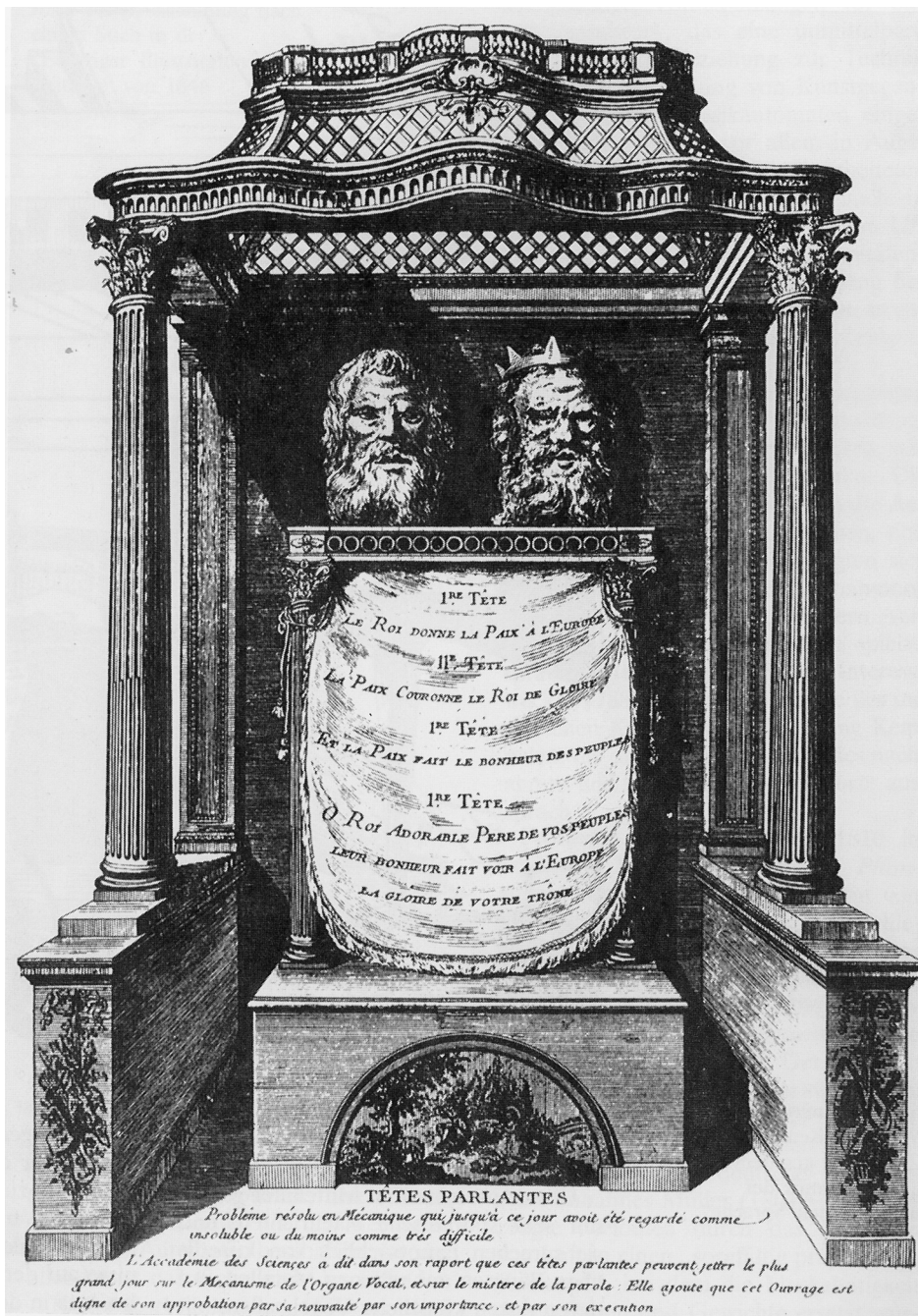


Abb. 4

Abbé Micals Sprechende Köpfe

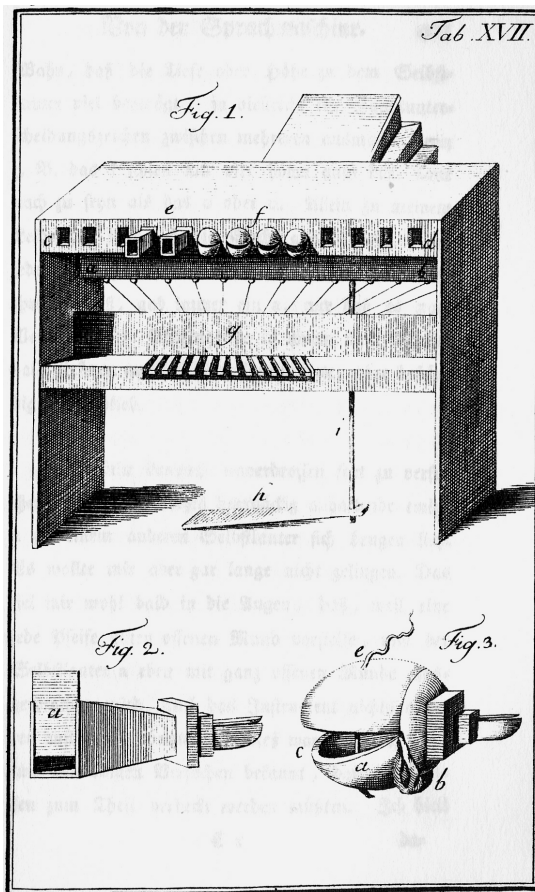
Abbé Mical zerstörte bald nach deren Erfindung seine *Sprechenden Köpfe* wieder, so dass über ihre weitere Funktionsweise nicht mehr bekannt wurde, denn er hatte auch der Kommission untersagt, darüber zu sprechen. Seinem Misstrauen ist es geschuldet, dass sich in der Kombination von Membranen und Zungenpfeifen eine Übernahme der Theorien von Dodart nur andeutet, aber nicht belegen lässt. Diese Membranen und Zungenpfeifen wurden in Micals Konstruktion jedoch um den wesentlichen Bereich der Kanäle oder allgemeiner der unterschiedlich geformten Schallräume erweitert, in denen man eine Würdigung der stimmlichen Beteiligung des Ansatzrohres erblicken kann. Die ersten Versuche von Kratzenstein und Mical kennzeichnen sich durch ihre Definition der Vokale und Konsonanten als Produkte der menschlichen Sprechorgane, die es daher nachzubauen galt und ihre Konstruktionen künstlicher Kehldeckel und Stimmlippen.

Der Techniker und österreichische Hofmaschinist Wolfgang von Kempelen, der bereits durch eine berühmte Schachspielmaschine mit einem versteckten kleinwüchsigen Schachspieler aufgefallen war, trat 1791 in Kenntnis der Abhandlung Kratzensteins<sup>39</sup> mit einer ausgefeilten Maschine an die Öffentlichkeit, die durch die erweiterte Erzeugung von Konsonanten beeindruckte und sogar französische oder lateinische Sätze aufsagte.<sup>40</sup> Es gab jedoch noch weitere Unterschiede zwischen von Kempelen und den beiden anderen Konstrukteuren Kratzenstein und Mical, die nicht nur auf technischem Gebiet lagen. Nährte sich Kratzenstein dem Thema mit seiner physiologischen Ausbildung und damit auch unter der weiterführenden Fragestellung eines Zusammenhangs von Seele und Stimme oder ihrer mechanischen Erzeugung, so erweckte Abbé Mical aus seinen technischen Ursprüngen in der Tradition der Automatenbauer den Anschein einer verbindlichen Standardisierung von Hochsprache. Wolfgang von Kempelen hingegen stammte eher aus der Tradition des mechanischen Apparatebaus, der Konstruktion von Automaten und befand sich als hoher Beamter in engem Kontakt zur österreichischen Aristokratie. Bei seinen 22 Jahre währenden Arbeiten an seiner *Sprechenden Maschine* stützte sich von Kempelen nach eigenen Angaben insbesondere auf die Stimmtheorie von Dodart.<sup>41</sup> Er stellte schließlich nach zweimaligem Abbruch der Arbeit an anderen Prototypen eine nur unter Einsatz hoher körperlicher Geschicklichkeit zu betreibende Sprechmaschine vor, die zur Stimmerzeugung weiterhin den ganzen Körper benötigte. Über einen Blasebalg X musste Luft mit dem rechten Arm gepumpt werden, während die rechte Hand die Konsonanten-Vorrichtungen R, SCH, N, M und S bediente. Die linke Hand hingegen operierte vor der Öffnung C, um durch geschicktes Abdecken und andere Gesten die Vokale hervorzubringen. Die dazu notwendigen Gesten gab von Kempelen in seiner Abhandlung, einer »der ersten detaillierten Gebrauchsanwei-

Daniel Gethmann

222

Abb. 5

Kempelens Sprechmaschine  
Nr. 2 und Zungenpfeifen

sungen zu einer sprechenden Maschine<sup>42</sup> an und sah voraus, dass die fertige Maschine durchaus wie einer ihrer Vorgänger, die von Kempelen nicht zufrieden stellten, an eine Tastatur angeschlossen werden könne. Wenn man sich in Sprachen wie Französisch oder Italienisch mit starkem Vokalanteil versuche, sei man in der Lage, in drei Wochen erstaunliche Fortschritte in der Beherrschung der Maschine zu machen und »eine bewundernswerthe Fertigkeit im Spielen erlangen«.<sup>43</sup>

Bedeutsam für von Kempelens Sprechmaschine ist, dass sie trotz obiger Beschreibung keine Buchstaben artikuliert, sondern »Sprachtöne«,<sup>44</sup> dass mit ihr mithin die Geschichte der Phonetik beginnt.<sup>45</sup> Nicht die möglichst genaue Artikulation einer heiligen Schriftsprache hatte von Kempelen nämlich mit seiner Maschine im Sinn, sondern er erfasste das lautliche Vermögen der Sprache für den einzelnen Sprechenden. Seine *Sprechende Maschine* generierte daher auch ein neues eigenes »Hauptalphabeth«,<sup>46</sup> das SCH und CH verzeichnete, aber C, Q, X

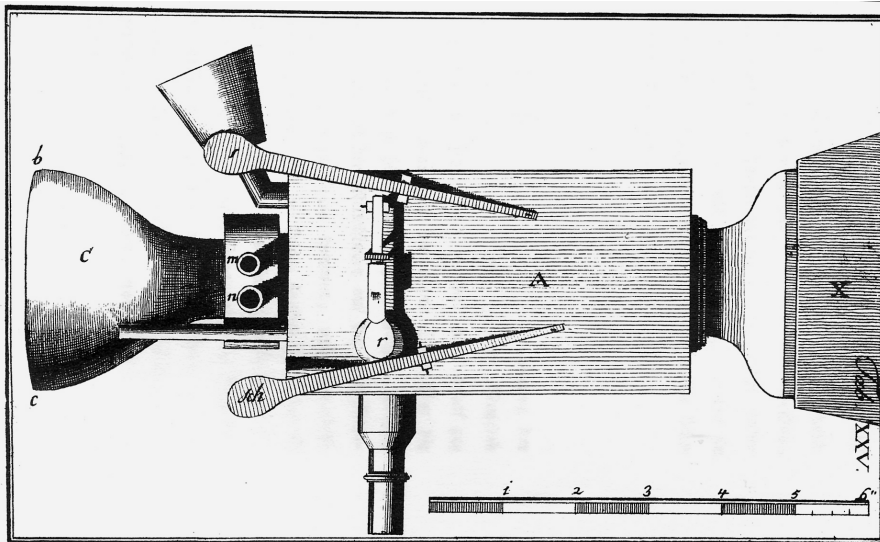


Abb. 6

Kempelens *Sprechende Maschine*

und Y aus klanglichen Gründen ausließ. »Buchstabe und Laut waren weit auseinandergetreten; dazwischen lavierte die Stimmphysiologie.«<sup>47</sup> Von dieser hatte sich Wolfgang von Kempelen jedoch dadurch gelöst, dass er Kratzensteins Bauweise einer Maschine, die noch am physiologischen Aufbau der menschlichen Stimmorgane orientiert war, durch eine Konstruktion überwand, die sich für die Funktion des Stimmorgans und insbesondere der Glottis interessierte. Die Veränderung bestand eben darin, nicht länger die menschliche Anatomie der an der Stimmerzeugung beteiligten Körperteile nachzuahmen, sondern bei der Konstruktion einer Sprechmaschine an erster Stelle deren Funktionsweise zu beachten. Dieser für die entstehende akustische Wissenschaft wegweisende Ansatz, eine Maschine der Stimme speziell unter Gesichtspunkten der Schallerzeugung zu konstruieren, wird allerdings durch die Tatsache verdeckt, dass auch weiterhin der ganze Körper zur Stimmerzeugung benötigt wurde.

Bei der Entstehung der synthetischen Stimme zeigt sich eine überraschend hohe Akzeptanz der Trennung des ursprünglichen Zusammenhangs von Stimme und Körper, dies hängt sicherlich auch mit der für ihre mechanische Erzeugung benötigten körperlichen Geschicklichkeit zusammen. Denn die Übernahme menschlicher Sprechfähigkeit durch eine Maschine wurde durchaus als weiterhin einer Person zugehörig verstanden. Am Ende seiner Gebrauchsanweisung der *Sprechenden Maschine* und der Beschreibung seiner Fähigkeiten, sie zu bedienen,

schrieb von Kempelen nämlich den auf die Beherrschung der Maschine bezogenen Satz: »Ich spreche ein jedes französisches oder italienisches Wort, das man mir vorsagt, auf der Stelle nach.«<sup>48</sup> Die Stimme des Ich wird insofern zu einer Funktion der Apparatur und das erste, was sie sagt, ist natürlich: »Papa« und »Mama.«<sup>49</sup> Hier gilt für die Maschine, was Schiller über die Seele sagte: Spricht die Maschine, so spricht, ach, die Maschine nicht mehr. Denn das Sprechen war auch weiterhin wie selbstverständlich an den Menschen gebunden, auch wenn es sich von den Gegebenheiten seiner Physiologie löste, blieb es dennoch vorerst der Kern seiner Souveränität. Diese Vereinnahmung der Maschine führte jedoch zu Konsequenzen: Als eine Nachbildung des menschlichen Stimmorgans zur Erzeugung von Sprachlauten nicht länger notwendig war, begann die Zeit, in der der Maschinendiskurs auf die Konzepte und Vorstellungen vom Körper wirkte. Nach 1791 weiß man also nichts über seine Sinne, bevor nicht Medien Modelle und Metaphern bereitstellen. Vor 1791 verdeutlicht die Geschichte der Sprechmaschine, dass das Wissen über die Medien vom Körper und seinen Studien in der Physiologie abgeleitet ist. Das Missverständnis der klassischen anthropozentrischen Techniktheorien, das darin besteht, Medien als Ausweitungen des menschlichen Körpers zu verkennen, statt ihre Entwicklungsgeschichte in Verbindung mit physiologischer Forschung genauer zu analysieren, äußert sich demzufolge bereits in dem Satz, mit dem von Kempelen die Stimme einer Sprechmaschine für seine Person reklamiert.

Am Beispiel der Stimme und ihrer Simulationsapparate, die in einer ihrer technischen Entwicklungslinien auch weiterhin von der Gestalt und Entnahme der entsprechenden Organe aus dem menschlichen Körper abhängig blieben, möchte ich abschließend zeigen, wie sich das Verhältnis von Sinnen und Medien nach dieser medienhistorisch bedeutsamen Umkehrung entwickelte. Dabei ist die Tatsache bemerkenswert, dass zur gleichen Zeit dieser Dissoziation von Stimme und Körper auch die Akustik als eigenes Fachgebiet entstanden ist. In Ernst Florens Friedrich Chladnis grundlegendem Werk »Die Akustik«, dessen Erscheinungsjahr 1802 häufig als Geburtsjahr dieser Wissenschaft genannt wird, finden sich Verweise auf von Kempelen und Kratzenstein sowie eine Beschreibung des Kehlkopfes, dessen Studium demnach in der damaligen Zeit Grundlagenerkenntnisse der Akustik vermittelt hat. In der Geschichte der Akustik wird ebenfalls häufig auf den Physiker Robert Willis verwiesen, der das Studium der Stimme in die Akustik integriert habe und sich ausdrücklich auf von Kempelen und Kratzenstein berief.<sup>50</sup> Willis ging ganz im Sinne der modernen Akustik so vor, dass er beschloss, »die Sprachwerkzeuge gänzlich zu vernachlässigen, und durch Versuche mit den gewöhnlichen akustischen Instrumenten wo möglich zu



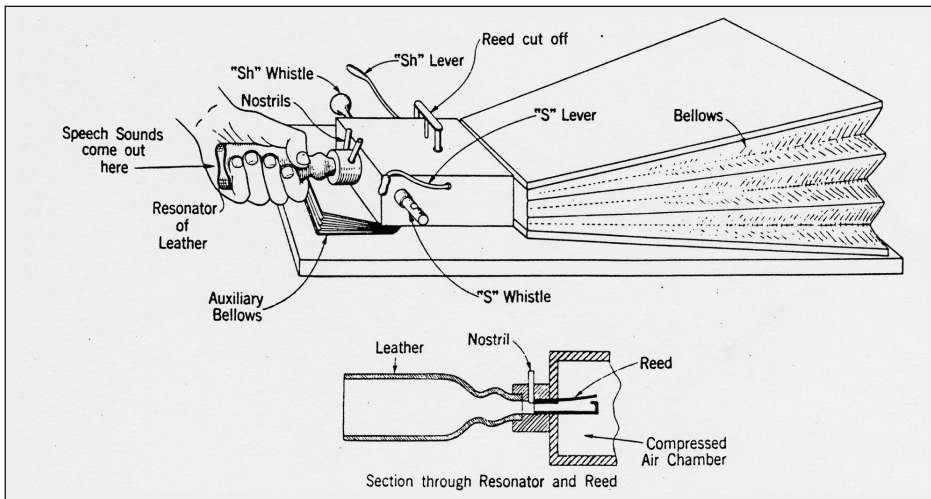


Abb. 7

Wheatstone's *Sprechmaschine*

bestimmen, welche Gestalten von Hohlkörpern, und welche andere Bedingungen zur Hervorbringung dieser Töne wesentlich seyen. «<sup>51</sup> Bei diesen empirischen Untersuchungen von Zungenpfeifen bestätigte sich die vormals von Albrecht von Haller physiologisch begründete These, dass ein weiterer wesentlicher Faktor der Stimmerzeugung in der Länge des Ansatzrohres zu suchen sei. Auch der englische Physiker Charles Wheatstone knüpfte an von Kempelen an, da er sich ebenfalls mit Sprechmaschinen befasste und von Kempelens *Sprechende Maschine* sogar nachbaute. Dieser Nachbau wurde auf einer Sitzung der »British Association for the Advancement of Sciences« im August 1835 vorgestellt, war bereits einfacher als das Original zu bedienen und ist klar als von Kempelens Erfindung erkennbar.<sup>52</sup>

In einem Essay äußerte Charles Wheatstone einige Jahre zuvor ein weiteres Motiv der Forschung an der Stimme, das eine kommende Medientechnik bereits projektierte. Seiner Meinung nach war nämlich keine andere Entwicklung auf dem Gebiet der Akustik so bedeutsam wie die einer Übertragung der Stimme. Dabei schlug er eine Kombination von Übertragungs- und Reproduktionstechnik vor, indem nicht die Stimme sondern Steuerungssignale übertragen werden sollten, mittels derer an einem anderen Ort befindliche Sprechmaschinen die Stimme und damit auch die Botschaft reproduzierten.

I have found by experiment that all these articulations, as well as the musical inflexions of the voice, may be perfectly, though feebly, transmitted

to any of the previous described reciprocating instruments [...] [C]ould articulations similar to those enounced by human organs of speech be produced immediately in solid bodies, their transmission might be effected with any required degree of intensity. Some recent investigations lead us to hope that we are not far from effecting these desiderata; and if all the articulations were once thus obtained, the construction of a machine for the arrangement of them into syllables, words, and sentences, would demand no knowledge beyond that we already possess.<sup>53</sup>

Wie zum Beweis fertigte Wheatstone dann seinen Nachbau von v. Kempelens *Sprechender Maschine* an. Sein Vorschlag sollte sich wiederum auf die Erforschung der menschlichen Stimme auswirken. Denn parallel zur akustischen Forschung und ihrer Durchsetzung einer veränderten Stimmlokalisation sorgten auch weiterhin Physiologen, die in der Nachfolge von Ferrein und Albrecht von Haller an Leichenteilen experimentierten, für Fortschritte auf dem Gebiet der künstlichen Stimmerzeugung. Diese Entwicklungslinie von Apparaturen der Stimme stand in enger Verbindung mit der anatomischen Forschung. So bemühte sich der Arzt Karl Friedrich Liscow in der Nachfolge von Ferrein am präparierten Kehlkopf unterschiedliche Stimmlagen zu erzeugen, damit sich seine Theorie der Stimme »durch den untrüglichen Ausspruch der Natur«<sup>54</sup> bekräftige. Der berühmte Physiologe Johannes Müller ging knapp einhundert Jahre nach Antoine Ferrein den Schritt von präparierten Kehlköpfen zu größeren Leichenteilen, um die Erkenntnisse von Ferrein und die Theorien von Willis am menschlichen Stimmobjekt zu überprüfen und gleichzeitig eine Sprechmaschine zu bauen.<sup>55</sup>

Johannes Müller schnitt den Kopf einer Leiche so ab, dass der ganze Stimmapparat mit einem Teile der Luftsäule daran hing. [...] Sodann hängte er den Kopf mit dem Stimmapparat auf, indem er die hintere Wand des Stimmapparates wie gewöhnlich an einem Pfeiler befestigte [...]. Wenn in dieser Anordnung der Kehlkopf der Leiche angeblasen wurde, so wurde der Klang der Stimme der menschlichen so ähnlich, dass aller Unterschied des lebenden Körpers und der Maschine verschwand. Bei passiver Bewegung der Lippen konnte Johannes Müller sogar zur Bildung einiger Konsonanten schreiten. Das M und das W gelangen sehr leicht, auch die Vokale U und A wurden durch die nötigen Veränderungen der Mundöffnung erzeugt, so dass die Leiche also noch klare menschliche Sprachlaute produzieren konnte. Das, was bei diesem Versuch besonders auffiel, war die absolute Menschenähnlichkeit der so erzeugten Leichen-

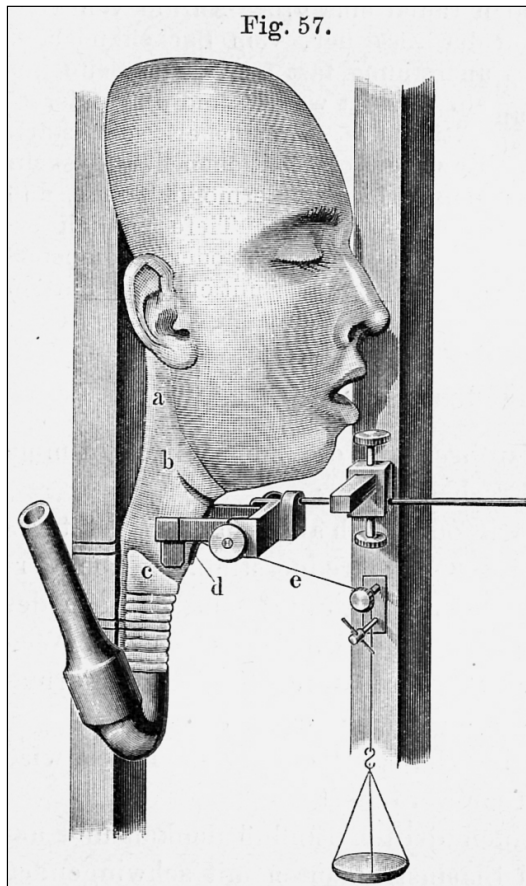


Abb. 8

Müllers Präparierter Kopf

stimme, was man von einem einfachen isolierten Leichenkehlkopfe kaum sagen kann. Man kann demnach wohl ohne Übertreibung die Behauptung aufstellen, dass das eigentlich Menschenähnliche in der Stimme durch die Einwirkung des Ansatzrohres zustande kommt.<sup>56</sup>

Damit hatte Johannes Müller die Ergebnisse der Akustik von Willis und Wheatstone durch physiologische Experimente bestätigt, die Zeit um 1830 brachte viele weitere Theorien und Experimente zur Stimmproduktion hervor, die eine Eigenresonanz des Ansatzrohres untersuchten, oder sich auf sein Zusammenwirken mit Atem und Stimmlippen zur Stimmproduktion verständigten.

Im Jahre 1835 trat der gebürtige Freiburger Joseph Faber mit seiner Sprechmaschine *Euphonia* auf, die die bis dahin beste Sprechleistung erbracht haben soll.<sup>57</sup> Sie unterschied sich von ihren Vorgängern darin, dass unter Einbeziehung

neuerer Forschungen der Physiologie großer Wert auf eine exakte Nachbildung der menschlichen Sprechorgane gelegt wurde und weniger ein Musikinstrument Pate stand. So ließ sich mittels sechzehn Tasten der *Euphonia* eine unerreichte Wiedergabetreue der menschlichen Sprache erreichen. Mit dieser Tastatur wurden die unterschiedliche Position des Unterkiefers, die Öffnung der Lippen und die Bewegungen der Zunge nachgeahmt, also das Ansatzrohr funktionell nachgebaut, was auch Wolfgang von Kempelen in der vierten, nicht-realisierten Version seiner *Sprechenden Maschine* vorgeschlagen hatte. Über diese Tastatur konnte ebenfalls die Stellung der Membran, die innerhalb eines Gummizylinders mit einer kleinen Pfeife und aufschlagender Elfenbeinzunge für die Funktion des Kehlkopfes und insbesondere der Stimmbänder zuständig war, gesteuert werden; über ein Pedal wurde der Blasebalg, der die Funktion der Lunge übernahm, betrieben. Insofern war *Euphonia* nicht einfach zu bedienen, so dass dem Techniker unbekannte Sätze undeutlicher gelangen als die häufig aufgesagten. In der verkleideten anthropomorphisierenden Variante der Sprechmaschine hing eine Puppe am Ende des Ansatzrohres, die ihren Mund beim Sprechen bewegte und die Augenlider senkte. Faber selbst schrieb nicht über seine Erfindung, sie wurde jedoch kurz und voll des Lobes in den Annalen der Physik erwähnt. Der Autor berichtete, dass man ihre Stimmhöhe und Lautstärke verändern könne.<sup>58</sup> *Euphonia* hat ihrem Erfinder ebenso wenig wie die anderen kommerziell verwerteten Sprechmaschinen der damaligen Zeit finanziellen Erfolg gebracht, sie wurde zwar vom Ehemann der Nichte Fabers im Gefolge von P. T. Barnum noch 1873 in den Vereinigten Staaten vorgeführt,<sup>59</sup> doch verliert sich dann ihre Spur und ihr weiterer Verbleib ist unbekannt.<sup>60</sup> Eine Imitationsmaschine der menschlichen Stimmorgane konnte Mitte des 19. Jahrhunderts nicht mehr auf wissenschaftliche Anerkennung hoffen, da die Akustik längst Apparate nach rein akustischen und nicht länger physiologischen Charakteristika baute.<sup>61</sup> Hermann von Helmholtz' elektrischer Vokalapparat ist hierfür ein gutes Beispiel, mit dem er seine Theorie des Ohrs bei der Wahrnehmung von bestimmten musikalischen Klangfarben oder Vokalen der gesprochenen Sprache und des Gesangs in Abhängigkeit von Obertönen stützte.<sup>62</sup>

Auch Alexander Graham Bell, der im Jahre 1863 eine Vorstellung *Euphonia*s in London sah, konnte danach bei einem Besuch bei Charles Wheatstone dessen Nachbau der *Sprechenden Maschine* von v. Kempelen besichtigen und versuchte sich an der Konstruktion eines Sprechapparates. Bell gelangte mit der erklärten Absicht, eher die Natur als von Kempelen zu kopieren, bis zur Aussprache einzelner Vokallaute auf dem Weg der mechanischen Reproduktion. Für seine Konstruktion nutzte er die genaueste Fassung des Ansatzrohres, die sich finden ließ, nämlich einen Totenschädel mit einer Mundhöhle aus Gummi und einem ebensolchen Keh-



Abb. 9

Fabers *Euphonia*

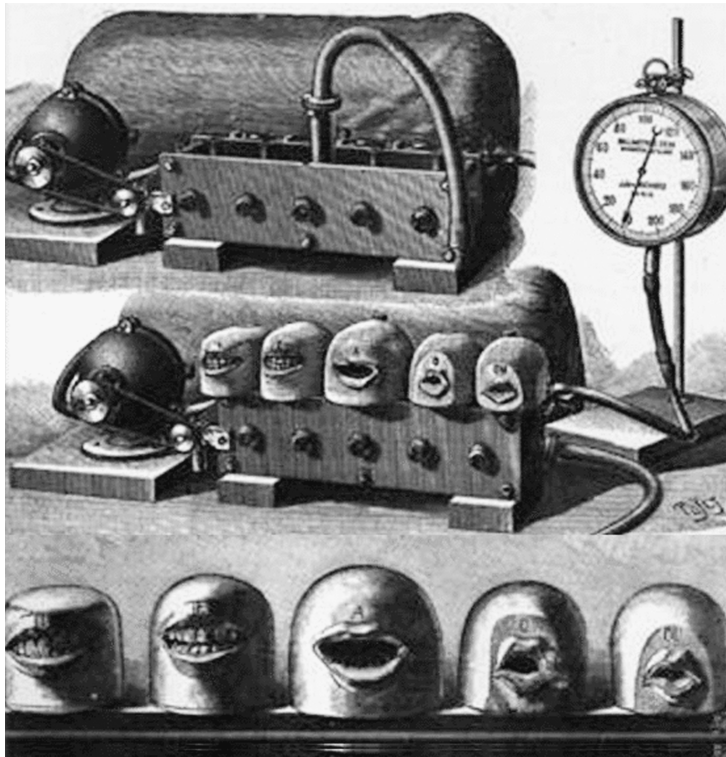
kopf.<sup>63</sup> Bei dessen Konzeptualisierung hatte auch eine Katze aus der Nachbarschaft ihr Leben verloren, deren Kehlkopf ebenfalls zu Studienzwecken benötigt wurde. Mit dieser gescheiterten Konstruktion war übrigens auch Wheatstones Idee einer Kommunikationsübertragung über den für Bell naheliegenden Anschluss von Sprechmaschinen an Telegraphenleitungen erledigt, die physiologische Forschung beeinflusste bekanntlich eher auf der Hörerseite des Telephons eine eigene Konstruktion wie den Bell'schen *Phonautographen* aus dem Jahre 1874, der in wesentlichen Bestandteilen aus seziierten menschlichen Hörorganen bestand.

Zuletzt trat um die Jahrhundertwende der französische Physiologe und Professor für Stimme an der Sorbonne, Georges René Marie Marage, mit dem Vorhaben an die Öffentlichkeit, eine möglichst naturgetreue Nachbildung der menschlichen Stimmorgane zur künstlichen Stimmerzeugung zu nutzen und nachzubauen. Dabei konnte er sich auf viele Geräte zur Sichtbarmachung des Schalls stützen, insbesondere auf Edouard Leon Scotts *Phonautographen* als auch auf Photographien manometrischer Flammen, wie sie 1862 vom akustischen Instrumentenbauer Rudolph König in Paris vorgestellt wurden. Marage fertigte Drehscheiben für Lochsirenen, die von den Charakteristika der optischen Schallaufzeichnung eines Lautes ausgehend konstruiert waren. Seine Ansatzrohre waren hingegen nach Gießformen gefertigte, detailgetreue Kopien der Mundhöhle mit Lippen und Zähnen.<sup>64</sup> Diese Abdrücke der Vokalartikulation gingen um 1900 in

Daniel Gethmann

230

Abb. 10

Marage's  
Sprechmaschine

ihrer anthropomorphen Gestaltung den Schritt zur Maschine, die nicht mehr nur ihre Bauprinzipien von der Erforschung der menschlichen Physiologie bezog, sondern im Unterricht von Gesangsschülern oder in der Behandlung von Sprechstörungen mit ihrer synthetischen Stimme die Funktion eines Vorbilds für die menschliche Stimmbildung übernahm.

- 1 Jean Paul: Nachflor und Spätlinge des Taschenbuchs, in: Jean Paul's Werke, Bd. 48: Vermischte Aufsätze und Dichtungen aus Zeitschriften, Taschenbüchern und Jean Paul's handschriftlichem Nachlass, Berlin 1901, S. 193. Jean Paul zählte Wolfgang von Kempelen zu seinen »Todfeinden« und widmete ihm im Jahre 1789: Unterthänigste Vorstellung unser, der sämtlichen Spieler und redenden Damen in Europa, entgegen und wider die Einführung der Kempelischen Spiel- und Sprachmaschinen, in: ebd., Bd. 4: Auswahl aus des Teufels Papieren, Berlin 1901, S. 51–65.
- 2 Vgl. zur Vorgeschichte: Heinrich Maximilian Brunner: Ausführliche Beschreibung der Sprachmaschinen oder sprechenden Figuren, Nürnberg 1798; David Brewster: Letters on Natural Magic. Addressed to Sir Walter Scott, London 1832; Walter Niemann: Sprechende Figuren. Ein Beitrag zur Vorgeschichte des Phonographen, in: Geschichtsblätter für Technik und Industrie, Bd. VII (1922), S. 2–30; Herbert Heckmann: Die andere Schöpfung. Geschichte der frühen Automaten in Wirklichkeit und Dichtung, Frankfurt/M. 1982; Reinhold Hammerstein: Macht und Klang. Tönende Automaten als Realität und Fiktion in der alten und mittelalterlichen Welt, Bern 1986.
- 3 Bis ins 20. Jahrhundert hinein basieren Erkenntnisse der Phonetik auf physiologischen Untersu-

- chungen. Vgl. beispielhaft die Anwendung der Röntgenphotographie bei: Oscar George Russell: *The Vowel. Some X-Ray and Photo Laryngoperiskopik Evidence*. Columbus, Ohio 1928; ders.: *The Mechanism of Speech*, in: *Journal of the Acoustical Society of America* 1 (1929), S. 83–109.
- 4 François Rabelais: *Gargantua und Pantagruel* [1532], hg. v. Horst und Edith Heintze, 2 Bde., Bd. 2, (Viertes Buch), Frankfurt/M. 1974, S. 167.
- 5 Ebd., S. 169.
- 6 François Rabelais: *Le Quart Livre* [1532], in: ders.: *Œuvres Complètes*, hg. v. Guy Demerson, Paris 1973, S. 1074.
- 7 Francis Bacon: *Nova Atlantis, Fragmentorum alterum*, London 1637; deutsche Übersetzung: Ders.: *Neu-Atlantis*, in: ders.: *Der utopische Staat*, hg. v. Klaus J. Heinisch, Reinbek 1960, S. 211.
- 8 Zur Einführung: Thomas L. Hankins/Robert J. Silverman: *Instruments and Imagination*, Princeton 1995; Joachim Gessinger: *Auge und Ohr. Studien zur Erforschung der Sprache am Menschen 1700–1850*, Berlin 1994; Jens-Peter Köster: *Historische Entwicklung von Syntheseapparaten zur Erzeugung statischer und vokalartiger Signale nebst Untersuchungen zur Synthese deutscher Vokale*, Hamburg 1973.
- 9 Vgl. zur Geschichte der Automaten: Alfred Chapuis/Édouard Gélis: *Le monde des automates*, 2 Bde., Paris 1928; ders./Edmond Droz: *Automata. A Historical and Technical Study*, Nechatel 1958; Derek J. de Solla Price: *Automata and the Origins of Mechanism and Mechanistic Philosophy*, in: *Technology and Culture* 5 (1964), S. 9–24; Silvio A. Bedini: *The Role of Automata in the History of Technology*, in: *Technology and Culture* 5 (1964), S. 24–42; Klaus Völker: *Künstliche Menschen*, München 1971; Pia Müller-Tamm/Katharina Sykora: *Puppen, Körper, Automaten. Phantasmen der Moderne*, Köln 1999.
- 10 »Der Körper gleicht einer Uhr.« Julien Offray de La Mettrie: *L'Homme Machine* [1748], Paris 1921, S. 124.
- 11 »Der Mensch ist nur ein Tier bzw. eine Maschinerie von Triebfedern, die sich gegenseitig spannen.« Ebd., S. 123.
- 12 »[Vaucanson] müsste sich noch mehr anstrengen, um einen Sprecher herzustellen; eine solche Maschine darf nicht mehr als eine Unmöglichkeit angesehen werden.« Ebd., S. 130.
- 13 Vgl. Hans Baumgarten: *Galen. Über die Stimme. Testimonien der verlorenen Schrift Peri phones. Pseudo-Galen: De vove et hanelitu. Kommentar*, Diss., Göttingen 1962.
- 14 Aristoteles: *De Audilibus*, in: *Aristoteles Werke in deutscher Übersetzung*, hg. v. Hellmut Flashar, Bd. 18: *Opuscula II/III*, Darmstadt 1972, S. 157.
- 15 Vgl. Denis Dodart: *Mémoire sur les causes de la voix de l'homme, et de ses différens tons. (Über die Ursachen der menschlichen Stimme und die verschiedenen Laute)*, in: *Mémoires de l'Académie royale des sciences* [1700], S. 238–287; Ders.: *Supplément au mémoire sur la voix et sur les tons*. In: *Mémoires de l'Académie royale des sciences* [1706], S. 136–148 und 388–410 sowie ders.: *Supplément au mémoire sur la voix et sur les tons*, in: *Mémoires de l'Académie royale des sciences* [1707], S. 66–81.
- 16 Gessinger: *Auge und Ohr* (Anm. 8), S. 455.
- 17 »Galen, und nach ihm alle Alten, ja auch viele Neue waren der Meinung, daß die Luftröhre einen ganz besondern, und nothwendigen Antheil an der Stimme hat, bis Dodart klar zeigte, daß sie nur in so weit zur Stimme beyträgt, als sie der nothwendige Kanal ist, der die Luft aus der Lunge dem Stimmhäutchen zuführt, daß sie aber zum Zittern der Luft gar nichts wirkt, sondern hier nur das ist, was bey der Orgel die Windlade.« Wolfgang von Kempelen: *Mechanismus der menschlichen Sprache nebst der Beschreibung seiner sprechenden Maschine* [1791], Stuttgart/Bad Cannstadt 1970, S. 71.
- 18 Vgl. Jamie C. Kassler: *Man – A Musical Instrument. Models of the Brain and Mental Functioning Before the Computer*, in: *History of Science* 22 (1984), S. 59–92.
- 19 Vgl. Antoine Ferrein: *De la Formation de la Voix de l'Homme*, in: *Mémoires de l'Académie royale des sciences* (1741), S. 409–432.
- 20 Ferrein: *De la Formation de la Voix de l'Homme* (Anm. 19), S. 417.
- 21 Vgl. Denis Diderot: *Les bijoux indiscrets*, in: ders.: *Œuvres*, hg. v. André Billy, Paris 1951, S. 1–234. In der Geschichte tritt Antoine Ferrein unter dem Namen Docteur Orcotome vor der Académie de Banza auf, wiederholt seine Versuche und erklärt seine Theorie: »Die Luft, die von außen eindringt, übernimmt die Rolle eines Bogens auf den sehnigen Fasern der Flügel, die ich Stimmbänder nennen möchte. Der sanfte Zusammenprall der Luft mit diesen Stimmbändern bringt sie zum Schwingen; und durch die mehr oder weniger raschen Schwingungen werden verschiedenelei Töne he-

- rausgebracht.« Denis Diderot: Die indiskreten Kleinode, in: ders.: Das erzählerische Gesamtwerk von Denis Diderot, Bd. 2, hg. v. Hans Hinterhäuser, Berlin 1966, S. 42.
- 22 Ferrein: De la Formation de la Voix de l'Homme (Anm. 19), S. 430 f. Zit. n. und übers. v.: Gessinger: Auge und Ohr (Anm. 8), S. 462.
- 23 Christian Wolff: Vernünftige Gedancken von dem Gebrauche der Theile in Menschen, Thieren und Pflantzen, Frankfurt/M./Leipzig 1725, S. 495 f.
- 24 Der Naturforscher Christoph Friedrich Hellweg überprüfte seine im Jahre 1781 in Tübingen unter dem Titel: »De formatione loquela« veröffentlichten Theorien zur Stimme und zum Sprechen, indem er Johann Beck untersuchte, der in Europa vorgeführt wurde, da sein vom Wundbrand zerstörtes Gesicht und der offene Gaumen einen Blick auf den Funktionsmechanismus des Kehlkopfes ermöglichte. Vgl. Brigitte Felderer: Stimm-Maschinen. Zur Konstruktion und Sichtbarmachung menschlicher Sprache im 18. Jahrhundert, in: Friedrich Kittler (Hg.): Zwischen Rauschen und Offenbarung. Zur Kultur- und Mediengeschichte der Stimme, Berlin 2002, S. 257–278 (hier: S. 262 ff.).
- 25 Vgl. Anonym: Über Herrn Doktor Müllers redende Maschine und über redende Maschinen überhaupt, Nürnberg 1788. Als Verfasser wird der Mathematiker Adam Gottfried Wetzel genannt. Vgl. Johann August Donndorf: Geschichte der Erfindungen. Bd. 3, Quedlinburg/Leipzig 1817, S. 370 f.
- 26 Leonhard Euler: Briefe an eine deutsche Prinzessin über verschiedene Gegenstände aus der Physik und Philosophie [1769], Reprint, Braunschweig 1986, S. 158 f. (Brief Nr. 137 vom 16. Juni 1761). Die Briefe sind an Friederike Charlotte Ludovica Luise von Brandenburg-Schwedt gerichtet. Vgl. das Vorwort von Andreas Speiser in ebd., S. XXI.
- 27 Allerdings wies Euler in der Nachfolge von Albrecht von Haller auf die Bedeutung des Ansatzrohres für die Stimmerzeugung hin, als er schrieb, »daß man zur Aussprache dieser verschiednen Vocalen der Höhlung des Mundes eine verschiedne Gestalt geben muß.« Ebd., S. 158.
- 28 Vgl. Christian Gottlieb Kratzenstein: Tentamen resolvendi problema ab Academia Scientiarum Imperiali Petropolitana ad annum 1780 publice propositum: 1) Qualis sit natura et character sonorum litterarum vocalium a, e, i, o, u, tam insigniter se diversorum; 2) Annon construi queant instrumenta ordini tuborum organicorum, sub termino vocis humanae noto, similia, quae litterarum vocalium a, e, i, o, u, sonos expriment, Petersburg 1781; ders.: Sur la naissance & la formation des voyelles, in: Journal de Physique 21 (1782), S. 358–380.
- 29 Vgl. Herbert Jüttemann: Mechanische Musikinstrumente. Einführung in Technik und Geschichte (Das Musikinstrument Bd. 45), Frankfurt/M. 1987, S. 76 ff.
- 30 Vgl. Gessinger: Auge und Ohr (Anm. 8), S. 568–582; Köster: Historische Entwicklung von Syntheseparametern (Anm. 8), S. 68–74.
- 31 Im Jahre 1819 führte Charles Cagniard de la Tour den Begriff »Sirene« für eine Apparatur ein, die einen kontinuierlichen Luftstrom in periodischen Zeitintervallen unterbrach und so Schall erzeugte. (Lochsirene) Vgl. Charles Cagniard de la Tour: Sur la Sirène. Nouvelle Machine d'Acoustique destinée à mesurer les vibrations de l'Air qui constituent le son, in: Annales de Chimie et de Physique 12 (1819), S. 167 ff.
- 32 Vgl. Albrecht von Haller: Elementa Physiologiae Corporis Humani. Bd. 3: Respiratio, Vox, Lausanne 1761. Deutsche Übersetzung von Johann Samuel: ders.: Albrecht von Hallers Anfangsgründe der Physiologie des menschlichen Körpers. Bd. III: Das Atemholen, Die Stimme, Berlin 1766.
- 33 Haller: Anfangsgründe der Physiologie (Anm. 32), S. 705.
- 34 Gessinger: Auge und Ohr (Anm. 8), S. 539.
- 35 Handschriftliches Manuskript in den »Procès Verbaux de l'Académie des Sciences et Belles Lettres«. Zit. n. ebd., S. 540 f. Vgl. auch: Felix Vicq d'Azyr: Premier Mémoire sur la voix, in: Mémoire de l'Académie royale des Sciences de Paris. Année 1779, Paris 1782, S. 178–206.
- 36 Abbé Mical soll geplant haben, allen großen Bibliotheken der Welt seine Maschine zur Verfügung zu stellen, um den Spracherwerb und die korrekte Aussprache der französischen Hochsprache zu fördern.
- 37 1. Kopf: Der König schenke Europa Frieden. 2. Kopf: Der Frieden bekränze den König mit Ruhm. 1. Kopf: Der Frieden macht das Volk glücklich. 2. Kopf: O erhabener König, Vater Eures Volkes, Euer Wohltun erweckt in Europa Ruhm für Euren Thron.
- 38 Antoine de Rivarol: Lettre a Monsieur le Président de \*\*\* sur la globe Airostatique, sur les Têtes parlantes, & sur L'état présent de l'opinion publique à Paris [1783]. Zit. n. Rene Simmen: Der mechanische Mensch. Texte und Dokumente über Automaten, Androiden und Roboter, Zürich 1967, S. 69.
- 39 Vgl. Kempelen: Mechanismus der menschlichen Sprache (Anm. 17), S. 197 ff.



- 40 »Vous eres mon ami – je vous aime de tout mon Coeur, oder in der lateinischen Sprache: Leopoldus Secundus – Romanorum Imperator – Semper Augustus«. Ebd., S. 456.
- 41 Vgl. ebd., S. 71.
- 42 Felderer: Stimm-Maschinen (Anm. 24), S. 267.
- 43 Kempelen: Mechanismus der menschlichen Sprache (Anm. 17), S. 455.
- 44 Ebd. S. 183.
- 45 »Es verdient festgehalten zu werden, dass die Experimentalphonetik eben ihre Wurzeln in Kempelen, d. h. vor der Ausbildung der sprachvergleichenden Methode hat«. Eberhard Zwirner/Kurt Zwirner: Grundfragen der Phonometrie, Basel <sup>2</sup>1966, S. 86.
- 46 Kempelen: Mechanismus der menschlichen Sprache (Anm. 17), S. 179.
- 47 Wolfgang Scherer: Klaviaturen, Visible Speech und Phonographie. Marginalien zur technischen Einstellung der Sinne im 19. Jahrhundert, in: Friedrich Kittler/Manfred Schneider/Samuel Weber (Hg.): Diskursanalysen. 1. Medien, Opladen 1987, S. 39.
- 48 Kempelen: Mechanismus der menschlichen Sprache (Anm. 17), S. 455.
- 49 Eine zeitgenössische Sprechmaschinenkonstruktion von Erasmus Darwin, die er im Verlauf seiner Studien zur Kurzschrift anfertigte, lieferte nur die Laute P, B, M und A, konnte jedoch auch bereits »Mama« und »Papa« sagen. »At that time I contrived a wooden mouth with lips of soft leather, and with a valve over the back part of it for nostrils, both which could be quickly opened or closed by the pressure of the fingers, the vocality was given by a silk ribbon about an inch long and a quarter of an inch wide stretched between two bits of smooth wood a little hollowed; so that when a gentle current of air from bellows was blown on the edge of the ribbon, it gave an agreeable tone, as it vibrated between the wooden sides, much like a human voice. This head pronounced the p, b, m, and the vowel a, with so great nicety as to deceive all who heard it unseen, when it pronounced the words mama, papa, map, and pam; and had a most plaintive tone, when the lips were gradually closed [...] or which if built in gigantic form, might speak so loud as to command an army or instruct a crowd.« Erasmus Darwin: The Temple of Nature. Or: The Origin of Society, London 1803, S. 119 f.
- 50 Vgl. Robert Willis: On the Vowel Sounds and on Reed Organ Pipes, in: Trans. Cambridge Philosophical Society, Bd. 3 (1830), S. 231 ff. Deutsche Übersetzung: ders.: Über Vocaltöne und Zungenpfeifen. In: Wiedemann's Annalen der Physik und Chemie 24 (1832), S. 397–437, (hier: S. 398).
- 51 Willis: Über Vocaltöne (Anm. 50), S. 400.
- 52 Zum Nachbau vgl. Charles Wheatstone: Reed Organ-Pipes, Speaking Machines etc. [1837], in: ders.: The Scientific Papers of Sir Charles Wheatstone, London 1879, S. 348–367. Ein Exemplar der »Sprechenden Maschine« befindet sich im Deutschen Museum, München, Abteilung für Musikinstrumente.
- 53 Charles Wheatstone: On the Transmission of Musical Sounds through Solid Linear Conductors, and on Their Subsequent Reciprocation [1831], in: ders.: The Scientific Papers (Anm. 52), S. 62 f.; vgl. Hankins/Silverman: Instruments (Anm. 8), S. 202 f.
- 54 Vgl. Karl Friedrich Salomon Liskovius: Theorie der Stimme, Leipzig 1814, S. 29–35 (hier: S. 35).
- 55 Vgl. Johannes Müller: Handbuch der Physiologie des Menschen, Koblenz 1837.
- 56 Hermann Gutzmann: Physiologie der Stimme und Sprache, Braunschweig 1909, S. 70 f.
- 57 Vgl. Hankins/Silverman: Instruments (Anm. 8), S. 214 ff.
- 58 Vgl. Anonym: Faber's Sprechmaschine, in: Annalen der Physik 58 (1843), S. 175–176. Zu Faber vgl. Köster: Historische Entwicklung von Syntheseapparaten (Anm. 8), S. 142–148.
- 59 Vgl. Phineas T. Barnum: Struggles and Triumphs or: The Life of P. T. Barnum. Written by Himself, hrsg. v. George S. Bryan, 2 Bde., Bd. 1, New York 1927, S. 400.
- 60 An Joseph Fabers Euphonia können die unterschiedlichen Gewichtungen der wissenschaftlichen Forschung zu Sprechmaschinen als Auswahl eines wissenschaftlichen Umfelds erkannt werden, das Faber entweder in eine Reihe mit Thomas A. Edison stellt, so der amerikanische Physiker Alfred Marshall Mayer in seinem Text: On Edison's Talking-Machine, in: Popular Science Monthly 3/12 (1878), S. 719, oder Faber abwertend als »Nachahmer« bezeichnet (Joachim Gessinger), als sei der Schritt von der Artikulation einzelner »Sprachtöne« zu der ganzer Sätze nach Wunsch des Publikums wegen Kempelens grundlegenden Vorarbeiten zu Fabers Euphonia zu vernachlässigen. Vgl. Gessinger: Auge und Ohr (Anm. 8), S. 622.
- 61 Auch das akustische Spielzeug funktionierte ebenfalls ausschließlich nach akustischen Gesetzen; das erste Patent einer Sprechpuppe erhielt am 31. Januar 1824 der Erfinder des Metronoms, Jean Maelzel, in Paris, der eine Poupée parlante schützen ließ, die »Papa« und »Mama« sagen konnte. Nach einer Streckung des linken Arms der Puppe bewegte sich mit Federkraft ein kleiner Blase-

Daniel Gethmann

234

balg im Inneren, dessen ausströmende Luft zu einem Ventil geleitet wurde, das sich zweimal öffnete, um das Wort »Papa« erklingen zu lassen. Durch die Streckung des rechten Arms der Puppe wurde ein weiteres Ventil in der Luftleitung geöffnet, das dann den Gesamtmechanismus zum Erklingen des Worts »Mama« veränderte.

- 62 Vgl. Hermann von Helmholtz: Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik [1863], Braunschweig 1913, S. 194 ff.
- 63 Alexander Graham Bell: Prehistoric Telephone Days, in: National Geographic 41 (1922), S. 233–237. Die Resonanz eines Totenschädels nutzte in einer verbesserten Version des Bell'schen Apparates ebenfalls E. W. Scripture im Jahre 1905, dessen Erfindung auch Stimmbänder aus Gummi vorsah und vom Erfinder speziell für musikalische Aufführungen gedacht war. Vgl. Edward Wheeler Scripture: Report on the Construction of a Vowel Organ, in: Smithsonian Miscellaneous Collections 47 (1905), S. 360–364.
- 64 Vgl. Georges René Marie Marage: Petit Manuel de Physiologie de la Voix à l'Usage des Chanteurs et des Orateurs, Paris 1911, S. 93.

**Marcus Erbe**

## **SPRACHLICHE UND SPRACHLOSE STIMMEN IN DER ELEKTROAKUSTISCHEN MUSIK**

Die kompositorische Instrumentalisierung elektroakustischer Medien, die seit der zweiten Hälfte des 20. Jahrhunderts zu nachhaltigen Erweiterungen musikalischer Ausdrucksmittel führte, zog auch apparative Metamorphosen der menschlichen Stimme nach sich. Im Sinne einer Annäherung an die Elektronische Sprachkomposition als Partialtradition des hier ins Auge gefassten Gegenstandes sollen daher zunächst paradigmatische Stationen medialer Stimminszenierungen und vokaler Performanzen im Bezirk der Elektroakustischen Musik ganz allgemein erörtert werden, wobei der Versuch einer Skizzierung musikhistorischer Schlüsselstationen in dem gegebenen Zusammenhang freilich keinen Anspruch auf Vollständigkeit erheben kann. Vielmehr soll es darum gehen, einen Blick auf charakteristische Formationsweisen der Stimme in den jeweiligen Stücken zu werfen. Anschließend wird die Elektronische Sprachkomposition am Beispiel des Werkes *Epitaph für Aikichi Kuboyama* (1960–62) von Herbert Eimert, welches in seiner Funktion als instruktiver Prototyp des Genres das bis zu diesem Zeitpunkt wohl radikalste Modell der Überführung von Sprachklang in Musik verkörpert, einer genaueren Betrachtung unterzogen. Obwohl sich Eimerts Stück mit Rücksicht auf die angewandten Klangformungstechniken der Elektronischen Musik zurechnen lässt, enthält es keinen einzigen synthetisch hergestellten Klang. Sämtliche Spektren wurden allein aus dem rezitierten Text gewonnen. Eimert etablierte dergestalt eine eigenständige kompositorische Methode, die sich ihm zufolge durch die Neufundierung des Verhältnisses von Dichtung und Musik auszeichne, das angesichts der apparativen Sprachverarbeitung »keinen ›Dualismus‹ zwischen Wort und Klang zu überwinden braucht, weil es ihn nicht mehr kennt.«<sup>1</sup> Die in der zitierten Äußerung vorgenommene kompositionsgeschichtliche Standortbestimmung gibt Anlass, zum einen Eimerts Sprachkomposition, zum anderen aber auch weitere der im Rahmen des vorliegenden Textes behandelten Vokalwerke auf die jeweilige Operationalisierung des Verhältnisses zwischen Text und Musik hin zu befragen.

Das seitens Eimert exponierte Wort-Ton-Problem hat im Laufe der Musikgeschichte zahlreiche Umdeutungen erfahren. Man denke nur an die heftige Auseinandersetzung zwischen dem Musiktheoretiker Giovanni Maria Artusi und dem Komponisten Claudio Monteverdi zu Beginn des 17. Jahrhunderts, die sich an Monteverdis unkonventioneller Manier der Textbehandlung entzündete. Artusi, der die musikalische Textausdeutung einer strikten Befolgung der damals

verbindlichen Tonsatzlehre unterworfen wissen wollte, empörte sich über augenscheinliche Verletzungen satztechnischer Regeln in Monteverdis Madrigalen, ohne die eigentliche Intention des Komponisten erkannt zu haben. Dieser rechtfertigte seine satztechnischen Entscheidungen mit dem Hinweis auf den Text als Ausgangspunkt der musikalischen Erfindung und teilte die Musik in zwei Schreibarten auf: die ältere *Prima pratica*, die den zu singenden Text der Musik unterordnete sowie die neuere *Seconda pratica*, in der die Musik durch die besonderen Affektqualitäten eines spezifischen Textes und die Ausdruckskraft der Worte motiviert sei. Diese zweite Schreibweise markierte insofern einen geistesgeschichtlichen Umbruch, als die darin propagierte Idee des Textes, welcher zu genuinen musikalischen Einfällen inspiriert, für die gesamte Barockmusik leitend werden und noch weit darüber hinaus nachwirken sollte. Eine nicht minder signifikante Transzendierung überkommener Wort-Ton-Kategorien lässt sich bezüglich Eimerts *Epitaph* konstatieren, wobei es allerdings zu bedenken gilt, dass gegenüber früheren Exempeln europäischer Vokalmusik erstens der dem Werk zugrunde gelegte Text ein gesprochener und kein gesungener war und zweitens die musikalischen Beziehungen zum Textinhalt nicht mit Hilfe melodisch-rhythmischer Formeln, sondern – nur so viel sei schon hier gesagt – auf der Ebene der Klangmodifikation verankert wurden. Sowohl die dem Medium Stimme eingeschriebenen materiellen Eigenschaften als auch die phonetischen Merkmale des sprachlich artikulierten Textes bilden im Kontext des Stückes die Grundbausteine einer kompositorischen Faktur, deren ästhetischer Reiz speziell im Wechselspiel zwischen dem verständlichen Wort und den polymorphen, bislang ungehörten elektroakustischen Manipulationen desselben liegt.

I.

Neben einem Stück wie Ernst Kreneks vokal-elektronischem Pfingstoratorium *Spiritus Intelligentiae Sanctus* (1955), das traditionellen Wort-Ton-Konfigurationen grundsätzlich verpflichtet bleibt, stehen zahlreiche Werke, die die Grenzen dieser Relation in zunehmendem Maße überschreiten. Konfrontiert man Kreneks Pfingstoratorium mit weiteren Kompositionen, innerhalb derer vokales Material und elektronische Klänge aufeinander bezogen sind, so fällt in erster Linie die strenge Sprachgebundenheit der elektronischen Klangstrukturen nach dem Vorbild einer *Imitatione della parola* auf. Die Entfaltung des lateinischen Textes, der chorisch vorgetragen und nur minimalen Transformationen unterzogen wird, korreliert mit musikalischen Wendungen, die den Textsinn unterstreichen, in-

dem sie zentrale Begriffe hervorheben oder den Affektgehalt des Textes mime-tisch nachzeichnen. Reste einer derartigen Wort-Ton-Behandlung finden sich noch in Karlheinz Stockhausens *Gesang der Jünglinge* (1955–56), dessen Silben und Wörter dem *Gesang der Jünglinge im Feuerofen* aus dem 3. Buch Daniel entstammen und von einer Knabenstimme nach strikten Vorgaben deklamiert werden. Rezitiert etwa der Knabe bei Zeitindex 6'18" die Wortfolge »starrer Winter«, so entspricht dieser Starre das plötzliche Ausbleiben des zuvor noch sehr bewegten elektronischen Klanges. Diese Isolation der Stimme ist zudem auf einer weiteren Ebene wirksam, die dem Hörer der stereophonen CD-Reduktion zwangsläufig verschlossen bleibt. Beim Anhören des originalen Vierspurtonbandes<sup>2</sup> wird man indes bemerken, dass auch die räumliche Bewegung des Sprachklangs in einem der vier rings um den Zuhörer verteilten Lautsprecher einfriert, bis die elektronische Klangschicht nach einer dreisekündigen Generalpause mit einem klirrend-hohen Sinustonkomplex wieder einsetzt. Es waren vermutlich solche und ähnliche Veranschaulichungen der Textsemantik, die Dieter Schnebel zu dem Urteil veranlassten, das Werk »sei eigentlich noch ›vertont«.<sup>3</sup> Im Unterschied zu Kreneks *Spiritus Intelligentiae Sanctus* unterwarf Stockhausen gleichwohl die Sprache selbst dem kompositorischen Akt. Er zielte darauf ab, die gesungenen Laute den synthetischen Klangfarben im Sinne eines gemeinsamen Klangkontinuums anzunähern. Angeregt durch phonetische Fachstudien bei Werner Meyer-Eppler an der Bonner Universität nahm Stockhausen akustisch-phonetische Differenzierungen der Lautsubstanz vor. Kraft systematisch ermittelter struktureller und spektraler Korrespondenzen zwischen der menschlichen Stimme und den elektronischen Klangfamilien konnte er unterschiedliche Vermittlungsgrade bis hin zu einer Verschmelzung der beiden Sphären komponieren. Abbildung 1 zeigt in Sonagrammdarstellung einen Ausschnitt des Werkes, in dem unter gezielter Ausnutzung spektraler Überschneidungsgebiete der stimmlose Konsonant [s] in einen gefilterten Rauschkomplex überführt wird. Des Weiteren ist die für diese Passage so markante Akkumulation der Laute [s] und [t] anhand der markierten Frequenzspitzen deutlich zu erkennen, welche aus der mehrfachen Überlagerung

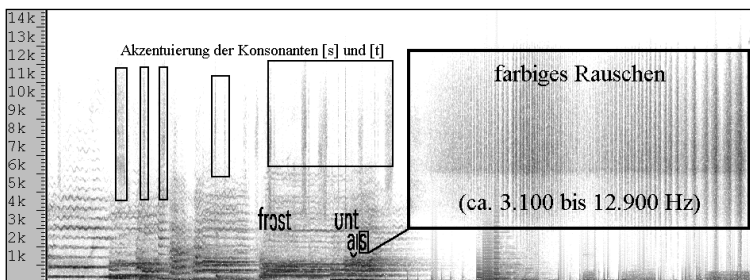


Abb. 1

Sonographische Repräsentation *Gesang der Jünglinge* (7'35" bis 7'55")

der Textpartikel »Frost« und »Eis« resultiert. Nicht zuletzt vermittelte Stockhausen mit dieser Emanzipation des konsonantischen Sprachanteils, dem in der herkömmlichen Vokalmusik eine rein akzidentelle Rolle zukam, der zeitgenössischen (wohlgemerkt nicht nur elektronischen) Vokalkomposition entscheidende Innovationsimpulse.

Einen nächsten Schritt auf dem Wege der Vermittlung zwischen stimmlicher Ausdrucksebene und elektronischer Klangwelt unternahm Luciano Berio mit der Realisation des Tonbandstückes *Visage* (1961). Doch anders als Stockhausen, der darauf bedacht war, die Kernaussage des apokryphen Textes trotz Silbenermutationen und anderer die Sprachsemantik destrukturierender Verarbeitungsstrategien zu erhalten – immerhin handelt es sich beim *Gesang der Jünglinge* um ein geistliches Werk, dessen religiöser Gehalt im Text beschlossen liegt und den Stockhausen durch die mehrfache Wiederholung der Zeilen »Preiset den Herrn« sowie »Jubelt dem Herrn« intentional verfestigte –, strebte Berio die emotive Affektion des Hörers auf der Grundlage textfreier Expressivität an, die zwischen Freude, Erregung, Furcht und Trauer changieren. Die breite Skala zuweilen extrem überformter artikulatorischer und parasprachlicher Merkmale in *Visage* entfesselt eine – so Christoph von Blumröders treffende Kennzeichnung – eigenständige »Emotionssemantik«,<sup>4</sup> die in den elektronischen Klangstrukturen des Werkes ihre Entsprechung findet. Nahezu sämtliche vokalen Performanzen der Sängerin Cathy Berberian (deren exzeptionelle Stimme dem Stück einen Großteil seiner musikalischen Wirkung verleiht) sind auf elektronische Klänge bezogen, die mal als Dialogpartner, mal als träumerische Reflexion und mal als Absorptionsmedium der sprachlosen Stimme fungieren.<sup>5</sup> Es ist hier gerade das Nicht-Gesagte, das zu jedem Zeitpunkt die emotionale Auseinandersetzung des Adressaten mit dem Gehörten provoziert, ganz im Sinne Roland Barthes:

Die menschliche Stimme ist tatsächlich der privilegierte (eidetische) Ort des Unterschieds: ein Ort, der sich jeder Wissenschaft entzieht, da es keine Wissenschaft gibt [...], die der Stimme gerecht wird: Man ordnet und kommentiert die Musik historisch, soziologisch, ästhetisch und technisch, es wird immer ein Rest bleiben, ein Zusatz, ein Lapsus, etwas Unausgesprochenes, das auf sich selbst verweist: die Stimme. [...] Es gibt keine menschliche Stimme auf der Welt, die nicht Objekt des Begehrens wäre – oder des Abscheus: Es gibt keine neutrale Stimme [...]. Jeder Bezug zu einer Stimme ist zwangsläufig einer der Liebe, und gerade deshalb bricht die Differenz der Musik, ihr Zwang zur Bewertung, zur Behauptung, in der Stimme hervor.<sup>6</sup>

Aber nicht nur in Berios *Visage* fungiert die Stimme als emotive Signatur eines bestimmten Menschen. Ebenso wie Berios Stück als Hommage an Cathy Berberian gehört werden kann,<sup>7</sup> lässt sich Makoto Shinoharas *Personnage* (1968/73) als ganz persönliches Porträt des ihm befreundeten Sängers Günther Weber begreifen. Es leuchtet unmittelbar ein, dass bereits die Titel beider Werke auf die Untrennbarkeit von Stimme und Identität der sich artikulierenden Person anspielen, wobei Shinoharas Stück den Physiognomiediskurs weiter vorantreibt: Aufgrund der zusätzlichen Integration akustischer Großaufnahmen genuin körperlicher Vorgänge ist *Personnage* Vokalkomposition und Körpermusik zugleich. Zudem durchbrach Shinohara die im Bezirk der autonomen Tonbandmusik unvermeidliche mediale Suspendierung der Koexpressivität von Stimme und Leib, indem er die räumliche Projektion des Tonträgers um einen live zu den Klängen agierenden Mimen erweiterte. Die Klimax des Stückes bildet eine Passage ab Zeitindex 4'45", die den Atem, Herzschlag, ja sogar den Schluckauf sowie das Schlürfen, Spucken und Schnarchen des Interpreten simultan erfasst und somit ein besonders eindrückliches Beispiel für die der Komposition inhärente Pluralität der Ausdrucksebenen darstellt.

Sämtliche der bislang berücksichtigten Werke verbindet das Moment der kompositorischen Ergründung elektroakustischer Transformationsmöglichkeiten spezifisch menschlicher Stimmen, hier mit dem Ziel der Auslotung textsprachlicher Qualitäten, dort mit der Absicht einer musikalischen Intensivierung emotionaler Archetypen. Gänzlich anders verfuhr György Ligeti. Mit dem überaus humorvollen Stück *Artikulation* (1957) gelang ihm die Komposition einer Kunstsprache, ohne auf stimmliches Material überhaupt zurückgreifen zu müssen. Diese auf den ersten Blick paradox anmutende Feststellung lässt sich dahingehend konkretisieren, dass sich Ligeti im Zuge der rein elektronischen Klangherstellung an nonverbalen Begleiterscheinungen der Rede, an Stimmcharakteristoren wie Lachen, Schluchzen, Seufzen, an Stimmsegregaten wie Pfeifen, Schnalzen, Atmen sowie an Stimmqualifikatoren orientierte, um synthetische Klangverläufe in Anlehnung an die mündlich realisierte Sprache hervorzubringen, die er sodann zu virtuellen Monologen, Dialogen und vielstimmigen Disputen verknüpfte. Eine Betrachtung unter dem Wort-Ton-Aspekt lässt die dabei entstandene imaginäre Gesprächsdramaturgie ebenso wenig wie die Singstimme in Berios *Visage* zu, da zwar sprachähnliche Strukturen beziehungsweise emotiv-vokale Gesten das jeweilige klangliche Geschehen dominieren, jedoch die textliche Vorlage als Referenzpunkt für die Analyse der stimmlichen Gestaltung fehlt.

Zum Abschluss der hier vorgenommenen querschnittartigen Präsentation vokaler Konfigurationen in der Elektroakustischen Musik sei auf zwei Stücke verwiesen, in denen die spektralen Merkmale der Stimme dank der Verwendung

stimm synthetischer Technologien tiefgreifenden Klangmutationen unterworfen werden konnten. John Chownings *Phoné* (1980–81) lässt sich auf diesem Gebiet als Pionierwerk bezeichnen, da Chowning darin zum ersten Mal die Verwandlung digital simulierter Stimmklänge in gleichermaßen synthetisch produzierte Instrumentalfarben ergründete. Formal gliedert sich die Komposition in eine Folge vokaler Klangzellen weiblicher und männlicher Timbres, die in stets neuen Verbindungen mit vertrauten Klangfärbungen aufgehen. Bemerkenswerterweise bewegt sich die musikalische Behandlung der künstlich erzeugten Stimmen in überkommenen Bahnen, indem ausschließlich metrisch streng gebundene Vokallaute als Melodieträger fungieren, obwohl das in Anschlag gebrachte Prinzip der FM-Synthese mit Sicherheit kühnere Klangexperimente zugelassen hätte. Es kann an dieser Stelle nicht entschieden werden, ob Chownings anachronistische Handhabung des vokalen Parts durch die Erzielung eines möglichst hohen Verschmelzungsgrades mit den primär harmonischen Frequenzspektren der instrumentalen Klangereignisse motiviert war. Der springende Punkt jedoch ist, dass sich trotz der Abwesenheit eines Textes infolge der Vokaldominanz ein weitgehend bekanntes Klangbild einstellt und somit die Komposition wie kaum ein anderes der erwähnten Werke (Kreneks Pflingstoratorium einmal ausgenommen) in die Nähe überlieferter vokaler Musik rückt. Weitaus differenziertere hybride Mixturen der menschlichen Stimme mit synthetischen Klangfeldern verwirklichte Jean-Claude Risset in seinem Stück *Invisible* für Sopran und zweikanaliges Tonband (1996). Die Stimme der Sopranistin Irène Jarsky führt den Hörer durch eine surreale akustische Umgebung, entgleitet diesem aber immer wieder aufgrund ihrer spektralen Verschmelzung mit Windgeräuschen und elektroakustischen Klangstrukturen. Zur zeitlichen Mitte des Werkes hin verkündet die Interpretin das Motto »La voix et le son s'harmonisent«, das anschließend in mannigfaltigen Durchführungen musikalisch entwickelt wird. Farblich schillernd bewegen sich glockenähnliche Klänge im Rhythmus eines gesprochenen Textes, überlagert von kunstvollen Melismen und vocoderartigen Vervielfachungen der Gesangsstimme.

## II.

Wie im ersten Teil des vorliegenden Textes bereits deutlich geworden sein dürfte, vollzog sich die Auseinandersetzung mit dem Phänomen Stimme in der Elektroakustischen Musik im Sinne einer produktiven Aneignung technisch induzierter Möglichkeiten fernab jeglichen Vertrauensverlusts in die Individualität der Stimme, wie ihn etwa Reinhart Meyer-Kalkus mit Blick auf die zeitgenössischen



vokalen Künste behauptet.<sup>8</sup> Und wenn auch die musikalischen Resultate in ästhetischer Hinsicht höchst unterschiedlich ausfielen, so konnten doch diverse Verbindungslinien zwischen den Werken, wie etwa die schöpferische Vorstellung eines Kontinuums zwischen Stimmklang und elektronischem Klang, aufgezeigt werden. Vor dem Horizont der seit Mitte der 1950er Jahre angestrebten Verschmelzung vokaler und elektronischer Klangereignisse erschien es nur konsequent, die gesamte Klangwelt eines Werkes allein aus der Sprache hervorgehen zu lassen; eine Herausforderung, der sich Herbert Eimert im *Epitaph für Aikichi Kuboyama* stellte. Zum Movens der Tombeau-Komposition wurde das Schicksal des japanischen Funkoffiziers Kuboyama, der zusammen mit 23 weiteren Besatzungsmitgliedern eines Thunfisch-Trawlers am Morgen des 1. März 1954 in die Nähe eines US-amerikanischen Wasserstoffbomben-Tests vor den Marshall-Inseln geriet und an den Spätfolgen der radioaktiven Verseuchung wenige Monate später verstarb.<sup>9</sup> Der Fall Kuboyama wurde zum Gegenstand internationaler Medienberichterstattung. Die weltweite Anteilnahme am Tod des Fischers, welcher insbesondere auf Seiten der japanischen Bevölkerung Erinnerungen an Hiroshima und Nagasaki weckte, verlieh der internationalen Anti-Atombewegung neuen Auftrieb. Als deren Mitinitiator nahm Günther Anders in seinem Buch *Die Antiquiertheit des Menschen* das Einzelschicksal Kuboyamas zum Anlass, die kollektive Bedrohung der Menschheit durch Atomwaffen herauszustellen. Anders zufolge schließe jeder nukleare Versuch die Eventualität der massenhaften Vernichtung mit ein:

Die Effekte sind so ungeheuer, dass im Moment des Experiments das ›Laboratorium‹ ko-extensiv mit dem Globus wird. Das aber bedeutet nichts anderes, als dass zwischen ›Probe‹ und ›Durchführung‹ zu unterscheiden, seinen Sinn verloren hat; dass jedes ›Experiment‹ zu einem ›Ernstfall‹ geworden ist. Tatsächlich haben ja die zahlreich durchgeführten ›Experimente‹ bereits ihre Effekte gehabt: dass das erste Opfer der Wasserstoffbombe, der japanische Fischer Aikichi Kuboyama, der im August 1954 starb, experimentell verschieden sei, wird man ja wohl nicht behaupten können; und ebenso wenig, dass der japanische Kutter Fikuryumaru, der 130 km vom Detonationszentrum entfernt, verseucht wurde, sich experimenti causa dort herumgetrieben habe.<sup>10</sup>

Unter Berufung auf den Sydney Chronicle vom 3. März 1955 druckte Anders eine deutsche Übersetzung der vermeintlich authentischen Inschrift auf dem Gedenkstein Kuboyamas ab. Wie man heute weiß, hat Anders jene Zeilen gleichsam als

literarisches Echo seiner philosophischen Reflexionen frei erfunden.<sup>11</sup> Eimert machte die Verse zur Grundlage seiner Komposition. Zusammen mit einer knappen Erläuterung der Todesumstände Kuboyamas stellte er das durch den Schauspieler Richard Münch vorgetragene Gedicht dem eigentlichen Werk voran:

Du kleiner Fischermann, / wir wissen nicht, ob du Verdienste hattest. / (Wo kämen wir hin, wenn Jedermann Verdienste hätte?) / Aber du hattest Mühen wie wir, / wie wir irgendwo die Gräber deiner Eltern, / irgendwo am Strande eine Frau, die auf dich wartete, / und zu Haus die Kinder, die dir entgegenliefen. / Trotz deiner Mühen / fandest du es gut, da zu sein. / Genau wie wir. Und recht hattest du, Aikichi Kuboyama / Du kleiner Fischermann Aikichi Kuboyama, / wenn auch dein fremdländischer Name keinen Verdienst anzeigt, / wir wollen ihn auswendig lernen für unsere kurze Frist / Aikichi Kuboyama. / Als Wort für unsere Schande / Aikichi Kuboyama. / Als unseren Warnungsruf / Aikichi Kuboyama. / Aber auch, / Aikichi Kuboyama, / als Namen unserer Hoffnung: Denn ob du uns / vorangingst mit deinem Sterben oder nur / fortgingst an unserer statt – / nur von uns hängt das ab, auch heute noch, / nur von uns, deinen Brüdern, / Aikichi Kuboyama.<sup>12</sup>

Die seitens Eimert angestrebte, bereits eingangs skizzierte Überwindung des Dualismus zwischen gesprochenem Text und musikalischem Klang konnte mit Hilfe äußerst ausgeklügelter Tonbandmanipulationen und Filterungstechniken erzielt werden, deren vollständige Aufzählung den Rahmen dieses Beitrags sprengen würde.<sup>13</sup> Im gegebenen Zusammenhang genügt ein exemplarischer Blick auf grundlegende Metamorphosen des sprachlichen Ausgangsmaterials, um zentrale realisationspraktische Verfahrensweisen nachzuvollziehen. Die Klänge des *Epitaphs* basieren im Wesentlichen auf vier Arten der elektroakustischen Variantenbildung: Filterung, Spreizung, Impulsherstellung und Transposition. Ab Zeitindex 3'33" des originalen Mehrspurbandes lassen sich zum ersten Mal sogenannte Artikulationsimpulse vernehmen, die in unterschiedlichen Ausprägungen und Komplexitätsgraden im Verlauf der Komposition wiederholt begegnen. Diese stoßartigen Klangverläufe konnten vermittels der gezielten Unterdrückung bestimmter spektraler Sprachanteile bei gleichzeitiger Hervorhebung anderer Teilbereiche gewonnen werden und führten je nach Einstellung der Filterparameter zu verschiedenen Graden der Sprachverständlichkeit (so schält sich bei 3'42" der Textausschnitt »auswendig lernen für unsere kurze Frist« aus den abstrakten Artikulationsklängen heraus). Extrem tiefe Klangfolgen, die Eimert selbst als posau-

nenartig charakterisierte, setzen bei 14'34" in mehrfacher Überlagerung ein. Bei nur oberflächlichem Zuhören fällt es schwer zu glauben, dass auch diese imposanten Klänge stimmlichen Ursprungs sind. Schenkt man jedoch den rhythmischen Konstellationen ein wenig mehr Aufmerksamkeit, so wird man feststellen, dass der Tonfall des Gedichts trotz der Tilgung des konsonantischen Sprachanteils sowie der Abwärtstransposition und zeitlichen Spreizung des Vokalbestandes subliminal erhalten blieb. Um perkussive, hart artikulierte Klänge aus dem stimmlichen Material herzuleiten, wurde ein Impulsgenerator verwendet. Neben dessen Funktion als Klanggeber vermag das Gerät den Amplitudenverlauf eines eingespeisten Klangvorgangs zu verändern. Hierzu erzeugt es Schwingungen im unhörbaren Bereich, die das akustische Rohmaterial periodisch zerhacken. Die Geschwindigkeit dieser Modulation lässt sich am Generator stufenlos regeln. Mit zunehmender Steigerung ereignet sich solch ein Impulsaufbau von 11'02" bis 12'40". Extreme Accelerandi und Ritardandi einzelner Wörter wurden mit Hilfe rigoroser Modifikationen der Bandlaufgeschwindigkeit realisiert, was zu raketenartigen Aufschwüngen ausgewählter Textfragmente führte, wie sie bei Zeitindex 18'16" sowie am Schluss der Komposition in Erscheinung treten.<sup>14</sup>

Dass phonetische Vorüberlegungen in die Behandlung des sprachlichen Materials eingeflossen sind, liegt auf der Hand. Und auch hier wieder hat Meyer-Eppler als Theoretiker Pate gestanden, dessen enge Zusammenarbeit mit Eimert bis ins Jahr 1951 zurückdatiert. Dennoch darf nicht übersehen werden, dass ein Teil der Sogwirkung, die die Klänge des *Epitaphs* auf den Hörer ausüben, weniger phonetischen Materialanalysen als der Klangphantasie des Toningenieurs Leopold von Knobelsdorff geschuldet ist, dem Eimert die klangliche Realisation des Stückes übertrug. Aus der vergleichsweise monoton vorgetragenen fiktiven Grabinschrift fertigte von Knobelsdorff in spielerischer Erkundung der Studioapparatur weit mehr als 90 Klangstrukturen, die ihre Expressivität einzig aus der apparativen Behandlung des Ausgangsmaterials beziehen.<sup>15</sup> Indessen darf man mit Sicherheit davon ausgehen, dass sich die musikalische Durchschlagskraft des *Epitaphs* auf die handwerkliche Kompetenz Eimerts zurückführen lässt, der den mahnenden Unterton der literarischen Vorlage akustisch zu profilieren vermochte. Ein in diesem Kontext äußerst eindrucksvoller Abschnitt (17'25" bis 18'22") vereinigt Imitationen schriller Schreie mit dem Heulen sirenenartiger Spektren sowie einem allmählich anwachsenden, ohrenbetäubenden Klangwirbel. Der musikalisch-rhetorischen Figur einer *Abruptio* gleich, deren Bedeutungsgehalt bereits im Zusammenhang barocker Affektmusik mit der Sphäre des Todes assoziiert war, reißt das klangliche Geschehen bei Minute 17'53" plötzlich ab. In die Stille hinein ruft die mehrfach mit sich selbst überlagerte Sprecher-

stimme »Wo kämen wir hin«, woraufhin sich die apokalyptisch anmutende Musik mit aller Gewalt fortsetzt, bis schließlich Kaskaden rauschhafter Klangwirbel in die Vokabel »Warnungsruf« einmünden.

Will man nun das im *Epitaph* geschaffene Verhältnis zwischen Text und musikalischer Faktur einer Bewertung unterziehen, so muss man einerseits Eimerts Selbsteinschätzung einer Neufundierung desselben beipflichten. Die intermediale Vermittlung beider Ebenen, die erst durch die konsequente Applikation der analogen Studioteknik auf Sprache möglich wurde, könnte inniger kaum sein. Andererseits lässt sich nicht übersehen, dass ungeachtet der Ausschöpfung fortschrittlicher Transformationstechniken explizite Verbindungen zum musikalischen Erbe bestehen. Der Rekurs auf musikalisch-rhetorische Figuren zwecks Untermuerung der Textbedeutung wurde bereits angedeutet.<sup>16</sup> Ferner ist die Faktur des Werkes von Eimerts geradezu humanistischer Hochschätzung des Wortes beherrscht, die die komponierten Klanggebilde bei aller Sezierung des sprachlichen Materials immer wieder zum Wort selbst zurückfinden lässt<sup>17</sup> und auch den Versrhythmus, der in zahlreichen Klangverbindungen unterschwellig anwesend bleibt, im Einklang mit der von der späten Renaissancetheorie geforderten engen Verzahnung von Dichtung und Musik nicht negiert. Nur noch am Rande seien hier sowohl die Identifikation der Musik mit der leidenden Person als auch die Einbeziehung der Adressatenwirkung vermerkt, die eine musikhistorische Parallelstelle in Monteverdis affektgeladenen Madrigalen des anbrechenden 17. Jahrhunderts haben.

Angesichts der oben umrissenen akustischen Evokationen der Katastrophe und des Vernichtungsschreckens nimmt es überhaupt nicht Wunder, dass der Gehalt des Werkes immer wieder in der Zergliederung der *Vox humana* als Symbol einer nuklearen Verwüstung des Individuums erblickt wird. Heinz Gramann beispielsweise hört die langgezogenen Glissandostrukturen im ersten Teil der Komposition als Bomberflugzeuge und assoziiert dementsprechend mit den tief gefilterten und abwärts transponierten Vokalen den Aufprall der Bombe.<sup>18</sup> Auch Hubert Steins betont das Vermögen dieser Musik, konnotative Bedeutungsfelder zu verfestigen:

Kristalline Wortsplinter, scheinbar von einer Detonationswelle erfasst, bersten wie eine Glasscheibe. [...] Die allgegenwärtige unterirdische Präsenz der Sprecherstimme scheint aus der Tiefe der Unterwelt zu uns zu flehen. Dass wir uns in einer akustischen Inszenierung des Weltgerichts befinden könnten, wird zumindest durch die posaunenartigen Klänge zum Ende der Komposition bekräftigt.<sup>19</sup>

Der Blickwinkel, den diese Interpretationen auf das Werk eröffnen, führt unweigerlich zu der Frage, ob Eimert nicht ebenso gut auf eine Verknüpfung der gesprochenen Grabinschrift mit elektronischen und konkreten Klängen hätte zurückgreifen können, um die gehaltliche Intention verstehbar zu machen. Möglicherweise wäre sogar ein gänzlich anders geartetes textliches und stimmliches Material zur Herleitung der akustischen Schreckensbilder geeignet gewesen. Dies suggeriert jedenfalls Hans G. Helms in einer Vorlesung über *Komposition in Sprache* aus dem Jahr 1964.<sup>20</sup> Darin polemisiert er gegen die seinerzeit aktuellen Tendenzen der Sprachkomposition und wirft ihren Exponenten eine Deformation der Dichtung vor. Zeitgenössische vokale Werke seien inkommunikabel, die literarischen Vorwürfe austauschbar beziehungsweise durch asemantisches phonetisches Material ersetzbar. Allein der Werktitel diene oftmals als Ersatz für die abhanden gekommene Bedeutungsebene. Helms Einwände sind insoweit berechtigt, als dass sich prinzipiell jeder Sprachlaut unter Zuhilfenahme entsprechender technischer Gestaltungsmittel in einen musikalisch verwertbaren Klang verwandeln lässt. Sofern aber die semantische Dimension des Materials mit der Faktur des Musikwerks in Einklang stehen soll, kann die der Komposition zugrunde gelegte Textsubstanz nicht beliebig sein. Unterzieht man im Hinblick darauf die Klangformungsprozesse im *Epitaph* einer näheren Betrachtung, wird man feststellen, dass das gesprochene Wort und der entsprachlichte Klang organisch auseinander erwachsen; dass sich der kompositorisch intendierte Gehalt anhand der Interaktion der semantischen Sprache, die sich im Erklingen der Musik als Moment des Wiedererkennens und Erinnerns in regelmäßigen Abständen spürbar mitteilt, und des elektronisch verarbeiteten Stimmklanges ersehen lässt. Zur Verdeutlichung sei ein 20 Sekunden dauernder Klangkomplex zu Beginn des Stückes herangezogen. Die Stimme des Sprechers, die ihrer nahen Mikrophonierung halber Intimität erzeugt, wird von der metallischen, infolge der starken Verhallung den Eindruck der Ferne hervorrufenden Klangwelt aufgesogen, bis aus dem latent vorhandenen Wortstrom<sup>21</sup> die Vokabel »sterben« sich löst. Wie weiter oben schon angeklungen ist, ließen sich noch zahlreiche weitere Klangereignisse innerhalb des Werkes benennen, an denen sich sinnhafte Elemente des gesprochenen Textes sowie spektrale Spuren<sup>22</sup> der sprachlichen Stimme erhalten haben. Nicht selten wirkt das Wort via Transformation sinnvoll – ein Effekt, der sich der fortwährenden Oszillation zwischen verständlicher Sprache und verklanglichter Stimme verdankt. Die elektroakustische Metamorphose des gesprochenen Wortes potenziert mithin dessen Bedeutungsgehalt und gestattet unter Ausnutzung klanglich-konnotativer Appelle eine musikalische Re-Lektüre des literarischen Prätextes. Eimert selbst machte in diesem Zusammenhang auf seine *Sechs Studien*

(1962) aufmerksam, die ebenso wie das *Epitaph* aus dem Lautbestand des Anderschen Gedichtes komponiert wurden. Die kompakten Stücke lassen jeden semantisierenden Kontext vermissen, da in ihnen die Vernehmbarkeit des gesprochenen Wortes bewusst ausgeklammert bleibt und der Sprachklang ausschließlich als ästhetisierte Klangrealie dargeboten wird.<sup>23</sup> Vergleicht man die sechs locker aneinandergereihten Einzelstudien mit dem *Epitaph*, so ergeben sich direkte Übereinstimmungen hinsichtlich der Klangdisposition dergestalt, dass komplette Teilabschnitte aus den Studien als eigenständige Materialschichten des *Epitaphs* – in der Regel unverändert, zuweilen aber auch geringfügig geschnitten oder in der Lautstärke angepasst – wiederkehren. Die folgende Tabelle soll diesen Sachverhalt illustrieren:

<i>Sechs Studien</i>	<i>Epitaph für Aikichi Kuboyama</i>	Kommentar
S3: 0' bis 0'15"	2'44" bis 2'59"	Eröffnungsklang <i>Epitaph</i> nach der Rezitation des Gedichts
S4: 0' bis 1'40"	11' bis 12'40"	Impulsaufbau
S4: 1'49" bis 2'27"	13'19" bis 13'57"	rhythmisch akzentuierte Wortsplitter
S5: 0' bis 0'44"	8'11" bis 8'39"	Glissando-Strukturen
S6: 0' bis 0'54"	17'24" bis 17'40"	schrille Rückkopplungen vokaler Sprachanteile
S6: 0'55" bis 1'46"	15'04" bis 15'45"	Spreizungen des Stimmklangs

Die soeben festgestellten frappanten Übereinstimmungen lassen Zweifel an Eimerts chronologischer Einordnung der *Sechs Studien* aufkommen (ihm zufolge entstanden diese erst im Anschluss an das *Epitaph*).<sup>24</sup> Vor dem aufgezeigten Hintergrund – so die abschließende These – erwecken die elektroakustischen Miniaturen vielmehr den Anschein textfreier, auf nichts Gegenständliches verweisender Vorübungen zum *Epitaph*, die Eimert einen Grundstock elaborierter musikalischer Abläufe für die Realisation des Werkes lieferten, welche durch die Herstellung wortsemantisch stark aufgeladener Klanggebilde gehaltlich angereichert wurden. Gestützt wird diese Annahme durch den im Titel akzentuierten Etüdencharakter der *Sechs Studien*, denn gerade im Bereich der Elektroakustischen Musik diente die Studie stets der Erprobung neuartiger Klangformungsmöglichkeiten, denen entweder im weiteren Schaffen des Komponisten eine besondere Stellung zukam oder die verworfen wurden, da sie dessen ästhetischem Anspruch nicht Stand zu halten vermochten. Die nachstehende Abbildung 3 bezieht sich auf die unterste Tabellenzeile und führt die Kombination einer Passage der sechsten Studie mit eigens für das *Epitaph* erarbeiteten wortbezogenen Spektren vor Augen. Deutlich hervorgehoben ist die Erweiterung der präexistenten Klangschicht um grell resonierende Sprachklänge, aus deren Fluss einzelne Silben

des Namens Aikichi Kuboyama flehentlich emportauchen. Eine derartig planvolle Berücksichtigung der sprachlichen Sinnebene erst konnte das wechselseitig sich bedingende Ineinander von teils verständlicher, teils dekomponierter Sprache und musikalisierter Stimme als solches identifizierbar werden lassen. Im hörenden Nachvollzug der Komposition wird somit die Mediendifferenz zwischen der Textvorlage und ihrer musikalischen Transformation aufgrund der Signalwirkung stimmlicher Reste sowie sprachlicher Bruchstücke auditiv erfahrbar. Die sich im Verlauf des Stückes unablässig ereignende Entsemantisierung des Textes, der nicht zufällig in seiner ursprünglichen, unverarbeiteten Gestalt gleichsam als Referenzpunkt für den Hörer noch vor dem eigentlichen Werk erklingt, lässt sich in Anbetracht der aufgezeigten Zusammenhänge zugleich als Resemantisierung des textlichen Gehalts interpretieren. Denn der destruktive Impetus, den Helms den Protagonisten der Elektronischen Sprachkomposition zum Vorwurf machte, zielte im Falle des *Epitaphs* keineswegs auf eine Negation der Dichtung. Im Gegenteil: Die seitens Eimert und seines Assistenten von Knobelsdorff praktizierte phonetische und spektrale Dekomposition der sprachlichen Stimme vollzog sich stets mit Blick auf die Möglichkeiten ihrer musikalischen Neuordnung innerhalb eines Geflechts aus sinnhafter Sprache und sinnlichem Klang. Dieses weist gegenüber der literarischen Vorlage ganz eigene Erlebnisqualitäten auf, die mit sprachlichen Mitteln allein nicht hervorzubringen sind. Auch wenn sich nicht leugnen lässt, dass der letztgenannte Aspekt auf Vokalmusik im Allgemeinen zutrifft, so hat das musikalische Resultat der Sprachkomposition mit einer Textvertonung im herkömmlichen Sinne nur wenig gemein, da sich die emotionale und somatische Aktivierung des Zuhörers als Folge einer akustischen Inskription des freien Klangexperimentes im elektronischen Studio einstellt – jenseits der strengen Kodierung von Stimme durch notenschriftliche Aufzeichnungen.

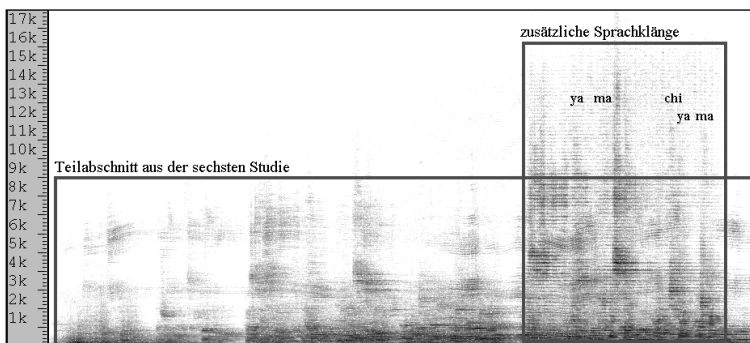


Abb. 2

Sonographische Repräsentation *Epitaph für Aikichi Kuboyama* (15'04" bis 15'45")

- 1 Herbert Eimert: Notizen zum Epitaph und den Sechs Studien Beiheft zur Wergo-Schallplatte 60014, S. 3. Ex aequo schrieb Eimerts langjähriger Schüler Hans Ulrich Humpert: »Die Vorstellung kontinuierlicher Übergänge zwischen Musik und Sprache führte dazu, Sprache als reine Klänge zu betrachten und zu behandeln, was das alte Verhältnis von Wort und Ton in ein neues verwandelte: Wort als Ton.« (Stimmen – Chöre – Klänge, in: Bernd Enders/Niels Knolle (Hg.): KlangArt-Kongress 1995, Osnabrück 1999, S. 151–158, hier: S. 151 f.).
- 2 Neben einer multiphonen Aufnahme vom *Gesang der Jünglinge* waren dem Verf. mehrspurige Digitalkopien einiger weiterer der im Rahmen des vorliegenden Aufsatzes angesprochenen Kompositionen zugänglich, die im Klangstudio des Musikwissenschaftlichen Instituts der Universität zu Köln archiviert sind. Sämtliche der im weiteren Verlauf des Textes gemachten Zeitangaben beziehen sich auf die entsprechenden Mehrkanalbänder abzüglich des jeweiligen Bandvorlaufs von einer Minute.
- 3 Zit. nach Werner Klüppelholz: Sprache als Musik. Studien zur Vokalkomposition seit 1956, Herrenberg 1976, S. 60.
- 4 Christoph von Blumröder Die Anfänge der Elektronischen Musik in Italien, in: ders./Norbert Bolin/Imke Misch (Hg.): Aspetti musicali. Musikhistorische Dimensionen Italiens 1600 bis 2000, Köln 2001, S. 263–270 (hier: S. 268).
- 5 Siehe das Beiheft zur CD, o.S.
- 6 Roland Barthes: Die Musik, die Stimme, die Sprache, in: ders.: Der entgegenkommende und der stumpfe Sinn. Kritische Essays III, Frankfurt/M. 1990, S. 279–285 (hier: S. 280).
- 7 Vgl. von Blumröder: Die Anfänge (Anm. 4), S. 268.
- 8 Reinhart Meyer-Kalkus: Stimme und Sprechkünste im 20. Jahrhundert, Berlin 2001, S. 61.
- 9 Zu den näheren Umständen siehe Ralph E. Lapp: Die Reise des Glücklichen Drachen, Düsseldorf 1958.
- 10 Günther Anders: Die Antiquiertheit des Menschen. Bd. I: Über die Seele im Zeitalter der zweiten industriellen Revolution [1956], München 1994, S. 260.
- 11 Vgl. Helmut Kirchmeyer: Kleine Monographie über Herbert Eimert (= Abhandlungen der Sächsischen Akademie der Wissenschaften zu Leipzig, Philologisch-Historische Klasse 75/6), Stuttgart/Leipzig 1998, S. 44.
- 12 Anders: Die Antiquiertheit des Menschen (Anm. 10), S. 346.
- 13 Eine ausführliche Beschreibung findet sich bei Hans Ulrich Humpert: Elektronische Musik. Geschichte – Technik – Kompositionen, Mainz 1987, S. 162–174.
- 14 Audiokommentare Eimerts zur Klanggewinnung können unter <http://www.elektropolis.de/audio.html> (24.2.2003) heruntergeladen werden.
- 15 Im Rahmen eines Gesprächs, das von Knobelsdorff am 14. Januar 1999 mit Studenten des Musikwissenschaftlichen Instituts der Universität zu Köln führte und dessen Mitschnitt dem Verf. vorlag, vermochte der Toningenieur sich kindlicher Sprachspiele zu erinnern, die er mit seinem Bruder durchführte und die ihn zur Realisation markanter Silbenpermutationen anregten, was auf bisweilen sehr persönliche Momente im Kontext der Klanggenese schließen lässt.
- 16 Auf eine nähere Beschreibung weiterer, hinsichtlich des Stückes durchaus relevanter Klassen wie Hypotyposis, Anabasis oder Circulatio musste aus Platzgründen verzichtet werden.
- 17 »Aber auch das Wort selbst, mit seinem Ausdruck, Sinn und Erlebnisinhalt, schlägt immer wieder durch und steigert sich mit dem ganzen Gewicht seiner Bedeutung zur Sinnbezeichnung einzelner Worte oder Wortverbindungen [...]« (Eimert: Notizen zum Epitaph, Anm. 1, S. 3).
- 18 Vgl. Heinz Gramann: Die Ästhetisierung des Schreckens in der europäischen Musik des 20. Jahrhunderts (= Orpheus-Schriftenreihe zu Grundfragen der Musik 38), Bonn 1984, S. 154.
- 19 Hubert Steins: Der Sprachklang als Inschrift. Herbert Eimerts »Epitaph für Aikichi Kuboyama«, in: Lukas Hellermann (Hg.): KlangRaum Kreuzeskirche 99, Saarbrücken 1999, S. 57–64 (hier: S. 63 f.).
- 20 Hans G. Helms: Komposition in Sprache, in: Heinz-Klaus Metzger/Rainer Riehn (Hg.): Darmstadt-Dokumente I (= Musik-Konzepte Sonderband), München 1999, S. 241–250.
- 21 Vgl. Eimert: Notizen zum Epitaph (Anm. 1), S. 4.
- 22 Um Verwechslungen mit dem Spur-Begriff Sybille Krämers auszuschließen, der in den aktuellen Medienkulturwissenschaften eine prominente Rolle spielt, sei ausdrücklich betont, dass die in Rede stehenden stimmlichen Bruchstücke absichtsvoll im klanglichen Gefüge pointiert wurden. Dagegen abzugrenzen sind eng an die Materialität der Realisationsmedien gebundene, nicht intentionale Spuren, die Rückschlüsse auf die während des Klangproduktionsprozesses verwendeten Apparate zulassen und sich als charakteristische Begleiterscheinungen der aktivierten Magnetton-



technik dem *Epitaph* eingewohnt haben (siehe Sybille Krämer: Das Medium als Spur und Apparat, in: dies. (Hg.): *Medium – Computer – Realität. Wirklichkeitsvorstellungen und Neue Medien*, Frankfurt/M. 1998, S. 73–94).

- 23 Lediglich die zweite Studie lässt den stimmlichen Ursprung der Klänge ansatzweise erahnen, jedoch bleibt die Semantik der Textvorlage auch unter größten Höranstrengungen verborgen.
- 24 Vgl. Eimert: Notizen zum *Epitaph* (Anm. 1), S. 4.

Jin Hyun Kim

**DIE SINGSTIMME ALS AUSDRUCKSZEICHEN.**

**ZUR MEDIALEN FUNKTION DER STIMME IN DER MUSIK**

**1. DIE VORPRÄDIKATIVE AUSDRUCKSFUNKTION DER STIMME UND DER MUSIK**

Das Phänomen der Stimme, das in den schriftsprachlich orientierten linguistisch-zeichentheoretischen Forschungen der vergangenen Jahre wenig thematisch geworden ist,<sup>1</sup> gewinnt im gegenwärtigen medientheoretisch-kulturwissenschaftlichen Diskurs insbesondere im Hinblick auf die Wiederentdeckung der Performativität der Rede neue Aufmerksamkeit. Der Medialität der Stimme wird in diesem Debatten-Kontext – jenseits der semantisch-prädikativen Funktionen der Kommunikation – eine vorprädikative Leistung zugeschrieben, der sich ein in der Rede »nicht intendierte[r] Überschuss an Sinn« verdankt.<sup>2</sup> Diese Überlegungen bilden – insbesondere hinsichtlich des Aspektes einer der Stimme geschuldeten vorprädikativen Sinnerzeugung – den Ausgangspunkt des vorliegenden Beitrags zur musikalischen Stimme. Die Untersuchungsperspektive richtet sich dabei auf eine mediale Ausdrucksfunktion der Stimme, die bereits in gewissem Umfang der Sprechstimme zukommt, die aber erst in der Singstimme ihren genuinen Ort zu erhalten scheint, und zwar insofern, als besonders die Singstimme es erlaubt, im Vollzug der »raum-zeitlich situierten Performanz des Zeichenverhaltens«<sup>3</sup> innere Zustände von Zeichenbenutzern zum Ausdruck zu bringen.

Die vorprädikative Funktion der Stimme, deren Leistung im Rahmen der sprachlichen Kommunikation darin gesehen werden kann, innere Zustände des Sprechers zum Ausdruck zu bringen, wurde bereits im Bühlerschen Organonmodell der Sprache als »Ausdrucksfunktion« von der Appell- und Darstellungsfunktion unterschieden. Der Stimme als »konkretem Schallphänomen« kommen, sofern sie zeichenhaft verwendet wird, die genannten drei Funktionen als »semantische Funktionen« zu: während »Darstellung« und »Appell« als »Symbol« und »Signal« die Beziehungen zu Gegenständen und Sachverhalten sowie zum Hörer organisieren, kommt es dem »Symptom« »kraft seiner Abhängigkeit vom Sender« in seiner Ausdrucksfunktion zu, dessen Innerlichkeit auszudrücken.<sup>4</sup> Die Stimme, ebenso wie ihr Produkt, der Laut, ist also Bühler zufolge »ein mediales Phänomen«, das nicht nur eine Relation zu den Dingen oder den Sachverhalten, sondern auch »je eine eigene Zeichenrelation« zum Adressaten sowie zum Subjekt der Sprechhandlung aufweist.<sup>5</sup> Obwohl die »Dominanz der Darstellungsfunktion der Sprache« für Bühler unbestritten gilt,<sup>6</sup> erhalten Appell- und Aus-

drucksfunktion in der Sprechsituation eine je eigene zeichensemantische Rolle. Beide können sich des »Tones« bedienen, der zwar »irrelevant ist für die Darstellung«, aber als »musikalische Modulation« der Stimme eine eigene ausdruckssemantische Kraft entfaltet: »Darum kann der Sprecher im Musikalischen seiner Seele die Zügel schießen lassen, kann den Ärger oder die Freude, wenn es sein muss, Jubel oder Verzweiflung erklingen lassen, ohne den reinen Darstellungssinn des Wortes im mindesten zu tangieren.«<sup>7</sup>

Die Bühlerschen Überlegungen wurden von Werner Meyer-Eppler<sup>8</sup> aufgegriffen und aus dem Bereich der sprachlichen Kommunikation in den Bereich der musikalischen Kommunikation übertragen. Nach Meyer-Eppler sind die akustischen Signale, die sowohl eine sprachliche als auch eine musikalische »Kommunikationskette« zwischen Expedienten, den Sendern, und Perzipienten durchlaufen, Träger unterschiedlicher Merkmale, die nicht nur sprachliche oder musikalische, sondern auch andere – etwa »diagnostische« – Merkmale mit einschließen.<sup>9</sup> Meyer-Eppler führt daher den Begriff »Kommunikationssphäre« ein, mit dem er darauf verweist, dass die akustischen Signale »mehrseitige kommunikative Dienste« simultan leisten.<sup>10</sup> Im wesentlichen unterscheidet er zwei Sphären: Die eine ist die »semantische Sphäre«, die nach Meyer-Eppler »die Träger der Zeichenfunktion (Zeichenkörper)« bezeichnet, deren Merkmale wesentlich »diskontinuierlich« sind und der Darstellung der Dinge oder Sachverhalte dienen. Die andere Kommunikationssphäre, die »ektosemantische Sphäre«, zeichnet sich dadurch aus, dass die »kontinuierlichen« Merkmale ihrer akustischen Signale entweder der Identifikation des Senders dienen oder auf Gefühle der Kommunikationspartner verweisen.<sup>11</sup> Entsprechend unterteilt Meyer-Eppler die ektosemantische Sphäre in eine »identifikatorisch-diagnostische« und eine »emotionale Sphäre«.<sup>12</sup>

Der vorliegende Beitrag ist der Versuch, die vorprädikative Ausdrucksfunktion der ektosemantischen Sphäre von Stimme und Musik – insbesondere in Relation zur singenden Person – zu verdeutlichen. In der Singstimme scheint die ektosemantische Sphäre für den Zeichenprozess der Musik aus folgenden Gründen eine entscheidende Rolle zu spielen: Erstens weist Musik nur eine untergeordnete Darstellungsfunktion auf. Zweitens besteht die Musiksemiose hauptsächlich in dem performativen, verkörpernden Akt. Der performative Zeichenprozess der Musik ist einerseits als ein Prozess der Interpretation der notierten Musik, z. B. einer Partitur, anzusehen, andererseits stellt die performative Inszenierung selbst einen Zeichenprozess dar, der mit dem Ausdrucksvermögen des Performers, der an der Performanz leiblich teilhat und damit sich selbst zeigt, eng gekoppelt ist. Während der erste Zeichenprozess sich auf die Verkörperung der komponierten bzw. notierten Musik durch die Singstimme bezieht, bezieht sich der zweite Zei-

chenprozess auf den Ausdruck des Sängers bzw. dessen Innerlichkeit durch seine leibliche Präsenz sowie Inszenierung. Im Horizont der Performativität der Musik stellen die vom Performer produzierten akustischen Phänomene, die nicht notwendigerweise in der musikalischen Partitur kodiert sein müssen, aber real in der Aufführung vorhanden sind, einen Hauptuntersuchungsgegenstand des musikalischen Zeichenprozesses dar. Drittens ist der Ausdruck innerer Einstellungen und emotionaler Zustände durch die Singstimme nicht bloß ein unmittelbarer, unkontrolliert erzeugter Überschuss an Bedeutung, sondern immer auch ein vom Sänger durch elaborierte und artifizielle Steuerung der Stimme inszeniertes Produkt. Die Ausdrucksfunktion der Singstimme ist zwar – wie zu Beginn für die Stimme erwähnt wurde – vorprädikativ und prädiskursiv, weist aber im Vergleich zur ektosemantischen Sphäre der Sprechstimme einen intendierten Zeichencharakter auf. Insofern die Singstimme in ihrer expressiven Dimension somit aufgrund ihrer intendierten Erzeugung als Artefakt gelten kann, stellt sich die Frage, inwieweit sich in Bezug auf die Ausdrucksfunktion der Singstimme Regelmäßigkeiten feststellen lassen.

Im folgenden soll diese Fragestellung anhand von psychologischen und neurowissenschaftlichen Forschungen zur Sprech- und Singstimme, beschränkt auf die emotionale Sphäre der Singstimme, diskutiert werden. Im Vordergrund wird dabei die Frage stehen, inwieweit und in welchen akustischen Merkmalen der Singstimme sich Emotionen manifestieren, bzw. inwieweit sich innere emotionale Zustände des Sängers durch die Singstimme ausdrücken und erschließen lassen.

## 2. ERFORSCHUNG DER EMOTIONALEN BEREICHE VON SPRECH- UND SINGSTIMME

Psychologische und physiologische Forschungen über die Sprech- und Singstimme belegen, dass je verschiedene innere emotionale Zustände zu je spezifischen Veränderungen in der Atmung, der Stimmgebung und der Artikulation führen, die sich in den akustischen Parametern der Stimme zeigen.<sup>13</sup> So beeinflussen emotionale Zustände etwa das Atmungsmuster, das wiederum Einflüsse auf den Druck der Stimmritze hat (siehe Abbildung 1). Die Erhöhung dieses Drucks führt zu einer Steigerung der Lautstärke und in gewissem Umfang auch zu einer Steigerung der Phonationsfrequenz.<sup>14</sup>

Frühere Untersuchungen zur Sprechstimme belegen, dass der emotionale Zustand bzw. die Haltung eines Sprechers Auswirkungen auf die Phonation und die Artikulation haben. Ivan Fónagy beispielsweise wies in seinen Studien zum

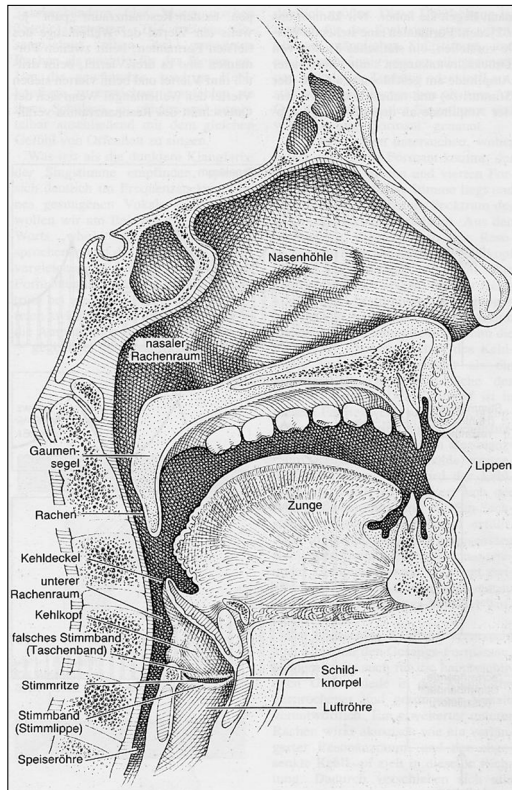


Abb. 1  
Stimmorgan

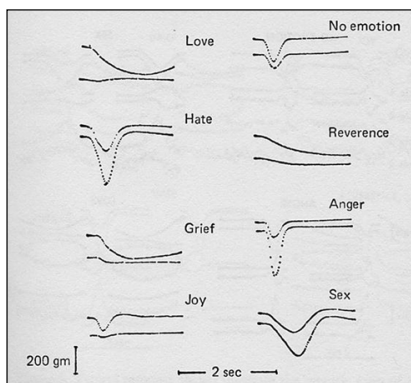
Verhalten der Glottis beim gefühlsbetonten Sprechen mit den Methoden der Kehlkopfspiegelung und der Röntgentomographie Einflüsse der emotionalen Zustände auf die Kehlkopfstellungen nach.<sup>15</sup> Diese stimmungsabhängigen Glottiseinstellungen bezeichnet er als »vorbewusst existente Ausdruckshaltungen«.<sup>16</sup> Ähnliche Ergebnisse erzielte Fónagy auch bezüglich der Auswirkungen der emotionalen Haltung auf die Artikulation. In seinen Untersuchungen konnte er Beziehungen zwischen verschiedenen Mustern der Zungenlage und der Mundöffnung beim Sprechen und der emotionalen Grundhaltung aufzeigen. Demnach zeichne sich etwa »Wut«, so Fónagy, durch rasche Bewegungen zwischen extremen Artikulationspositionen aus, während langsame und weichere Bewegungen »Zärtlichkeit« kennzeichnen sollen.<sup>17</sup>

Johan Sundberg sieht eine mögliche Erklärung für derartige Vorgänge in Phonation und Artikulation beim Ausdruck von Emotionen darin, »dass jede Emotion und jede Haltung ihre eigenen typischen Bewegungsmuster haben, welche das Verhalten des gesamten Körpers (einschließlich des Stimmorgans) beeinflussen«.<sup>18</sup> Als Beispiele führt er »Traurigkeit« und »depressive Haltung« an, bei

denen ein Großteil der muskulären Aktivität minimiert wird. Solch eine Interdependenz zwischen Bewegung und Emotion spiegelt sich bereits im Wortstamm der Emotion: Das lateinische Wort »emotio« bedeutet heftige Bewegungen.<sup>19</sup>

Auch neurowissenschaftliche Forschungen stützen die Annahme eines Zusammenhangs zwischen Emotion und Bewegungsmuster. Verschiedene emotionale Zustände weisen laut bisheriger neurowissenschaftlicher Forschungen einen Zusammenhang mit der Erregung des autonomen Nervensystems auf: Ärger und Freude stehen in Relation zur Aktivität des sympathischen Nervensystems, während Langeweile und Trauer mit der Aktivität des parasympathischen Nervensystems zusammenhängen.<sup>20</sup> Nach der neurobiologischen Emotionstheorie von Manfred Clynes sind phonatorische und artikulatorische Gebärden Manifestationen der von ihm sogenannten »sentic forms«, die wiederum zeitliche Muster für neuromuskuläre Aktivitäten in Bezug auf verschiedene Emotionen repräsentieren.<sup>21</sup> Diese »sentic forms« sind durch ihre räumlich-zeitliche Struktur gekennzeichnet. Die folgende Abbildung zeigt verschiedene »sentic forms«, die durch den von Clynes entwickelten »sentograph« beim Fingerdruck gemessen wurden.

**Abb. 2**  
»sentic forms«



In der Stimme manifestieren sich diese »sentic forms« generell in den akustischen Merkmalen von Frequenz, Dauer und Amplitude. Diese Merkmale, die beim Sprechen als Tonhöhenverlauf, Dauer und Lautstärke wahrgenommen werden, sind als Mittel zum Ausdruck unterschiedlicher Emotionen in der Sprechstimme anzusehen. Dieser Befund wird auch durch Forschungsergebnisse aus der Phonetik und Experimentalpsychologie bestätigt. Insbesondere in Bezug auf eine Differenzierung verschiedener Erregungen – wie beispielsweise »verärgert« vs. »fasziniert« – besteht Einigkeit darüber, dass sich diese in Änderungen der Tonhöhe, der Intensität und des Tempos in der Intonation manifestieren.<sup>22</sup>

Insoweit lässt sich der emotionale Ausdruck der Sprechstimme schon bei den früheren Forschungen über die Sprechstimme durch die Untersuchung der akustischen Merkmale der Sprechstimme nachweisen. Im Gegensatz dazu sind die akustischen Mittel zum emotionalen Ausdruck in der Singstimme nicht eingehender untersucht worden. Bei vielen früheren psychologischen bzw. musikpsychologischen Forschungen über die Singstimme handelt es sich hauptsächlich um Befragungen zum subjektiven Eindruck des Hörers.<sup>23</sup> Erst in einigen neueren psychologischen und akustischen Forschungen wird die Frage nach den formalen akustischen Merkmalen der Singstimme hinsichtlich ihres emotionalen Ausdrucks mit einbezogen.<sup>24</sup> Die Hauptuntersuchungsperspektive dieser neueren Forschungen konzentriert sich auf die Aufdeckung der formalen akustischen Mittel der Singstimme, welche vom Sänger zur Enkodierung der Emotionen eingesetzt werden.

Untersuchungen zur Emotionskodierung der Singstimme konzentrierten sich zunächst auf die vom Komponisten im Hinblick auf den emotionalen Ausdruck festgelegten musikalischen Merkmale, die in Form von Handlungsanweisungen in der Partitur vorliegen, da die Singstimme deren interpretativer Realisation dient. Das Ergebnis der Untersuchungen, in denen musikalische Merkmale bezüglich ihres emotionalen Ausdrucks erforscht wurden, weist darauf hin, dass Sprech- und Singstimme eine Gemeinsamkeit aufweisen, welche in der musikkonstanten Struktur besteht. Die Merkmale der Sprechstimme wie Tonhöhenverlauf, Intensität und Lautstärke, denen in der Sprache emotive Bedeutungen zukommen, finden sich beim Singen bereits in der vom Komponisten notierten Musik in der Partitur. In der Musik sind somit die sogenannten »sentic forms« nicht nur auf der Aufführungsebene, sondern auch auf der Kompositionsebene, bzw. in der Partitur zu analysieren. Insbesondere der vom Komponisten festgelegte Verlauf der Tonhöhenkurve wird von einigen Forschern wie Johan Sundberg und Eliezer Rapoport als eine Art »Makrointonation« bezeichnet.<sup>25</sup> Beim Singen kann der Sänger diese vorgegebene Makrointonation während der Aufführung durch die Änderung der Lautstärke, der Tondauer, des Tempos und der Konnektion von Staccato und Legato variieren und damit den emotionalen Ausdruck des Musikstückes erweitern.

Ein interessantes Untersuchungsfeld der Singstimme stellt darüber hinaus aber auch die sogenannte »Mikrointonation« dar. Die Singstimme ermöglicht aufgrund der flexiblen Änderungsmöglichkeiten des Resonators eine zusätzliche Gestaltung der zeitlichen Struktur von einzelnen Tönen, welche durch den Wechsel der Tonhöhe innerhalb eines einzelnen Tons erfolgt. Eine solche Intonation auf der Mikroskala innerhalb eines individuellen Tons wird Mikrointonation

genannt. Beim Singen kommt dem Sänger also über die Bindung an die vom Komponisten festgelegte Makrointonation hinaus eine weitere Möglichkeit zum emotionalen Ausdruck zu: Der Sänger kann die Emotionen auch auf der Mikroebene innerhalb eines individuellen Tons kodieren.<sup>26</sup>

Die Emotionskodierung der Singstimme durch Mikrointonation tritt hauptsächlich während der musikalischen Performanz auf. Die Mikrointonation der Singstimme ist einerseits als ein unmittelbares Produkt des physischen Leibes, welches innere Zustände des Sängers übermitteln kann, anzusehen. Andererseits ist die Mikrointonation der Singstimme in der künstlerisch-musikalischen Performanz ein höchst inszenierter Ausdruck der Emotionen und wird somit zu einem symbolischen Phänomen. So erweitert die Untersuchung emotionalen Ausdrucks am Beispiel der Mikrointonation in der Singstimme das Spektrum der Erforschung musikalischen Ausdrucks, der einen durch die körperliche Performanz zu konstituierenden Sinn des musikalischen Zeichenprozesses darstellt.

### 3. ANALYSE DER MIKROINTONATION IN DER SINGSTIMME MITHILFE DES FFT-SONOGRAMMS

Der in der Mikrointonation eines einzelnen Tons durch die Singstimme hervorbrachte emotionale Ausdruck, der sich in den kleinsten Details der Frequenzverlaufskurve niederschlägt, kann in den akustischen Signalen entziffert werden. Da die Mikrointonation jedoch in den traditionellen Notationssystemen wie z. B. in einer musikalischen Partitur nicht notierbar ist, bedarf die Mikroanalyse der fluiden auditiven Schallphänomene eines neuen Aufzeichnungsmediums.

Der musikalische Ton, der die Grundeinheit der Musik darstellt, ist durch eine charakteristische zeitliche Struktur gekennzeichnet. Beim akustischen Signal eines musikalischen Tones bilden im Allgemeinen drei Segmente jeweils einen bestimmten Zeitabschnitt des Signals, dem charakteristische Signaleigenschaften zugeordnet werden: »Einschwingvorgang«, »quasistationäre Phase« und »Ausschwing- oder Übergangsvorgang«. Der Einschwingvorgang bezeichnet die »instationäre Phase« des Klangeinsatzes, welche dynamische Eigenschaften wie plötzliche Zeithüllenänderungen und rasche Grundfrequenzänderungen kennzeichnet. Daran schließt sich die »quasistationäre Phase« an, die sich durch zeitliche Invarianz von Kurzzeitspektrum, Grundfrequenz und Intensität auszeichnet. Dabei tragen überlagerte Amplituden- und Frequenzfluktuationen (Tremolo und Vibrato) in der Regel zur Belebung des statischen Klangeindrucks bei. Die letzte Phase eines musikalischen Tons bildet die Übergangsphase zum nächsten Ton, die durch graduelle Zeithüllenänderungen sowie kontinuierliche Grundfre-



quenzänderungen gekennzeichnet ist. Diese Übergangsphase zum nächsten Ton wird Ausschwing- oder Übergangsvorgang genannt.<sup>27</sup>

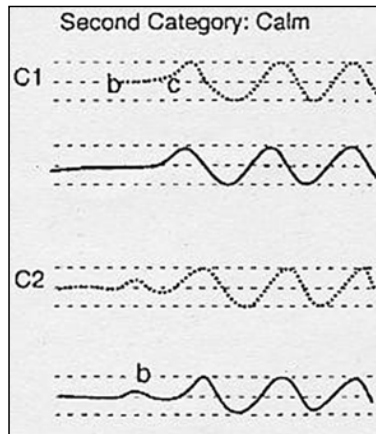
Für die Mikroanalyse der Singstimme, die innerhalb eines einzelnen musikalischen Tons durchgeführt werden soll, scheint ein Sonogramm mit der FFT Analyse (Fast Fourier Transform-Analyse) geeignet zu sein: Das FFT-Sonogramm ist ein Frequenz-Zeit Diagramm, das durch die Dekomposition eines Signals in dessen Frequenzkomponenten alle spektralen Informationen, d. h. alle Frequenzkomponenten des Signals und die ihnen entsprechenden Intensitäten in einer zeitlichen Einheit zeigt.

Einen bedeutenden Ansatz zur Untersuchung der Emotionskodierung innerhalb eines individuellen vokalen Tons liefert Eliezer Rapoport.<sup>28</sup> Für die FFT-Analyse der Mikrointonation der Singstimme entwickelte er verschiedene Kategorien der zeitlichen Struktur des vokalen Tons, welche aus der Kombination unterschiedlicher Merkmale der akustischen Signale bestehen. Diese Kategorien und Mechanismen der akustischen Signale eines einzelnen Tons, welche sich durch die FFT-Analyse nachweisen lassen, bilden die Analysemittel des FFT-Sonogramms für die Kodierung verschiedener Emotionen der Singstimme. Im folgenden werden zunächst die relevanten »Mechanismen« und Kategorien von Rapoport vorgestellt. Im Anschluss daran werden seine Forschungsergebnisse anhand eigener FFT-Analysen verschiedener Singstimmen nachvollzogen.

Rapoport zufolge wird die zeitliche Struktur des vokalen Tons insbesondere beim Toneinsatz durch ein Zusammenspiel der folgenden »Mechanismen« kontrolliert:<sup>29</sup> 1. Einsatz der Phonation; 2. Vibrato; 3. Verstärkung von höheren Partialen bzw. des Sängerformanten; 4. Übergang – eine graduelle Zunahme der Tonhöhe vom Toneinsatz bis zur quasistationären Phase; 5. Sforzando – eine plötzliche Zunahme der Tonhöhe beim Einschwingvorgang der Töne.<sup>30</sup> Die unterschiedlichen Typen der zeitlichen Struktur des vokalen Tons – die von ihm sogenannten »modes«<sup>31</sup> –, denen jeweils verschiedene emotionale Bedeutungen zugeschrieben werden, zeichnen sich durch die regelhafte Kombination dieser Mechanismen aus.

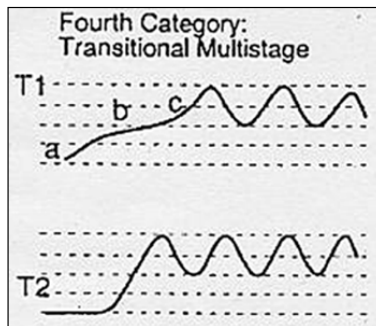
Rapoports kategorielle Differenzierung unterscheidet folgende »modes«, die auch in der von mir durchgeführten FFT-Analyse nachgewiesen werden konnten: 1. »Neutral-Soft«; 2. »Calm«; 3. »Expressive«; 4. »Transitional-Multistage«; 5. »Excited«.<sup>32</sup> Die erste Kategorie »Neutral-Soft« besitzt keine besondere Struktur; sie ist durch die Abwesenheit von Vibrato gekennzeichnet. Bei der zweiten Kategorie »Calm«, die beim Ausdruck der Ruhe, Traurigkeit und gedämpften Stimmung auftritt, wird mithilfe der drei Mechanismen »Phonation – Verstärkung von höheren Partialen – Vibrato« oder »Phonation – Vibrato – Verstärkung von höheren Partialen« graduell operiert.

Abb. 3

Kategorie  
»Calm«

Bei der dritten Kategorie »Expressive« dominiert das Vibrato. Die vierte Kategorie »Transitional-Multistage« zeigt sich beim Ausdruck von großer Faszination oder Aufregung. Sie enthält eine graduelle oder plötzliche Zunahme der Tonhöhe. Die typische Kombination dieser Kategorie besteht aus den Mechanismen »Phonation – Verstärkung von höheren Partialen – Übergang und wahlweise Vibrato«.

Abb. 4

Kategorie  
»Transitional-  
Multistage«

Die fünfte Kategorie »Excited«, die beim Ausdruck von Freude oder Spannung erkennbar ist, ist durch die Kombination der vier Mechanismen »Phonation – Sforzando – Verstärkung von höheren Partialen – Vibrato« gekennzeichnet.

Rapoport bestätigt durch seine eigene FFT-Analyse von Opernarien und Kunstliedern die Regelmäßigkeit der »modes«-Kategorien in Verbindung mit der Emotionskodierung der Singstimme. Um seine Untersuchungsergebnisse zu überprüfen, wurde von mir ein Experiment mit der FFT-Analyse der Singstimme durchgeführt. Die zur Analyse ausgewählten Gesangsstücke umfassen Opernarien, zeitgenössische Vokalwerke und asiatische bzw. koreanische Volkslieder. Als Quelle wurden die auf einer CD (compact disc) aufgenommenen Gesangsstü-

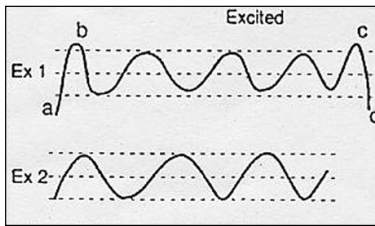


Abb. 5

Kategorie »Excited«

cke benutzt. Die ausgewählten Abschnitte jedes Stückes wurden auf die Festplatte eines Macintosh-Computers für die FFT-Analyse überspielt. Zur Analyse wurde das vom *Institut de Recherche et Coordination Acoustique/Musique* (IR-CAM) entwickelte Softwareprogramm »AudioSculpt« eingesetzt: Schrittgröße 512, Blackman-Fenster, ein Rahmen mit einer Länge von 4096 Abtastpunkten und die Grundfrequenz 53,8.

Bei der FFT-Analyse der Singstimme ließ sich in den Opernarien, in denen es um den Ausdruck gedämpfter Stimmung – wie etwa Einsamkeit oder Traurigkeit – geht, die Kategorie »Calm« nachweisen. Die folgenden Abbildungen stellen die Analyseergebnisse dieser Kategorie dar: Abbildung 6 ist ein Analyseabschnitt der Opernarie »La schia chio pianga« aus Händels Oper »Rinaldo«, bei der die Frauenstimme tiefe Trauer auf Grund des Abschieds mit ihrem Geliebten ausdrückt. Die Analyse einer weiteren Opernarie »Einsam in trüben Tagen« aus der Wagner Oper »Lohengrin« ist ebenfalls durch die Kategorien »Calm« gekennzeichnet. Dabei zeigt sich die Kombination der Mechanismen »Phonation – Verstärkung von höheren Partialen – Vibrato« (Siehe Abbildung 7).

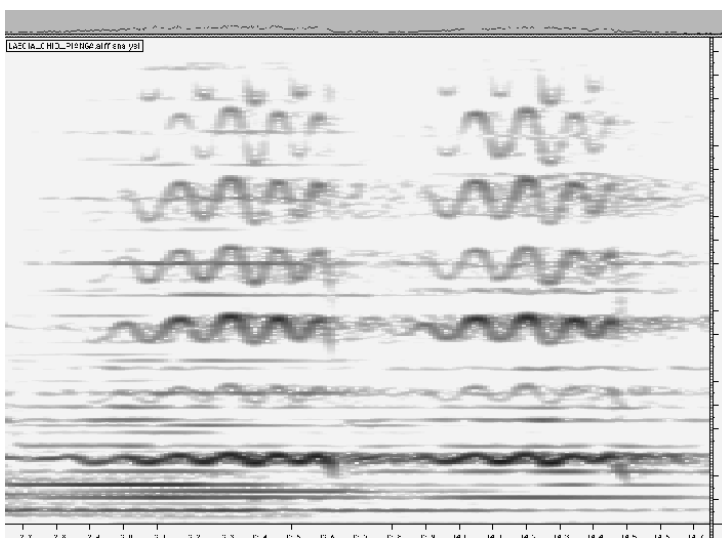


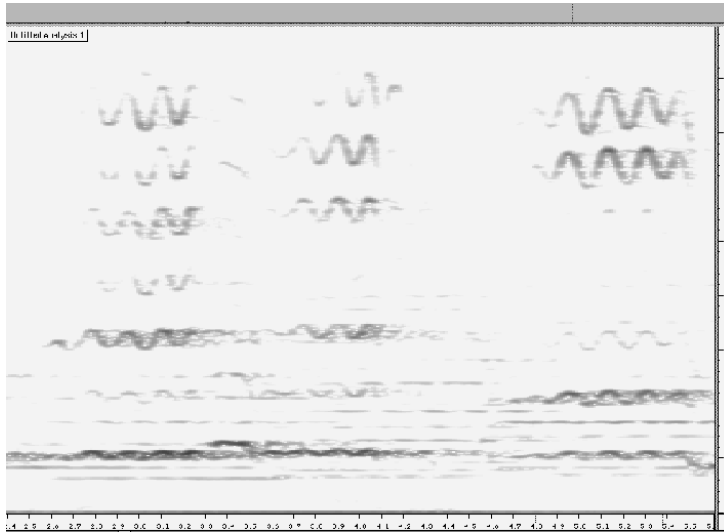
Abb. 6

»La schia chio pianga«

Jin Hyun Kim

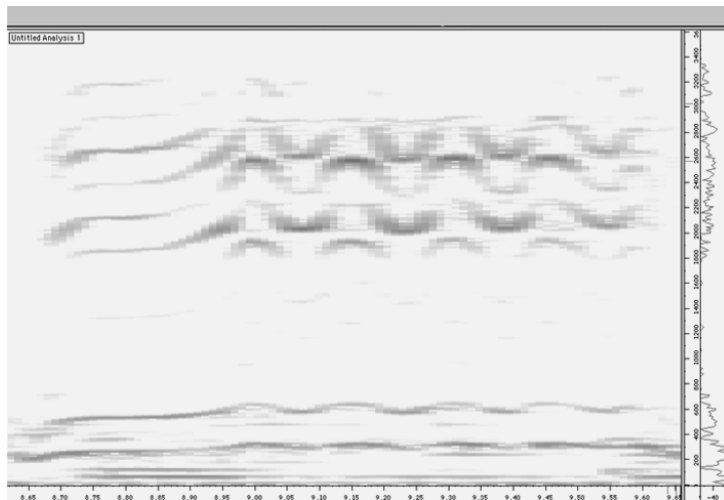
260

Abb. 7

»Einsam in trüben  
Tagen«

Die Kategorien »Transitional-Multistage« und »Excited« sind in der bekannten Opernarie »Un de felice, eterea« aus Verdis Oper »La Traviata« erkennbar, bei der es sich um große Freude auf Grund von Verliebtsein handelt. In dieser Opernarie zeigt sich die Kategorie »Transitional-Multistage« in der Männerstimme (Sänger: Tito Shipa) und die Kategorie »Excited« in der Frauenstimme (Sängerin: Emma Carelli). In der Männerstimme (Abbildung 8) findet man die Mechanismen in der Reihenfolge »Phonation – Verstärkung von höheren Partialen – Übergang (graduelle Zunahme der Tonhöhe) und wahlweise Vibrato«, in der Frau-

Abb. 8

»Un de felice,  
eterea« (Männer-  
stimme)

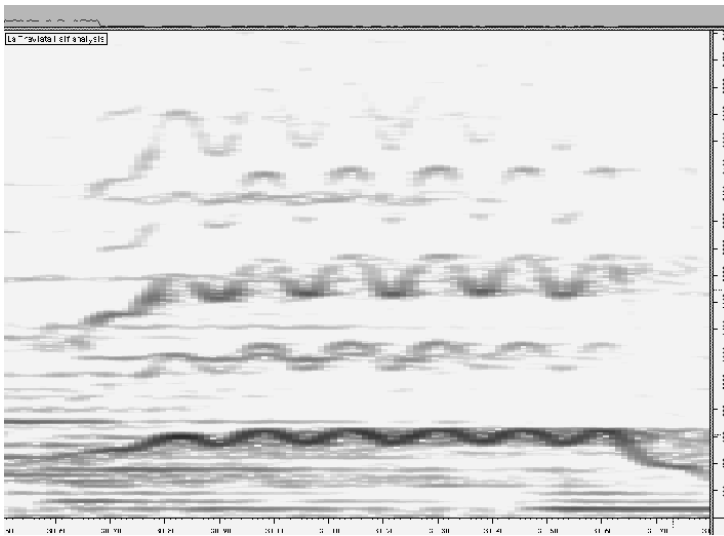


Abb. 9

»Un de felice, eterea«  
(Frauenstimme)

enstimme die Kombination der Mechanismen »Phonation – Sforzando (plötzliche Zunahme der Tonhöhe) – Verstärkung von höheren Partialen – Vibrato« (Siehe Abbildung 9).

Diese Kategorien lassen sich auch in der Singstimme mit asiatischer Singtechnik nachweisen. Bei einem koreanischen Volkslied, das große Freude aufgrund der reichlichen Ernte zum Ausdruck bringt, wird die Kategorie »Excited« mit der Kombination der Mechanismen »Phonation – Sforzando – Verstärkung von höheren Partialen – Vibrato« beobachtet (Abbildung 10).

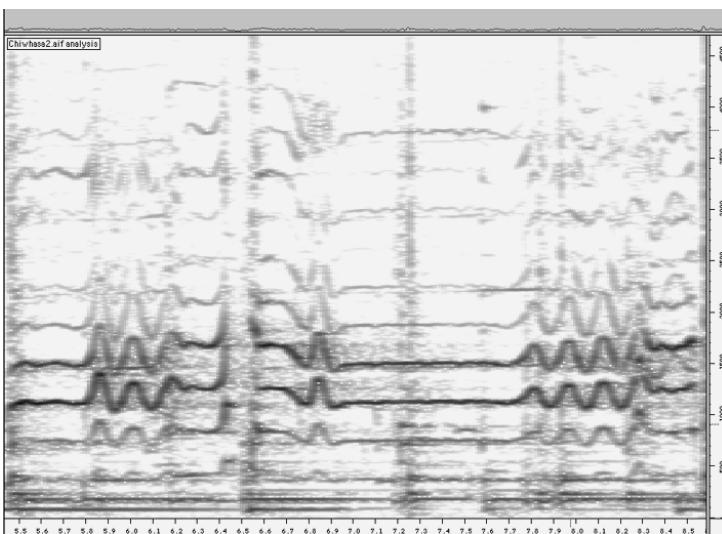


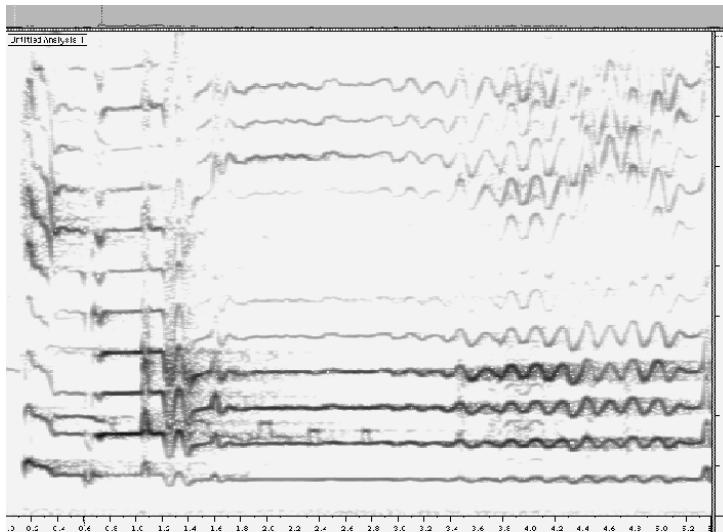
Abb. 10

»Chihwasan«

Jin Hyun Kim

262

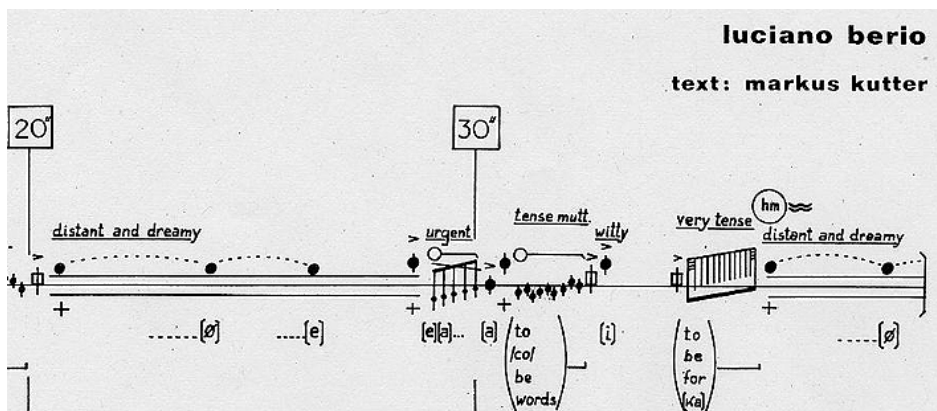
Abb. 11  
»Ibyulga«



Bei einem weiteren koreanischen traditionellen Lied mit dem Titel »Ibyulga [Abschied]«, bei dem großer Schmerz beim Abschied einer Frau von ihrem Geliebten zum Ausdruck gebracht wird, tritt die fünfte Kategorie »Excited« häufig auf. Abbildung 11 ist ein Abschnitt der Analyse dieses Stückes, wobei die Kategorien »Excited« und »Expressive« nacheinander auftreten.

Abb. 12

Notenbeispiel »sequenza III«



Als letzte FFT-Analyse wurde Luciano Berios Vokalwerk für eine Frauenstimme »sequenza III« (1966) ausgewählt, welches insbesondere mit der Möglichkeit der Stimme zum emotionalen Ausdruck experimentiert. Berio gibt in

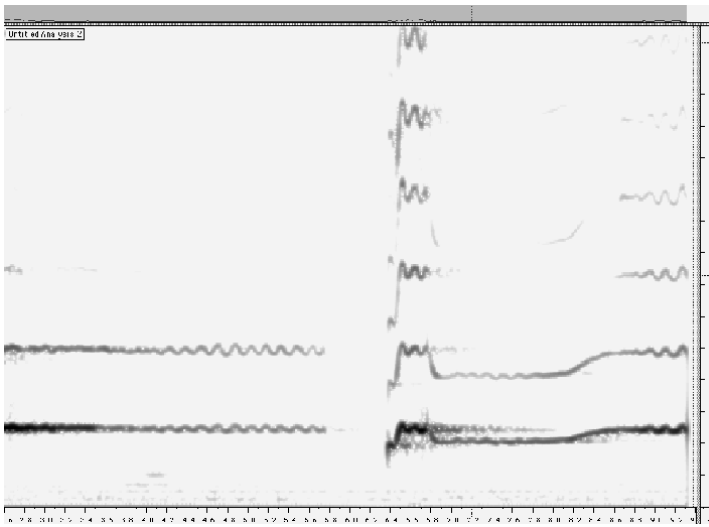


Abb. 13/1

Sonogramm  
»sequenza III«

diesem Stück genaue Anweisungen für die einzusetzende Stimmtechnik der Sprech- und der Singstimme wie etwa verschiedene Geschwindigkeitsgrade periodisch artikulierter Töne sowie die Bewegung der Stimme im Register. Des Weiteren gibt Berio eine verbalisierte Anweisung zum Gefühlsausdruck für jede kleine Einheit des Stückes wie z. B. »tender«, »tense«, »dreamy«, »joyful« usw. Teilweise verlangt er von der Sängerin einen sprunghaften Wechsel des emotionalen Ausdrucks innerhalb von einigen nacheinander folgenden Tönen, wie in der folgenden Partitur zu erkennen ist.

Die FFT-Analyse dieses Stückes hatte das Ziel, die Übereinstimmung der verbalen Anweisungen des Komponisten mit den im Sonogramm nachzuweisen den Kategorien der »modes« in den akustischen Signalen zu überprüfen. Die folgende Abbildung zeigt das Ergebnis der FFT-Analyse eines Abschnittes dieses Stückes, bei dem vier Töne nacheinander unterschiedliche Kategorien aufweisen:

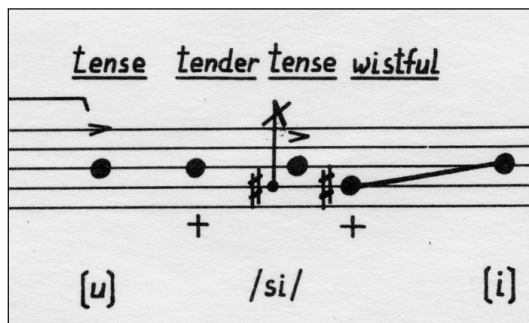


Abb. 13/2

Partitur »sequenza III«

Jin Hyun Kim

264

»Neutral-Soft«, »Expressive«, »Excited« und »Calm«. (Abbildung 13–1). In der Partitur stehen dafür Ausdrucksanweisungen in der Reihenfolge »tense – tender – tense – wistful«.

#### 4. RESÜMEE

Gegenstand des hier vorliegenden Beitrages war die Diskussion der akustischen Merkmale der Singstimme in Bezug auf ihre ektosemantische bzw. emotionale Sphäre. Im Zentrum dieser Überlegungen stand die Annahme, dass den akustischen Merkmalen der Singstimme in der ektosemantischen Sphäre eine Ausdrucksfunktion zukommt. Dieses vorprädikative und prädiskursive Moment der Stimme, in dem die (emotionale) Innerlichkeit des Senders zum Vorschein kommt, kann im Hinblick auf seine semiotische Funktion im Anschluss an Bühler »Ausdruckszeichen« genannt werden.<sup>33</sup> Das Ausdruckszeichen stellt in den bisherigen linguistisch-philosophischen Forschungen zur Sprache als einem Instrument der Symbolisierung nur ein wenig untersuchtes Phänomen dar, da bevorzugt die Gegenstandsrelation in den Blick genommen wurde. Vor dem an der Darstellungsfunktion der Sprache orientierten Hintergrund wurde häufig sogar von einer Defizienz der musikalischen Kommunikation gesprochen, ebenso wie die Medialität der Stimme wenig thematisch wurde. Im gegenwärtigen medientheoretischen Diskurs zum vorprädikativen Phänomen der Stimme, der vorwiegend auf die Sprechstimme Bezug nimmt, gilt das sprachliche Ausdruckszeichen als unmittelbar erzeugtes Phänomen, das vom Sprecher nicht intendiert und demnach nicht seiner Kontrolle unterworfen ist. Im Gegensatz dazu gewinnt das Ausdruckszeichen in der musikalischen Kommunikation einen zentralen Stellenwert, weil in der musikalischen Performanz die Ausdrucksebene der akustischen Signale nicht lediglich eine subordinierte Funktion der Darstellungsfunktion bildet, sondern zur dominierenden medialen Funktion aufsteigt. Gerade beim performativen Akt des Singens ist die Singstimme nicht nur ein Phänomen des physischen Leibes, sondern auch ein inszenierter symbolischer Zeichenausdruck. Das von der Singstimme hervorgebrachte Ausdruckszeichen, welches innere emotionale Zustände des Sängers zugleich unmittelbar und intendiert zu übermitteln vermag, ist kein nachgeordnetes Moment der musikalischen Kommunikation, sondern zentraler Gegenstand des performativen musikalischen Aktes. Die Singstimme, die sich durch die artifizielle Inszenierung von Persönlichkeit auszeichnet, ist nichts anderes als das Ausdrucksmedium, durch das der Sänger intentional kommuniziert. In der Performanz der Singstimme konstituieren sich



also nicht-sprachliche Ausdruckszeichen, die zwar nichts darstellen, aber gleichwohl etwas zeigen. Dieses ›Zeigen‹ der Singstimme ist aber – auch wenn es sich in einem nicht-diskursiven, vorprädikativen Modus entfaltet – nicht nur ein unwillkürliches Zum-Ausdruck-Kommen des emotionalen Innenlebens, sondern vor allem ein zeichenhafter Akt, der sich unter der Kontrolle des Sängers vollzieht. Es wurde deshalb in dem Beitrag die These vertreten, dass die Singstimme, obwohl sie auf die emotionale Sphäre beschränkt bleibt, dennoch in ihrer Performanz als regelgeleitet angesehen werden muss, wobei sich diese Regelmäßigkeit durch die Analyse ihrer akustischen Merkmale nachweisen lässt. Die Untersuchung der emotionalen Sphäre der Singstimme durch eine FFT-basierte Spektralanalyse ihrer in computergenerierten Sonogrammen dargestellten akustischen Merkmale zeigt, dass die Mikrointonation der Singstimme, die in der Partitur nicht analysiert werden kann, als intendiertes Ausdrucksmittel von Emotionen dient. Die Mikroanalyse einzelner Töne ergibt eine strukturelle Regelmäßigkeit der Singstimme, die sich aus der je variierenden Kombination musikalisch-akustischer Merkmale wie beispielsweise Toneinsatz, Vibrato, Verstärkung des Sängersformanten und Sforzando ableiten lässt. Diese Untersuchungsergebnisse können also die Annahme stützen, dass die Singstimme dem Sänger in der musikalischen Kommunikationssituation erlaubt, Ausdruckszeichen zu generieren, die Emotionen in einer intendierten, regelhaften Weise kodieren.

- 1 Vgl. hierzu etwa Christian Stetter: *Schrift und Sprache*, Frankfurt 1997.
- 2 Sybille Krämer: *Das Medium als Spur und als Apparat*, in: dies. (Hg.): *Medien – Computer – Realität. Wirklichkeitsvorstellungen und Neue Medien* [1998], Frankfurt/M. 2000, S. 73–94 (hier: S. 79).
- 3 Ebd., S. 90.
- 4 Karl Bühler: *Sprachtheorie. Die Darstellungsfunktion der Sprache* [1934], Stuttgart 1999, S. 28.
- 5 Ebd., S. 30.
- 6 Ebd.
- 7 Ebd., S. 46.
- 8 Werner Meyer-Eppler: *Grundlagen und Anwendungen der Informationstheorie* [1959], Berlin/Heidelberg/New York 1969.
- 9 Ebd., S. 3.
- 10 Ebd.
- 11 Ebd.
- 12 Werner Meyer-Eppler: *Informationstheoretische Probleme der musikalischen Kommunikation*, in: *Revue Belge de Musicologie* 13/1959, S. 44–49 (hier: S. 45).
- 13 Johan Sundberg: *Die Singstimme*, in: Klaus Winkler: *Die Physik der Musikinstrumente*, Heidelberg 1992, S. 14–21; Klaus R. Scherer: *Expression of Emotion in Voice and Music*, in: *Journal of Voice* 9/3 (1995), S. 235–248; Johan Sundberg: *Die Wissenschaft von der Singstimme*, Bonn 1997; Marcel Zentner/Klaus R. Scherer: *Emotionaler Ausdruck in Musik und Sprache*, in: Klaus-Ernst Behne/Günter Kleinen/Helga de la Motte-Haber (Hg.): *Musikpsychologie. Jahrbuch der Deutschen Gesellschaft für Musikpsychologie*, Bd. 13: *Musikalischer Ausdruck*, Göttingen 1998, S. 8–25.
- 14 Sundberg: *Die Wissenschaft von der Singstimme* (Anm. 13), S. 200.
- 15 Ivan Fónagy: *Mimik auf glottaler Ebene*, in: *Phonetica* 8 (1962), S. 209–219.
- 16 Ebd., S. 213.

Jin Hyun Kim

266

- 17 Ivan Fónagy: La mimique buccale, in: *Phonetica* 33 (1976), S. 31–44; Sundberg: Die Wissenschaft von der Singstimme (Anm. 13), S. 209.
- 18 Sundberg: Die Wissenschaft von der Singstimme (Anm. 13), S. 211.
- 19 Lexikon der Neurowissenschaft, 4 Bde., Heidelberg 2000, Bd. 1, 2000, Stichwort: Emotionen, S. 386–397 (hier: S. 386).
- 20 Josef Dudel/Randolf Menzel/Robert F. Schmidt (Hg.): Neurowissenschaft. Vom Molekül zur Kognition [1996], Berlin u. a. 2001, S. 478–486; Nils Birbaumer/Robert Schmidt: Biologische Psychologie [1991], Berlin/Heidelberg/New York 1999, S. 624–674.
- 21 Manfred Clynes: *Sentics. The Touch of Emotions*, London 1977.
- 22 Ivan Fónagy/Klara Magdics: Emotional Patterns in Intonation and Music, in: Karl Ammer/Otto von Essen/Fritz Hintze (Hg.): *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 16/4 (1963), S. 293–326; Johan Sundberg: Production and function of the »singing formant«, in: Henrik Glahn/Sören Sörensen/Peter Ryom (Hg.): *International Musicological Society. Report of the eleventh congress Copenhagen 1972*, Bd. II, Copenhagen 1974, S. 679–686; Ian R. Murray/John L. Arnott: Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion, in: *The Journal of the Acoustic Society of America* 93/2 (1993), S. 1097–1108; Bernd Tischer: Die vokale Kommunikation von Gefühlen, Weinheim 1993; Hervine Siegwart/Klaus R. Scherer: Acoustic Concomitants of Emotional Expression in Operatic Singing: The Case of Lucia in *Ardi gli incensi*, in: *Journal of Voice* 9/3 (1995), S. 249–260; Gordon Troup/Sinisa Djordjevic: The Effect of Emotions on the Spectra of the Singing Voice, in: *2nd International Conference on Acoustics and Musical Research (CIARM 95)*, Ferrara, Italy, S. 403–408.
- 23 Helmut Rösing: Musikalische Ausdrucksmodelle, in: ders./Herbert Bruhn/Rolf Oerter (Hg.): *Musikpsychologie. Ein Handbuch in Schlüsselbegriffen*, München/Wien/Baltimore 1985, S. 174–179; Reinhard Pekrun: Musik und Emotion, in: Herbert Bruhn/Rolf Oerter/Helmut Rösing (Hg.): *Musikpsychologie. Ein Handbuch*, Reinbek 1993, S. 180–188.
- 24 Genannt seien die aktuellen Forschungen von Patrik N. Juslin: A functionalistic perspective on emotional communication in music, in: *European Society for the Cognitive Sciences of Music Newsletter* 8 (1995), S. 11–16; Alf Gabrielsson/Erik Lindstrom: Emotional expression in synthesizer and sentograph performance, in: *Psychomusicology*, 14/1–2 (1995), S. 94–116; Patrik N. Juslin/John A. Sloboda (Hg.): *Music and Emotion. Theory and Research*, Oxford 2001.
- 25 Johan Sundberg: Die Wissenschaft von der Singstimme (Anm. 13); Eliezer Rapoport: Temporal Aspects of Timbre and Emotional Expression in Opera and Lied Singing. An FFT Study, in: *Systematische Musikwissenschaft* 4/1–2 (1996), S. 311–329; ders.: Emotional Expression Code in Opera and Lied Singing, in: *Journal of New Music Research* 25/2 (1996), S. 109–149.
- 26 Johan Sundberg: Die Wissenschaft von der Singstimme (Anm. 13), S. 199.
- 27 Wolfgang Fasold/Wolfgang Kraak/Werner Schirmer (Hg.): *Taschenbuch Akustik*, Bd.1, Berlin 1984, S. 236.
- 28 Rapoport: Temporal Aspects of Timbre and Emotional Expression in Opera and Lied Singing (Anm. 25); ders.: Emotional Expression Code in Opera and Lied Singing (Anm. 25).
- 29 Rapoport: Emotional Expression Code in Opera and Lied Singing (Anm. 25), S. 129.
- 30 Ebd., S. 129–133. Rapoport stellt dabei noch weitere Mechanismen wie Änderung der Tonhöhe innerhalb der Töne und Einheitspuls dar. Hier werden jedoch nur die in meiner eigenen Analyse bestätigten Mechanismen vorgestellt.
- 31 Ebd., S. 110; ders.: Temporal Aspects of Timbre and Emotional Expression in Opera and Lied Singing (Anm. 25), S. 314.
- 32 Ebd., S. 115–130; ders.: Temporal Aspects of Timbre and Emotional Expression in Opera and Lied Singing, S. 317–323.
- 33 Karl Bühler: *Sprachtheorie* (Anm. 4), S. 36.

Doris Kolesch

DIE SPUR DER STIMME.

ÜBERLEGUNGEN ZU EINER PERFORMATIVEN ÄSTHETIK

### 1. LAUTLOS

Die schriftliche Fassung meines Vortrages setzt ein mit einem zweifachen Verstummen oder Verklungen, das die mediale Differenz von Stimme und Schrift, von Sprechen und Schreiben ausstellt. Verklungen ist meine eigene Stimme, *in* und *mit* der ich *über* die Stimme und *von* ihr gesprochen habe, ohne dass beides jemals zur Deckung gekommen wäre. Verklungen ist aber auch die Reproduktion der Stimme von Marianne Hoppe, die Thomas Bernhards Erzählung *Der Stimmenimitator* liest. Mit der Wiedergabe dieser Aufnahme begann ich meinen Vortrag, um mich am Beispiel von Hoppes Stimme und der von ihr artikulierten Geschichte dem Phänomen Stimme zu nähern und auf dieser Grundlage über Stimme als Medium und über eine performative Ästhetik nachzudenken.

Zu sagen – und zu schreiben –, dass Marianne Hoppes Stimme inzwischen verklungen sei, impliziert eine signifikante Mehrdeutigkeit. Eine solche Aussage verweist auf den Ereignischarakter der Stimme, die singulär, einmalig und unwiederholbar ist. Die Stimme ist ein flüchtiges Phänomen, das im Hier und Jetzt existiert und dessen Existenz paradoxerweise im beständigen Verschwinden besteht, in einer Bewegung und Prozessualität, auf deren ambivalente materielle Gegebenheit ich noch eingehen werde.

Das Verklungen-Sein von Hoppes Stimme verweist neben der präsentischen, gegenwärtigen Dimension von Stimme aber auch auf deren existentielle Verfasstheit. Die Stimme gilt als Symbol und zugleich als Index, als Spur des Lebendig-Seins – ich erinnere nur an den Schrei des Neugeborenen, jene allererste akustische Lebensäußerung. Insofern die Stimme Ausweis des Lebendigen, des belebten Körpers ist, hat sie auch Teil an seiner Vergänglichkeit. Mit dem Körper hat die Stimme das Schicksal der Sterblichkeit gemein, das sie in ganz anderer Weise prägt als beispielsweise das technische Veralten oder Überholtwerden von Medien wie Grammophon, Schallplatte oder Tonband. Die Stimme ist die Spur des Körpers in der Sprache, und sie steht für dessen Identität: Wir erwarten am Telefon, dass eine Freundin, ein Freund uns allein an unserer Stimme erkennt. Die Stimme verändert sich mit dem Körper und verändert zugleich diesen Körper. So gehört es zu unserem Alltagswissen, dass Kinder andere Stimmen haben als Erwachsene und dass die männliche Stimme in der Pubertät »bricht«, auch

wenn wir uns signifikanterweise kaum mehr an unsere eigene Stimme als Kind erinnern können. Doch nicht nur das jeweilige Lebens-Alter prägt sich Stimmen ein, auch Krankheiten oder die jeweilige Stimmung und Befindlichkeit färben sie. Eben diese Unbeständigkeit und Sterblichkeit der menschlichen Stimme war seit Jahrhunderten *Movens* für Erfinder, Techniker und Ingenieure, die die stets schwindende Stimme zu verewigen und auf Dauer zu stellen suchten. Im Jahr 1779 schrieb die Akademie der Wissenschaften zu Sankt Petersburg einen Wettbewerb aus zur Untersuchung der Vokallaute und zur Konstruktion eines Instruments, das sie künstlich nachahmen sollte. Die Maschine von Kempelen gewann den Wettbewerb, sie konnte das Wort *exploitation* richtig aussprechen. Wenige Jahre später schreibt ein anonymes Autor über eine andere Sprechmaschine: »Ich wage vorherzusagen, dass nur die *Sprechenden Köpfe* diese ehrenwerte Vielfalt der französischen Sprache werden erhalten und gegen die Unbeständigkeit der menschlichen Dinge sichern können.«<sup>1</sup> All diese technischen Versuche und die gesamte Medientechnik der Stimme künden vom Paradox der Macht und zugleich Ohnmacht der Stimme. Heute besitzen wir zwar noch immer keine Möglichkeiten, die menschliche Stimme annähernd lebensecht künstlich nachzuahmen, doch existieren seit gut hundert Jahren – also noch nicht allzu lange – verschiedene akustische Medien, die Stimmen speichern, konservieren und übertragen können, und deren Auswirkungen auf das Sprechen und auf die Stimmen derjenigen, die mit diesen Techniken umgehen, noch weitgehend unerforscht sind. Das in den letzten Jahren verstärkte zu beobachtende Interesse zahlreicher Wissenschaften an stimmlichen Phänomenen dürfte maßgeblich durch diese Medien-Geschichte angestoßen sein.<sup>2</sup>

Die Stimme von Marianne Hoppe ist nicht nur im zeitlichen Sinne verklungen, sondern auch im existentiellen Sinne verstummt. Ihre technische Reproduktion offenbart gerade in der Wiederholung einen Riss, eine Differenz zwischen dem Sagen und dem Gesagten, zwischen den stimmlichen Lauten und den von ihnen artikulierten Zeichen, zwischen verschiedenen Formen von Anwesenheit. Einen Riss zwischen Körper und Sprache, der jede stimmliche Artikulation durchzieht, der jedoch in der technischen Reproduktion der Stimme einer Verstorbenen in besonderer Weise manifest wird. Wir hören nicht nur die abgründige Geschichte über einen Stimmenimitator, die uns Rätsel aufgibt und die wir, wenn wir es wollen, jederzeit nachlesen können.<sup>3</sup> Wir hören zugleich etwas, das ausgesprochen, aber nicht artikuliert, das gezeigt, aber nicht gesagt wird, das also den Charakter eines Zeichens übersteigt und die Einheit von Stimme und Sinn aufbricht: die eigentümliche Präsenz eines – im Falle des Tonbands oder auch des Telefons – abwesenden Körpers, die Körperlichkeit des Sprechens, die Michel Serres

als »Fleisch« und Roland Barthes als »Korn« der Stimme beschrieben haben. Wir hören nicht nur, was Marianne Hoppe sagt, wir hören Marianne Hoppe atmen, hören das Schmatzen und die Bewegungen ihrer Lippen, ihrer Zunge und ihres Gaumens, hören das leibliche Fundament des Sprechens und der Sprache. Auch werden wir – in den Pausen, im An- und Absetzen der Stimme – aufmerksam auf das Schweigen als Voraussetzung des Sagens: auf unser eigenes Schweigen während des Zuhörens ebenso wie auf das Schweigen, das dem Sagen vorausgeht und ihm nachfolgt. Zugleich hören wir nicht einfach einen Körper, sondern wir hören einen bestimmten Körper, dem wir ein Geschlecht, ein Alter, einen sozialen Status, bisweilen auch eine regionale Herkunft zuschreiben – wobei wir uns in und mit diesen Zuordnungen durchaus täuschen können. Zu hören gibt die Stimme Hoppes darüber hinaus eine bestimmte Technik und Tradition, eine geschulte und geübte Praxis, ja Virtuosität des schauspielerischen Sprechens. Und um diese vorläufige, keineswegs erschöpfende Bestandsaufnahme abzuschließen, höre ich in der Stimme der alten Marianne Hoppe – die Aufnahme des *Stimmenimitators* wurde 1999 gemacht, Hoppe war damals 90 Jahre alt – ein gelebtes Leben, die Spuren eines Jahrhunderts, die in der Reibung, der Verbrauchtheit und der gelassenen Souveränität dieser Stimme nachklingen. Das Tonband vermittelt mithin eine widersprüchliche, ja paradoxe, unerhörte Erfahrung: Im Jetzt des Hörens erleben wir ein Wirkliches in seinem vergangenen Zustand. Vor diesem Hintergrund möchte ich meine Überlegungen der im Oktober 2002 verstorbenen Marianne Hoppe widmen, denn die Faszination, die ihre Stimme auf mich auszuüben vermochte, stellt einen nicht wegzudenkenden Impuls für meine Auseinandersetzung mit Stimmen dar. Wenn ich angebe, dass Marianne Hoppes Stimme in Robert Wilsons Frankfurter *Lear*-Inszenierung oder in Heiner Müllers *Quartett* am Berliner Ensemble mein Nachdenken über Stimme maßgeblich inspiriert hat und noch heute inspiriert, dann geschieht dies weder, um eine subjektive Vorliebe von mir auszustellen, noch um der Stimme der Hoppe eine objektive oder gar paradigmatische Qualität zuzuweisen. Es geschieht vielmehr aus der Überlegung und der noch immer offenen Frage heraus, wie ich in angemessener Weise über Stimme sprechen und schreiben kann.<sup>4</sup> Wie also wäre der Stimme eine Stimme zu geben?

## 2. ATOPISCH

Bernhards *Stimmenimitator* gibt zwar keine Antwort auf diese Frage, aber doch einen bemerkenswerten Hinweis, dem ich nachgehen möchte und der mich über einige Umwege führen wird, bevor ich präzisieren kann, welche spezifischen Pro-

bleme die Stimme aufwirft, will man sich ihr (nicht nur) wissenschaftlich nähern. Der Stimmenimitator begeistert das Publikum, indem er mehr oder weniger berühmte Stimmen nachahmt und diese Imitationen beständig erweitert und wiederholt. »Als wir ihm jedoch den Vorschlag gemacht hatten,« – so die Erzählerin (wenn Hoppe den Text spricht) oder der Erzähler (wenn wir Bernhards Geschichte lesen) – »er solle am Ende seine eigene Stimme imitieren, sagte er, das könne er nicht.«<sup>5</sup> Der Imitator kann also andere, fremde Stimmen nachahmen, er kann sich das Fremde aneignen, in der Stimme einverleiben und zugleich entäußern. Es macht vielleicht gerade die Faszination des Stimmenimitators aus, dass er dem Publikum eine bekannte, wiedererkennbare Stimme vorführt, die gleichwohl aus einem anderen, fremden Körper kommt. Allerdings scheitert der Stimmkünstler an der Fremdheit, der Unzugänglichkeit seiner eigenen Stimme. Er scheitert an einer Alterität, die ausgerechnet in seinem Rollen- und Verwandlungs-Medium par excellence – der Stimme – ein Subjekt, eine Person unverwechselbar hindurchtönen lässt: *Personare* bedeutete im antiken griechischen Theater das Hindurchtönen der Stimme durch die trichterförmige Maske, die das Gesicht des Schauspielers bedeckte. So hallt im Begriff der Person die enge Verbindung von Stimme und Identität noch nach, meinte *persona* zunächst doch nichts anderes als die Rollenfigur, die sich im Sprechen artikuliert und verkörpert.

Von Bernhards *Stimmenimitator* ist zu lernen, dass die Stimme niemals allein vom Ort des Subjekts her analysiert werden kann, sie ist kein Objekt, das vom erkennenden Subjekt trennscharf zu scheiden wäre. Eine Phänomenologie der Stimme untersteht nicht dem Primat der Reflexion auf die eigene Stimme, denn meine eigene Stimme und die Stimme der oder des Anderen zeigen und enthüllen je anderes. Wahrscheinlich trägt dieser Umstand – neben dem Okularzentrismus des neuzeitlichen Europa, also der Dominanz des Visuellen sowohl in anthropologischen, ontologischen, erkenntnistheoretischen oder ästhetischen Diskursen als auch in zahlreichen Praktiken unserer Kultur – dazu bei, dass die konkret erklingende Stimme in den Überlegungen der Sprachphilosophie, der Kulturtheorie und Ästhetik aber auch der Soziologie bislang einen so begrenzten Raum einnimmt: Die Stimme erschließt sich nicht hinreichend in der Selbstreflexion, sie geht nicht in jenen Beziehungen der Selbstbespiegelung und jenen Techniken der Selbsterkenntnis auf, die unsere Kultur und insbesondere unser neuzeitliches Wissenschaftsverständnis spätestens seit Descartes privilegieren. Zwar exponiert die Stimme im Sich-selbst-sprechen-Hören eine prägnante Form des Selbstbezugs und der Selbstpräsenz, die Derrida jedoch genau deshalb als geradezu masturbatorische Selbstaffektion des gespaltenen Subjekts kritisiert, da da-

mit eine Nähe zum eigenen Selbst und eine Intimität zum anderen vorgegeben und die Differenzen im Eigenen wie zum Fremden negiert werden. Wenn ich mich selbst sprechen höre, kann es zwar geschehen, dass ich mich als eine andere vernehme, beispielsweise wenn ich vor dem Klang meiner eigenen Stimme erschrecke. Gleichwohl wird im Sprechen das Gesprochene nicht dadurch kenntlich, dass ich meiner eigenen Stimme lausche und mich darin des Gesagten versichere. Was die Bedeutung der Stimme und ihre Performanz ausmacht, kommt also weniger in der selbstreflexiven Struktur meines Sprechens zum Vorschein, sondern vielmehr in der Wirkung, die die fremde Stimme ausübt, die mich anspricht und die an mich appelliert.<sup>6</sup> Über diesen Appellcharakter der Stimme schreibt Roland Barthes:

Die Aufforderung zum Zuhören ist das vollständige Ansprechen eines Subjekts: Sie stellt den gleichsam körperlichen Kontakt zwischen diesen zwei Subjekten (durch die Stimme und das Ohr) über alles: Sie schafft die Übertragung: »Hör mir zu« heißt: *Berühre mich, wisse, dass ich existiere.*<sup>7</sup>

In diesem Appell, diesem Anspruch sind nicht nur Stimme und Ohr, Sprechen und Hören verschränkt. Barthes erweitert das Hören um das Fühlen, weist ihm eine taktile und existentielle Dimension zu. Sprechen wie Hören sind nicht isolierbar, sondern bleiben eingebunden in das Zusammenspiel aller Sinne. Kein Mensch lernt sprechen, der nicht hören kann. Die gesprochene Sprache tritt immer als ein Doppelpes auf: »als Passion des Ohrs und als Aktion der Stimme, die gegeneinander um eine Nuance verschoben sind.«<sup>8</sup> Von daher erscheint es schlüssig, dass eine literarische Urszene des Sprechens sowohl den Appellcharakter der Stimme als auch die Verschränkung von Sprechen, Hören und Fühlen betont: Kafkas *Bericht für eine Akademie*. Und es spricht für den von mir eingangs erwähnten Zusammenhang von Medientechnologie und Stimm-Bewusstsein, dass das erste Wort des sprechenden Affen Rotpeter das archetypische Wort moderner Tele-Kommunikation ist, das Telephonwort schlechthin: »Hallo«:

Was für ein Sieg dann allerdings – so der Affe in seinem Bericht – als ich eines Abends vor großem Zuschauerkreis – vielleicht war ein Fest, ein Grammophon spielte, ein Offizier erging sich zwischen den Leuten – als ich an diesem Abend, gerade unbeachtet, eine vor meinem Käfig versehentlich stehen gelassene Schnapsflasche ergriff, unter steigender Aufmerksamkeit der Gesellschaft sie schulgerecht entkorkte, an den Mund setzte und ohne Zögern, ohne Mundverziehen, als Trinker von Fach, mit

rund gewälzten Augen, schwappender Kehle, wirklich und wahrhaftig leer trank; nicht mehr als Verzweifelter, sondern als Künstler die Flasche hinwarf; zwar vergaß den Bauch zu streichen; dafür aber, weil ich nicht anders konnte, weil es mich drängte, weil mir die Sinne rauschten, kurz und gut »Hallo!« ausrief, in Menschenlaut ausbrach, mit diesem Ruf in die Menschengemeinschaft sprang und ihr Echo: »Hört nur, er spricht!« wie einen Kuss auf meinem ganzen schweißtriefenden Körper fühlte.<sup>9</sup>

Dieser Zusammenhang von Hören und Fühlen scheint mir besonders wichtig, weil er ein Konzept des Hörens zu entwickeln erlaubt, das das Hören nicht als ein bloß passives Empfangen, sondern als eine gerichtete Aktivität versteht. Kafka erinnert an eine Grundeinsicht der Theorien des Performativen, dass nämlich jeder, der spricht (oder schweigt), unter die Konsequenz seines Tuns gerät,<sup>10</sup> und er betont die Interdependenz von Hören und Fühlen, von Hören *als* Fühlen, die noch in dem unsäglichen Sprichwort »wer nicht hören will, muss fühlen« durchscheint.<sup>11</sup> Dagegen gilt in der neuzeitlichen Hierarchie der Sinne und des Sinns der Sprecher als aktiv, während der Hörer als passiv und allenfalls reagierend vorgestellt wird. Aus dieser asymmetrischen Relation speist sich auch die etymologische Nähe von Hören und Hörigkeit, Gehorsam. Gegen diese traditionelle Auffassung des vermeintlich passiven Hörens ist sowohl auf entwicklungsgeschichtlich-physiologischer Ebene als auch auf ästhetisch-philosophischer Ebene die Aktivität des Hörens zu entdecken. Hierzu zunächst ein kurzer entwicklungsgeschichtlicher Exkurs.

Nur wenige Tage nach der Befruchtung einer weiblichen Eizelle beginnt der Fötus, noch bevor er ein Embryo ist, die Organe des Ohres aus der gleichen Zellschicht zu bilden, aus der er auch seine Haut bildet. Dieser evolutionäre Zusammenhang von Haut und Ohr, von Fühlen und Hören setzt sich fort, wenn man das eigentliche Hörorgan, das sogenannte Labyrinth in der Cochlea genauer betrachtet. Das Innenohr verfügt über circa 30 000 Haarzellen, die dort in mehr oder weniger gleichmäßigen Reihen stehen. »Im Elektronenmikroskop sieht das aus, als seien sie aufgereiht wie die Tonfiguren-Armee der chinesischen Kaiser in Tonkin – jedes einzelne Härchen, Stereozilie genannt, eine winzigkleine Figur.«<sup>12</sup> Diese Härchen enthalten Eiweißelemente, wie sie auch im Muskelgewebe vorkommen und die Basis für Kontraktionen bilden. Ähnlich wie Muskeln können auch diese Härchen aktiv werden. Sie richten sich auf, wenn ein Ton erklingt, und es richten sich bemerkenswerter Weise immer auch die Zilien der nächstverwandten Obertöne – also Oktaven, Quinten, Quartan – auf, vergleichbar den »Menschen« – wie Joachim Berendt anfügt – »die sich aufrichten, um irgend et-



was, das sie gern hören möchten und das nur sehr leise wahrnehmbar ist, besser hören zu können.«

Die Aktivität des Ohres, die Berendt in einem nicht unproblematischen anthropomorphen Vergleich zu veranschaulichen sucht, bleibt ohne Mikroskop verborgen. Von außen betrachtet – und ich unterstreiche hier die visuelle Metapher des Betrachtens – erscheint das Ohr als passiv. Je stärker sich der neuzeitliche Mensch auf die Priorität und Dominanz des Sehens kaprizierte, desto einseitiger wurde das Hören als eine bloße Passivität gedacht. Karl Josef Pazzini vermerkt dazu:

Es ist eine dieser physiologischen Anekdoten, dass das Hören passiv sei, eine Anekdote, die das Auge erfunden hat. Was sieht das Auge am Ohr: Es bewegt sich nicht. Was schließt das Auge messerscharf? – Da tut sich nichts. Genauso wenig sieht das (männliche) Auge am Geschlecht der Frau. Da muss doch was sein! Und es weiß, dass da was ist. Deshalb muss es immer wieder hinsehen. Aber da ist nichts zu sehen. Oder es ist nicht das zu sehen, was es weiß. Das ist ärgerlich. Mit dem Ohr ist das sehr ähnlich.<sup>13</sup>

Die Zusammenführung von Ohr und Sexualität mag hier zunächst in ihrer Unvermitteltheit überraschen, sie beruht jedoch auf vielfältigen Bezügen. Die Tabuisierung und Vereinseitigung sowohl des Hörens – und damit auch der Hörigkeit, des Gehorsams – als auch der Sexualität, insbesondere der weiblichen, haben durchaus miteinander zu tun: Beide durchkreuzen die aufklärerische Vorstellung von der Autonomie des Subjekts. Gehörtes lässt sich weit weniger beherrschen als Gesehenes oder Ergriffenes. Wir können die Augen öffnen und schließen, wir können sie auf Ausschnitte richten und zentrieren, wir können mit unseren Händen gezielt nach etwas greifen oder es unterlassen. Die Wahrnehmungen des Ohres sind diffuser und weniger steuerbar, wir sind ihnen stärker ausgesetzt, sie dringen aufdringlich in uns ein. Selbst im Schlaf, der nur bei geschlossenen Augen möglich ist, ist der Körper über das Ohr noch mit der Umwelt verbunden und hört auf akustische Reize. Hören steht – wie die Sexualität – im Verdacht zu verführen und abhängig zu machen. Die Sirenen-Episode im 12. Gesang von Homers *Odyssee* ist der zur Mythe geronnene Beleg für diese Auffassung,<sup>14</sup> und auch der Mythos von Narziss, der zunächst von der Nymphe Echo, einem rein akustischen Wesen, verführt wird, bevor er seinem eigenen Spiegelbild verfällt, rekuriert darauf.

Hören und insbesondere Einer-Stimme-Zuhören ist weit mehr als ein bloß passives Aufnehmen äußerer Reize. In Anlehnung an Jacques Lacans Bestim-

Doris Kolesch

274

mung der Stimme als Triebobjekt hat Roland Barthes gerade die aktive und affektive Dimension des Stimmen-Hörens betont, die nicht nur eine oder einen jeweils Sprechenden voraussetzt, sondern immer auch das zuhörende Subjekt im Wortsinne ins Spiel bringt.

Es gibt keine menschliche Stimme auf der Welt, die nicht Objekt des Begehrens wäre – oder des Abscheus: Es gibt keine neutrale Stimme – und falls mitunter diese Neutralität, dieses Weiß der Stimme auftritt, so ist dies für uns ein großes Entsetzen, als entdeckten wir mit Schrecken eine erstarrte Welt, in der das Begehren tot wäre. Jeder Bezug zur Stimme ist zwangsläufig einer der Liebe [...].<sup>15</sup>

Damit schließt sich der Kreis zu meiner Ausgangsfrage, wie wir über Stimmen reden und nachdenken können, und es wird hoffentlich deutlich, weshalb ich von meiner Faszination für Marianne Hoppes Stimme gesprochen habe. Wenn das Wahrnehmen von Stimmen immer ein affektiver Prozess ist, der Sprechende und Hörende gleichermaßen involviert, dann hat dies weitreichende Konsequenzen dafür, mit welcher Perspektive und welcher Begrifflichkeit wir uns der Stimme in der Reflexion nähern können. Das bis heute geläufige Konzept eines distanziereten, unbeteiligten »Beobachters« auf der einen Seite und eines davon getrennt wahrgenommenen Gegenstandes auf der anderen Seite verfehlt offenbar wesentliche Aspekte von Stimme und Stimmlichkeit oder blendet sie aus. Die Stimme ist kein abgrenzbarer Gegenstand, kein Ding, sondern ein raumgreifendes, ja raumschaffendes akustisches Geschehen. Stimmen können uns umhüllen, wir können in sie eintauchen, sie können uns aber auch wie ein Pfeil, ein *punctum* treffen oder uns schleichend, hinterhältig einnehmen; immer jedoch bindet die Stimme in besonderer Weise die Sprecherin an ihr Sprechen und an ihre Zuhörer ebenso wie sie die Zuhörer mit der Sprecherin verbindet.

Wenn das bislang Gesagte zutrifft, dann überschreitet die Stimme etablierte Grenzen: Sie gehört weder dem Sprechenden noch dem zuhörenden Subjekt; sie ist weder nur bei der, die spricht, noch bei dem, der hört. Sie bedarf der Hohl- und Resonanzräume, um zu klingen, doch verweigert sie zugleich die Entscheidung darüber, ob Klang und Stimme nun drinnen oder draußen, innerhalb oder außerhalb dieser Resonanzräume sich ereignen. Sie ist weder nur präsentische, einmalige Gegenwart noch nur wiederholende Re-Iteration. Auch kann sie weder durch ihre materiellen, ihre körperlich-sinnlichen Aspekte hinreichend bestimmt werden noch durch ihre eher immateriell konnotierten Dimensionen wie Atem, Hauch, Seele.<sup>16</sup> Vergleichbar der komplexen Gegebenheit unseres Körpers,

die Plessner als Körper-Haben und Leib-Sein umschrieb, könnten wir für unsere Stimme formulieren: Wir *haben* eine Stimme und wir *sind* zugleich Stimme. Die Stimme erscheint so als paradigmatische Figur der Überschreitung. Sie ist eine Transgression, die die Grenze, die sie überschreitet, zugleich begründet. Sie wäre somit weder diesseits von Präsenz noch jenseits von Repräsentation, weder diesseits des Körpers noch jenseits der symbolischen Ordnung, sondern beides, dazwischen. Dies zeigt sich besonders eindrücklich an der Telefonstimme, die die Grenze, die sie überbrückt, hervorbringt und exponiert. Die Grenze zwischen Hier und Dort, zwischen Anwesenheit und Abwesenheit, zwischen Sagen und Zeigen wird von ihr gezogen und zugleich überschritten. Damit ist die Stimme nicht einfach etwas, das Sprecher und Hörer verbindet, sondern sie verkörpert in ihrer szenischen Qualität zwischen Aufführung und Wahrnehmung diese Relation selbst.

Wir stehen also vor dem Paradox, etwas auf den Begriff bringen zu wollen, das sich nicht auf einen, schon gar nicht auf den Begriff bringen lässt. Ich möchte für dieses vertrackte Problem auf den wertvollen Ausdruck des Atopischen zurückgreifen.

Die Atopie widersteht der Beschreibung, der Definition, der Sprache als *maja*, als Klassifikation von Namen (Fehlern). Als *atopos* lässt der Andere die Sprache erbeben: man kann nicht *von* ihm, über *ihn* sprechen; jedes Attribut ist falsch, schmerzhaft, taktlos, peinlich: der Andere ist *unqualifizierbar* (das wäre die genaue Bedeutung von *atopos*).<sup>17</sup>

Diese Nichtqualifizierbarkeit stellt keinen Mangel, sondern eine Qualität dar – die zu entdeckende Qualität des Vermischten, des Gemenges, des Beweglichen, Flüssigen, Vielfältigen, Passageren. Sie erscheint desto wertvoller, wenn wir uns vergegenwärtigen, wie wir normalerweise über Stimmen reden: in ärmlichen, hilflosen Adjektiven und Attributen. Wir beschreiben Stimmen als warm, als hoch, als rau, dünn oder auch sexy. Wenn ich nun die Stimme im Gegensatz dazu als atopisch bezeichne, als ein Phänomen, das sich systematischer Definition und Klassifikation widersetzt und sich der eindeutigen Verortung entzieht, ist die Paradoxie meines eigenen Sprechens und Schreibens zu betonen: Ein Diskurs, der Nicht-Diskursives einzufangen sucht, ist notwendigerweise charakterisiert durch Unvollständigkeiten, durch Verfehlungen, Lücken, Diskrepanzen. Und es ist zu fragen, ob und wie sich das prekäre Dazwischen der atopischen Stimme in ihrer Medialität niederschlägt.

### 3. AISTHETISCH

Als ein Anknüpfungspunkt bietet sich hier der Medienbegriff Marshal McLuhans an. Vor dem Hintergrund technologischer Entwicklungen und techniktheoretischer Überlegungen hat McLuhan schon in den sechziger Jahren die Eigensinnigkeit von Medien betont und die geläufige Auffassung widerlegt, Medien seien neutrale Vermittler, neutrale Vehikel des Sinns. Die bekannte und noch immer ernst zu nehmende These, das Medium selbst sei die Botschaft, fußt auf einem anthropomorphen Technikverständnis. McLuhan konzipiert das Verhältnis von Mensch und Technik, von Mensch und Medium nach dem Muster funktioneller Äquivalenz. Mechanische Techniken kommen entsprechend als Veräußerung körperlicher Funktionen des Menschen in den Blick, elektronische Medien als Exteriorisierung der Sinnesorgane und des Zentralnervensystems. Das Auto wäre eine mechanische Erweiterung und Steigerung der Bewegungsfähigkeit des Körpers, die Videokamera eine Veräußerung und Verstärkung des Auges. Nimmt man diesen McLuhanschen Medienbegriff zum Ausgangspunkt, zeigt sich die Differenz, die ins Spiel kommt, wenn es um die Stimme als Medium geht: Es kann in einzelnen Situationen durchaus plausibel sein, von der Stimme als einer Verlängerung, einer Erweiterung und Extension des Körpers zu sprechen. So kann ich bei einem Vortrag mit meiner Stimme auch die Personen in der hinteren Reihe erreichen, oder mich stimmlich in der Dunkelheit zu erkennen geben. Doch selbst angesichts derartiger Fälle springt die Einseitigkeit und Unzulänglichkeit einer solchen Bestimmung ins Auge. Die Stimme ist nicht einfach nur Verlängerung des Körpers, sie *ist* in gewisser Weise Körper und übersteigt diesen zugleich. Bernhards Erzählung *Der Stimmenimitator* handelt ja nicht zuletzt davon, dass wir die Stimme als ein Instrument benutzen können, dass sie zugleich aber mehr und anderes ist als ein kontrollierbares, entsprechend einsetzbares Instrument. Spricht die Stimme oder bedienen wir uns ihrer? Die Frage scheint falsch gestellt, trifft doch wohl beides gleichzeitig zu. Wer schon einmal in einer Prüfungssituation seine aufgeregte Stimme zu kontrollieren suchte, hat eine Erfahrung davon gemacht, dass die Stimme eine Spur bewahrt – eine Spur des Körpers, des Semiotischen im Sinne Kristevas, aber auch bestimmter kultureller Praktiken, Normen, Traditionen und Disziplinierungen. Sie fungiert gleichzeitig als Zeichen *von* etwas und als Anzeichen *für* etwas, als Index – dessen Besonderheit ja gerade darin liegt, dass er nicht – wie das sprachliche Zeichen – ein willkürliches Zeichen darstellt, dessen Beziehung zum Gegenstand durch eine Regel oder Konvention festgelegt wird. Der Index unterhält einen physischen, materiellen Zusammenhang zum Bezeichneten. Von daher gibt es keine »reine«, körperlose Stimme.

Die reine Stimme ist ein Mythos. Sie ist ohne Verbindung zum Ausdruck und zum Sinn nicht vorstellbar, obwohl sie in der Identität zugleich eine Differenz zu ihnen bewahrt, die allein es ermöglicht, dass sie als *Material* des Signifikanten *dienen* und zugleich *sich* als *Signifikant geben* kann.<sup>18</sup>

Der stimmlich artikulierte Diskurs muss damit als eine andere Form der Verkörperung von Sprache aufgefasst werden als beispielsweise der schriftliche Diskurs. Dem Fluidum der Stimme eignet eine andere raum-zeitliche Materialität als den diskreten Elementen der Buchstaben, Worte und Interpunktionszeichen. Entsprechend ziehen Stimme und Schrift eine jeweils differente Fokussierung unserer Aufmerksamkeit nach sich und konstellieren das Verhältnis von Sinn und Sinnen jeweils anders. Während der schriftliche, insbesondere der gedruckte Text von seinem Verfasser losgelöst zirkulieren kann, ist uns eine solch radikale Trennung einer stimmlichen Äußerung von der sie äußernden Person kaum möglich. Vor diesem Hintergrund erweist sich der stimmlich artikulierte Diskurs als Diskurs in einem spezifischen, durchaus überraschenden Sinn: *Discursus* meint ursprünglich die Bewegung des Hin- und Herlaufens, des Auseinanderdriftens, des Kommens und Gehens, der Verwicklungen. Der stimmliche Diskurs – so könnten wir nun präzisieren – übersteigt gerade seine Diskursivität, insofern ein nicht-diskursives, ein nicht arbiträres, analoges Moment die Bedingung seiner Möglichkeit darstellt. Im stimmlichen Diskurs läuft Diskursives und Nicht-Diskursives, laufen Sagbares und Zeigbares hin- und her und ineinander.

Die Nicht-Verortbarkeit, das Atopische durchzieht also auch die Stimme als Medium. Genau hier erreichen wir eine Grenze, an der die Stimme mehr und anderes ist als ein Medium der Sprache. Ich meine das Phänomen, dass die Stimme im Kommentar, in der Bekräftigung, der ironischen Hinterfragung oder gar der Subversion des Gesagten eine Dimension und Funktion erlangt, die nicht mehr in der sprachlichen *Semiosis* und der Stimme als Medium der Sprache aufgeht. Paul Zumthor führt dazu aus:

Wenn der geschriebene Text Stimme wird, wandelt er sich grundlegend, solange Hören und sprachlich-körperliche Gegenwart andauern. Jenseits der Gegenstände und der Bedeutungen, auf die sie sich bezieht, verweist die Rede auf das Unbenennbare. Die Rede ist nicht einfach die Vollstreckerin des Sprachsystems. Sie bestätigt nicht nur nicht vollständig dessen Vorgaben, sondern handelt ihm oft, in ihrer ganzen Körperlichkeit, zu unserer Überraschung und unserem Vergnügen zuwider. [...] Die Rede gibt sich als Erzählung, sie wird aber gleichzeitig im Klang der Stimme

und der Bewegung des Körpers, die ihr Ausdruck verleihen, zum Kommentar des Erzählten.<sup>19</sup>

Wir kommen damit an einen Punkt, an dem uns die Betrachtung der Stimme als Medium von Sprache nicht weiterführt. Denn selbst wenn wir die geläufige Definition von Medien als Vermittlungsinstanzen annehmen, also als Mittler und Vermittlung in einem, als Phänomene, die das, was sie übertragen, in dieser Übertragung erst hervorbringen, können wir wesentliche Dimensionen von Stimme, wie die oben von Zumthor genannten, nicht erfassen. Dies hat meines Erachtens weniger mit dem jeweiligen Medienbegriff zu tun als vielmehr mit dem jeweiligen Begriff und Verständnis von Sprache – genauer: mit der Reduktion von Sprache auf Kommunikation, auf den Austausch von Informationen, Meinungen, Sichtweisen und Positionen, die man mit dem Modell von Sender, Nachricht und Empfänger zu beschreiben sucht.

Jochen Hörisch hat auf eine andere, grundlegende Funktion von Medien aufmerksam gemacht, die ich hier ins Spiel bringen möchte: Medien koordinieren Interaktion: »Wir machen den Mund auf – so Hörisch –, um mit Leuten, die offene Ohren haben, das Tun unserer Hände und Füße zu koordinieren.«<sup>20</sup> Sprechen erscheint damit als mehr und anderes als Kommunikation im oben erwähnten Sinn.

Um das zu präzisieren, scheint mir wiederum ein kleiner evolutionsgeschichtlicher Exkurs von Nöten. Der Anthropologe André Leroi-Gourhan hat in seiner Schrift *Hand und Wort* darauf aufmerksam gemacht, dass der evolutionäre Erwerb der Zweifüßigkeit dem Menschen neben der Befreiung der Hand und damit der Entwicklung von Werkzeugen und Technik auch eine Befreiung des Mundes von den Funktionen des Zugreifens ermöglichte. Nun konnte der Mund zur Herausbildung von Sprache eingesetzt werden. Diese anthropologische Verquickung von Körper, Stimme und Sprache hat Roland Barthes zum Ausgang genommen, um eine andere, lustvolle und sinnliche Vision des Sprechens zu entwerfen:

Als er zum aufrechten Gang übergang, sah sich der Mensch freigeworden, die Sprache und die Liebe zu erfinden: es ist vielleicht die anthropologische Entstehung einer doppelten, zusammenwirkenden Perversion: das Sprechen und der Kuss. So betrachtet haben die Menschen, um so freier sie (mit dem Mund) waren, um so mehr gesprochen und geküsst; und logischerweise werden die Menschen, wenn sie sich durch den Fortschritt aller Handarbeit entledigt haben, nur noch reden und küssen!

Stellen wir uns für diese doppelte Funktion, die an einer Stelle loka-

liert ist, eine einzige Überschreitung vor, die aus dem gleichzeitigen Gebrauch von Sprechakt und Kuss entstehen würde: *beim Küssen sprechen, beim Sprechen küssen*. Man könnte annehmen, dass es diese Wollust gibt, da die Liebenden unaufhörlich »das Wort von den geliebten Lippen trinken«. Was sie da genießen, ist im Liebesringen das Spiel mit dem Sinn, der aufblüht und sich unterbricht: die Funktion *in der Verwirrung*: mit einem Wort: *der gestammelte Körper*.<sup>21</sup>

Diese beflügelte anthropologische Spekulation stilisiert das Sprechen des Menschen zum zweckfreien, nicht auf Kommunikation reduzierbaren lustvollen Akt. Vergleichbares deutete schon Kafka an, indem der Affe Rotpeter den Ausruf der ihn bestaunenden Menschenmenge »Hört nur, er spricht!« wie einen Kuss auf [s]einem ganzen schweißtriefenden Körper fühlte.« Nicht nur das Ohr wäre damit als sinnlich-sensuelles Organ zu rekonzeptualisieren, auch die Stimme, die verführen, umschmeicheln, überreden, bisweilen auch überwältigen will. Das Potential der Stimme zu verführen, zu überreden und zu überwältigen, ist grundsätzlich ambivalent: Es kann weder per se als positiv noch als negativ angesehen werden, und es kann mit jeweils unterschiedlicher, ja sich wechselseitig ausschließender Absicht eingesetzt werden. Die Stimmen der Liebenden, aber auch die dazu diametral entgegengesetzte Stimme des Führers, des Befehlsgebers, rekurrieren beide mit differenter Intention und Wirkung auf dieses Potential.

Hier setzt die Aufgabe einer performativen Ästhetik der Stimme an. Ästhetik ist dabei nicht zu verstehen als Lehre vom Schönen oder als Lehre vom Kunstwerk, sondern sie ist in einem weiteren Sinne als *Aisthesis* aufzufassen, als sinnliche Wahrnehmung und Erkenntnis. Die implizit szenische, ja theatrale Qualität der Stimme, die immer schon Aufführung und Wahrnehmung ist, hat zur Folge, dass eine Stimme gehört und beantwortet werden will. Sie ist Anspruch, Appell und Gabe in einem. Diese Gabe, dieser Anspruch zeigt sich nicht nur in dem, was gesagt wird, sondern vor allem auch im *Wie* des Sagens: in seinem Rhythmus, seiner Musikalität, seiner Klangfarbe und Modulation. Was in diesen ästhetischen, wechselseitig verschränkten Prozessen performiert wird, ist nicht bloß Kommunikation unter Menschen, sondern es ist eine spezifische, jeweils unterschiedlich zu bestimmende Form der Begegnung, der Nähe (oder auch der Distanz, der verweigerten, gefürchteten Nähe). In dieser Perspektive kommt das Ästhetische als Fundament einer Ethik in den Blick, nicht umgekehrt.

Die Begegnung einer Stimme, die Begegnung in der Stimme lässt sich nicht auf Verständigung oder Konsens zurückführen. Wenn Einigkeit und Konsens wirklich herrschen, dann kommen Kommunikation und Interaktion zwischen

Menschen zum erliegen und werden überflüssig. »Wer dem anderen zustimmt, hat weiter nichts zu sagen. Kommunikation würde, wenn Verständigungs- und Konsenstheoretiker [– wie Habermas oder Apel –; D. K.] denn recht hätten, ausgerechnet in dem Augenblick kollabieren, in dem sie ihr Ziel erreicht hätte.«<sup>22</sup>

Entsprechend kommt dem Dissens, der Reibung, dem Korn der Stimme und ihrem von mir explizierten ambivalenten, nämlich zugleich diskursiven und nicht-diskursiven, sagenden und zeigenden Diskurs ein besonderes Gewicht zu: das Gewicht, die Interaktionsfähigkeit und Interaktionsmöglichkeit des Menschen gerade nicht zu begrenzen und abzuschließen, sondern offen zu halten, einen Spiel-Raum des Handelns zu eröffnen. Die Stimme vermittelt, interagiert zwischen Instanzen, die zusammenkommen wollen, aber nicht zusammenkommen können – eine Situation, die die Stimme im Idealfall gerade nicht behebt, sondern weitertreibt. Dissens, Reibung oder auch das Korn der Stimme bestimme ich folglich nicht inhaltlich, sondern ästhetisch: als synchroner oder divergenter Rhythmus, als Kontrast oder Gleichklang des Lauten und Leisen, des Eigenen und Fremden, als wiederholtes Einstimmen oder Ausscheren in das Wechselspiel von Sprechen und Hören, von Sagen und Zeigen.

Hier schließt sich der Kreis zum Anfang meiner Ausführungen: Es ist gerade das vermeintliche Ende von Bernhards Erzählung, das in der vieldeutigen – allerdings schriftlich nicht repräsentierbaren und insofern in meinem Text *lautlos* bleibenden – Artikulation Marianne Hoppes *ein Ende* anzeigt (das Ende von Bernhards Erzählung), und zugleich *einen Beginn* – nämlich den Beginn meiner Auseinandersetzung mit dem, was sich da im Sagen zeigt.

1 Zitiert nach Jean-Loup Rivière: *Das Vage der Luft*, in: Dietmar Kamper/Christoph Wulf (Hg.): *Das Schwinden der Sinne*, Frankfurt/M. 1984, S. 103. Zu Wolfgang von Kempelens Stimm-Maschinen vgl. auch Brigitte Felderer: *Stimm-Maschinen. Zur Konstruktion und Sichtbarmachung menschlicher Sprache im 18. Jahrhundert*, in: Friedrich Kittler/Thomas Macho/Sigrid Weigel (Hg.): *Zwischen Rauschen und Offenbarung. Zur Kultur- und Mediengeschichte der Stimme*, Berlin 2002, S. 257–278.

2 Vgl. Karl-Heinz Göttert: *Geschichte der Stimme*, München 1998; Kittler/Macho/Weigel: *Zwischen Rauschen und Offenbarung* (Anm. 1).

3 Nämlich in: Thomas Bernhard: *Der Stimmenimitator*, Frankfurt/M. 1991, S. 9–10.

4 Unter »Stimme« verstehe ich im folgenden weder eine exemplarische Einzelstimme noch eine abstrakte, allen möglichen Stimmrealisierungen vorgelagerte Idee von Stimme, sondern die immer schon polyphone, ja heterophone Überlagerung verschiedener Stimmen in einer Stimme.

5 Bernhard: *Der Stimmenimitator* (Anm. 3), S. 10.

6 Vgl. Dieter Mersch: *Was sich zeigt. Materialität, Präsenz, Ereignis*, München 2002, S. 123–124.

7 Roland Barthes: *Zuhören*, in ders.: *Der entgegenkommende und der stumpfe Sinn*, Frankfurt/M. 1990, S. 255.

8 Dietmar Kamper: *Vom Hörensagen. Kleines Plädoyer für eine Sozio-Akustik*, in: Kamper/Wulf: *Das Schwinden der Sinne* (Anm. 1), S. 113. Wo das Hören ausfällt, kann die Rückkopplung auch über Tast- und Kontaktinformationen des Mund- und Racheninnenraumes geliefert werden, die ein dauerhaftes motorisches Programm (Sprechen) aufzubauen erlauben, welches auf die Sinnesrückmeldung schließlich sogar verzichten kann. Vgl. dazu Detlef B. Linke: *Die Stimme im Mund*, in: Hans



- Belting/Dietmar Kamper/Martin Schulz (Hg.): *Quel Corps? Eine Frage der Repräsentation*, München 2002, S. 467–475.
- 9 Franz Kafka: Ein Bericht für eine Akademie, in ders.: *Ein Landarzt und andere Drucke zu Lebzeiten. Gesammelte Werke Bd. 1.* (Nach der Kritischen Ausgabe hg. v. Hans-Gerd Koch), Frankfurt/M. 1994, S. 243.
- 10 Vgl. Kamper: *Vom Hörensagen* (Anm. 8), S. 114.
- 11 Horst Wenzel hat gezeigt, dass das Sprichwort ursprünglich den – inzwischen eher verschütteten – Zusammenhang aller Sinne bezeugt (vgl. Horst Wenzel: *Hören und Sehen – Schrift und Bild. Kultur und Gedächtnis im Mittelalter*, München 1995; ders.: *Die Empfängnis durch das Ohr. Zur multisensorischen Wahrnehmung im Mittelalter*, in: Thomas Vogel (Hg.): *Über das Hören. Einem Phänomen auf der Spur*, Tübingen 1998, S. 159–179).
- 12 Joachim Ernst Berendt: Ich höre, also bin ich, in: Vogel: *Über das Hören* (Anm. 11), S. 76.
- 13 Karl Josef Pazzini: Wer nicht hören will, muß fühlen. Einige Diskussionsbeiträge zum Hören in der Psychoanalyse und der Pädagogik, in: *Paragrana* 2/1–2 (1993), S. 20.
- 14 Homer: *Odyssee*, in: ders.: *Ilias – Odyssee* (übers. v. Johann Heinrich Voß), 1984, S. 601–608.
- 15 Roland Barthes: Die Musik, die Stimme, die Sprache, in: ders.: *Der entgegenkommende und der stumpfe Sinn* (Anm. 7), S. 280.
- 16 Zur metaphysischen Verbindung von Stimme, Seele, Präsenz und Identität im Logozentrismus vgl. Jacques Derrida: *De la grammatologie*, Paris 1967, zur kulturhistorischen Dimension von Atem, Seele, Hauch vgl. Gernot und Hartmut Böhme: *Feuer – Wasser – Erde – Luft. Eine Kulturgeschichte der Elemente*, München 1996.
- 17 Roland Barthes: *Fragmente einer Sprache der Liebe*, Frankfurt/M. 1988, S. 45.
- 18 Michael Wimmer: Verstimmte Ohren und unerhörte Stimmen, in: Kamper/Wulf: *Das Schwinden der Sinne* (Anm. 1), S. 125.
- 19 Paul Zumthor: Körper und Performanz, in: Hans Ulrich Gumbrecht/K. Ludwig Pfeiffer (Hg.): *Materialität der Kommunikation*, Frankfurt/M. 1988, S. 709.
- 20 Jochen Hörisch: *Der Sinn der Sinne. Eine Geschichte der Medien*, Frankfurt/M. 2001, S. 31–32.
- 21 Roland Barthes: *Über mich selbst*, München 1978, S. 153.
- 22 Hörisch: *Der Sinn der Sinne* (Anm. 20), S. 33.

**Anthony Moore / Siegfried Zielinski**  
**VOX PER → SONARE.**  
**EIN VORTRAGSDUETT IN ACHT TEILEN<sup>1</sup>**

1.

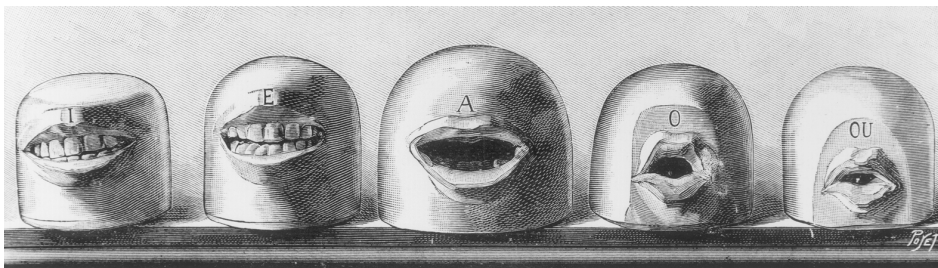
In einem Brief an Witold Gombrowicz schrieb der Autor der *Zimtläden* und des *Götzenbuchs*, Bruno Schulz, unsere Geschlechtlichkeit stamme aus einer anderen Entwicklungsepoche als unser geistiger Zustand. Die Stimme ist Teil unserer Geschlechtlichkeit. Sie stammt aus einer anderen Epoche als unser geistiger Zustand. Für sie gibt es keine Existenz außerhalb des lebendigen oder – in technischer Hinsicht – des gebauten vibrierenden Körpers. Es sei denn in der Form als Datenmaterial. Aber als solches ist sie nicht hörbar, sondern nur lesbar oder schreibbar. Für die Stimme ist der Körper Resonanz, die nicht zwei- oder dreidimensional projizierbar ist. Sie artikuliert sich aus seinem Inneren heraus und indem sie, wie durch eine Maske, durch ihn hindurch dringt; *personare*.

Die Herkunft des Wortes »Stimme« ist nicht ganz klar, aber am ehesten ist davon auszugehen, dass es die Weiterbildung eines Wortes für »Mund« (im Griechischen *stóma*, noch älter *staman* für »Maul«) ist, mit einer engen Verbindung zur »Kehle«. *Eústomos* heißt »schön redend«, was leicht verstanden werden kann als: »dessen Stimme schön ist«, wörtlich aber auch: »dessen / deren Mund schön ist«.

Stimme und Mund sind nicht voneinander zu trennen, genauso wenig wie die Zunge vom Mund. Es gibt unendlich viele Varianten unter den Grundaktionstypen der Zunge, der aufschlagenden Zunge, der durchschlagenden Zunge und der Zunge als Membrane. Der Kehlkopf ist unser innerer Mund. Er hat mehrere

**Abb. 1**

Photographie de la voix. Fig. 6. – Moulages de la bouche prononçant les voyelles.



Zungen. Das sind die Stimmbänder. Ihre Länge entscheidet auch über die Höhe oder Tiefe der Stimme.

Die Stimme ist ein Phänomen mit großem Variantenreichtum. Es gibt u. a. Bruststimmen und Fistelstimmen, helle und dunkle Stimmen, Kopfstimmen, Zwischenstimmen und die unendlichen Variationen zwischen Sprech- und Singstimmen. Noch Helmholtz sprach die Vermutung aus, dass Sprech- und Singstimme unterschiedliche Mechanismen zu eigen seien. Bei der Sprechstimme wären die Stimmbänder als gegenschlagende Zungen gestellt, bei der Singstimme als durchschlagende. Inzwischen wissen wir, dass Sprechen und Singen lediglich zwei unterschiedliche Arten, zwei Pole auf der Skala der Verwendung der Stimmwerkzeuge sind.

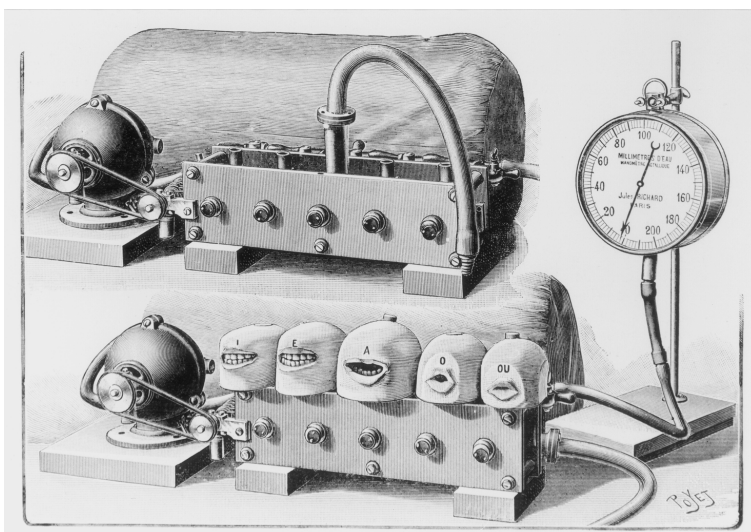
Je bestimmter im Singen die Töne festgehalten werden, desto besser ist es für den Gesang. Je weniger dagegen die Töne festgehalten und die Intervalle überschritten werden, desto besser für die Sprache. Widrigenfalls entsteht jenes widerliche Sprechen, welches man das singende nennt.<sup>2</sup>

Woher mag diese harsche Abwertung des singenden Sprechens als widerlich stammen?

Der in natürlichen Bahnen laufende menschliche Gesang ist in hohem Maße

Abb. 2

Photographie de la voix. Fig. 7. – Synthèse des voyelles. En haut, la sirène à voyelles. En bas, le même appareil surmonté des moulages.



der Produktionsfähigkeit der Drüsen, besonders der Keim- bzw. Geschlechtsdrüsen, unterworfen. Er ist von der innersekretorischen Drüsenfunktion abhängig. Unter innerer Sekretion versteht man »die von der Organumschlingung des nervus sympathicus geleistete Säftebildung in verschiedenen Organsystemen, denen sämtlich spezifische Funktionen zufallen«, schreibt Carl Ludwig Schleich 1916 in seinem Buch *Vom Schaltwerk der Gedanken*.<sup>3</sup> Der *nervus sympathicus* ist das unter dem Zwerchfell liegende Sonnengeflecht, das schon von den alten Griechen als der Sitz der Seele bezeichnet wurde . . . »Man könnte die Drüsen als Saftbereitungszentralen im Bau des Leibes bezeichnen, die vom nervus sympathicus angeregt und geleitet werden.«<sup>4</sup> – Ein *nervus antipathicus* ist im Körper bisher noch nicht ausgemacht worden.

## 2.

The timbral memory of another's voice rings as truly inside my head as ever it did in the outer air. And no magnetic field however malevolent may disarrange the particles or waves of its inner echo. Voices from behind the cinema screen are similarly disembodied and likewise emerging from the centre. Early experiments with the stereoscopic placing of the voice were abandoned as it was felt that the audience simply became distracted. With the sound of the voice coming from the middle of the screen however, even when the actor appears off to one side, the perception continues to be that the voice emerges from the mouth of the speaker. So, from the centre it does come, just as the inner sound of the remembered voice.

From the perspective of sound, the listener is at the vanishing point. What, for the eye, is in the distance, the point of convergence, is, for the ear, the closest. The vanishing point is the point of starting out and is located at the centre, the centre of an expanding sphere, a radial divergence that becomes a space which may contain other spaces. But at the centre, the very middle of anything, there is only room enough for one of us. It may be true that each point of view is also unique, separate. But you can look at someone looking at something and look at it too, a socialising triangulation. In a way, like thinking, you can never really know what people are hearing. Thinking and listening both start from the same place, the vanishing point between the ears. At the beginning of sound you are alone in listening, at the end of seeing you are alone in the vanishing point. At the beginning of seeing you can at least imagine others gazing down the sight-lines of the periphery, at the end of hearing, no longer at the centre but in the skin, you can at least imagine an expanding sphere that holds those same others. Seeing begins with the

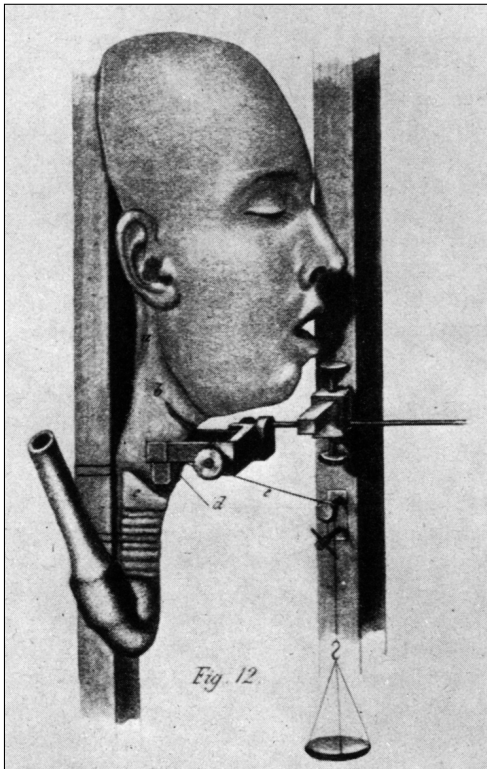


Abb. 3

Fig. 12. Darstellung eines Versuchs zum Studium der Stimm-bildung am isolierten menschlichen Kehlkopf (1839).

idea that you are not alone because you can't see behind you. The middle is in front and far away. Hearing begins with the idea of being alone because, though you can see someone looking at you, you can't hear anyone listening to you.

So, there is the voice emerging from the centre of the screen, and the inner sound of the remembered voice which likewise issues from the centre, the interior vanishing point and, in between these two approaching signals is the interface of the eardrum where the waves and

frequencies start to interfere. And real silences can come about when the inner and the outer, half a cycle apart, phase-cancel each other.

3.

Der Schrei und das Lallen sind nach Freud die ersten Äußerungen und Signale unbewusster sexueller Lust. Die sexuelle Begierde entwickelt sich zuallererst im Mund. Das frühe Lallen und Schreien ist nichts anderes als eine Stimulanz der

erogenen Mundzone, also eine Art autoerotische Abführung überschüssiger Triebenergien. Selbst brillianteste Mozart- oder Verdi-Arien sind in dieser Hinsicht nichts anderes als sorgfältig gepflegte, ausgebildete Töne der Lust, kosmetisierte und trainierte Töne des ursprünglichen Lallens. Im Belcanto, dem schönen Gesang, mit seinem Anspruch auf höchste Kultivierung der Stimme und formale Perfektion, erfährt diese erotische Dimension des Singens die größte Abstraktion. Die Form des Gesangs ist nahe am Bild.

Der Kehlkopf zeigt beim Menschen einen klar ausmachbaren Geschlechtstypus. Es bestehen geschlechtliche Differenzen in der Funktion des Kehlkopfes, und diese Differenzierung weist darauf hin, dass durch sie die Andersartigkeit des Geschlechts anregend und anziehend wirkt. Das Werden einer Stimme hängt wesentlich ab von den unterschiedlichen Absonderungen von Hormonen. Doch kommt der Geschlechtsdrüse nicht nur die Funktion zu, die Ausbildung der Geschlechtsmerkmale (am Kehlkopf) zu beeinflussen, sondern sie auch im Zustand höchster Ausbildung zu erhalten.

»Da man bei reproduzierenden Musikern und Schauspielern eine ziemlich



Abb. 4

Diejenigen, die am stärksten atmen, haben den besten Geruch. (Empedokles) – Von dem italienischen Arzt Geronimo Mercuriali (1530 bis 1606) empfohlene Atemübungen zur Stärkung der Stimme.

gesteigerte normale Sinnlichkeit, eine (mitunter) maßlose Betätigung des sonst physiologischen Geschlechtstriebes sich entwickeln sieht,« – was bei den Vertretern der Geisteswissenschaften, deren Energiequelle mehr in der Phantasie und im ästhetischen Empfinden wurzelt, verhältnismäßig häufig in sexuellen Abirrungen sich kundgibt –» so ist eine Überreizung der Geschlechtsnerven keine Seltenheit.<sup>5</sup>

In der Devianz, in der Abweichung drückt sich häufig Wesentliches aus. Wie wir mit dem Anderen, mit dem von uns Verschiedenen umgehen, wie wir es klassifizieren, markiert Existenzielles im Umgang mit den Phänomenen. Besetzungen der Wertungen durch Angst und Neid sind naheliegend. Dies zeigt sich auch in der Geschichte des in der Zivilisation bisher drastischsten Eingriffs in den geschlechtlichen Charakter der Stimme, die Kastration. Sie hat über mehr als zwei Jahrtausende sämtliche nur denkbaren gesellschaftlichen Wertungen erfahren: von der Unterwerfung zur Vergötterung, der Verherrlichung über die Inkriminierung bis zur massiven Diskriminierung. Die ersten kastrierten Sänger hießen »Galli = Nachtigallen«, die kastrierten Männer am Hofe des chinesischen Kaisers »Eunuchen«. Trotz vielfachen Verbots war die Kastration noch bis ins 19. Jahrhundert hinein gepflegte Praxis, war »das Messerchen«, mit dem sie betrieben wurde, in Italien noch hochgepriesen. Die Stimme der Kastraten ahmt die der Engel im Himmel nach, hieß eine populäre Phrase. Sie verwies auf das Agentenhafte, das Grenzgängertum der Kastraten zwischen dem Erreichten und dem Unerreichbaren.

Richard Wagner über den Kastraten Cassaroli, der in seinem Elternhaus häufig zusammen mit Carl Maria von Weber verkehrte:

Der italienische Sopransänger, ein ungeheurer, rundbäuchiger Koloss, entsetzte mich durch seine hohe Weiberstimme, seine erstaunliche Volubilität im Sprechen und sein kreischendes, stets bereites Lachen. Trotz seiner großen Gutmütigkeit und Beliebtheit, namentlich auch in meiner Familie, war dieser Mensch mir gespenstisch widerwärtig. Ihn italienisch sprechen und singen zu hören, erschien mir als das Teufelswerk dieser Spukmaschine. [...] Der skandalösen Gestalt Cassarolis gegenüber erfaßte mich Webers überaus zarte, leidende und geistverklärte Erscheinung mit ekstatischer Teilnahme.<sup>6</sup>

#### 4.

Some 2500 years ago Empedocles formulated a theory of perception that was significantly located between two alternative hypotheses. In the case of vision, sim-

ply put, one stated that objects emanated light that shone into the eye. The other concept was somehow the reverse, that the eye functioned like a projector, illuminating the object. The thinking of Empedocles embraced both these notions with the proposition that the eye did indeed engage the object by casting its gaze upon it. But at the same time the object shone back toward the viewer. This meant that the perception of the object was located more or less half way between the seer and the seen, at an interface, an ethereal skin of interference outside the body where two streams of light came together.

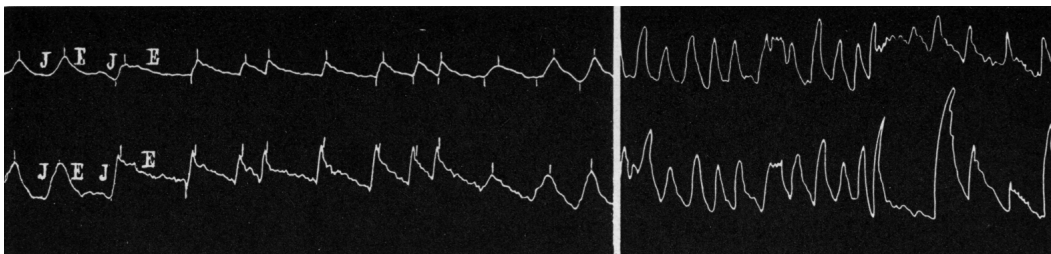
Empedocles' texts are only available to us in fragments.<sup>7</sup> One collection of fragments is called *The senses and the mind* (perhaps ›thinking‹ might be better). The greek poet-philosopher principally describes the sense of ›hearing‹ much as he does the other senses within his *Theory of pores*, and especially he remarks that auditive perception comes from inner noises (Geräusche), when the inner air produces a sound (Klang) by being moved through resonance (Schall). For Empedocles, the system of the ear (das Gehör), which he calls a fleshy twig or branch, is some kind of a bell (ground of resonance) that produces sounds to match those in the outer world. By being moved, the inner air is beating against the bell-like, solid parts and thus produces sound.

This suggests an inversion of function: the idea is partly concerned with two internal, sonic phenomena generated by the ear-brain, tinnitus and oto-acoustical emissions which are measurable sounds emerging from the ear. By addressing the issue of OAEs, we are reversing the normally accepted function of the ear purely as receiver.

Listening inwards is like looking at your own retina. There might be a tendency to think of ourselves as surrounded from without, on all sides information

#### Abb. 5

Sprechatemkurven, links eines gesunden Mädchens, rechts eines Stotterers. Die obere Kurve zeigt die Brust-, die untere die Bauchatmung. Rechts sind die Inspirationsspasmen und die Atemverschiebung deutlich erkennbar. Nach Max Nadoleczny »Die Sprach- und Stimmstörungen im Kindesalter« (1912).





flying at us, colliding with our sensors, meteorites of stuff hitting our skin. But noises are also generated from within, which may include the noise of our thoughts as well as the physical sound of the body at work.

So what is the connection between the sounding space of the world and the chamber or space of perception we tend to locate inside our skin? Sounds emerging from the ear meeting others on the way in, at the interface of the tympanic membrane suggests a model where signals pass through each other, retain their identity or get lost; a complex of waves that in some cases amplify each other and in others, cancel each other out.

In her book *Membranes*, Laura Otis writes, »anatomically, a membrane defines a cell's form; in essence, microscopists knew that cells existed because they could see their boundaries. More important though, the membrane plays a key physiological role.«<sup>8</sup> Discussing Virchow's concept of vitalism from 1858, she remarks

that the fundamental quality of life was simultaneous independence (Selbstständigkeit) and dependence (Abhängigkeit), a quality epitomised by the bounded cells he observed: ›Life‹, (wrote Virchow), ›consists of an exchange, but it would stop being life if this exchange did not have certain limits. These limits establish certain standards of moderation and regulation, as much in the simple cell as in the organism composed of cells. In the cell we have come to know the membrane and the nucleus as the regulators and moderators.‹ The membrane defined the cells not only because it made them visible and set their limits in a physical sense but also because it let them regulate their own inputs and outputs. While these semi-permeable membranes never entirely sealed cells off from their environments, they allowed cells to resist them and to select the molecules that could enter.<sup>9</sup>

5.

Kunst, die durch Medien hindurchgeht, ist Zeitkunst. Geräusche, Musik, Rede, alle Klänge sind Phänomene in der Zeit. Aber welcher Zeit? Indem sie dauern, sind sie Chronos aufs Engste verbunden, den die Griechen erfanden, um der Zeit, die *uns* hat, einen Namen zu geben, jener Zeit, die wir nur sehr bedingt selbst bestimmen können, die unsere Lebensstrecke ausmacht. Chronos ist die erlittene Zeit, die in dem, der seine eigenen Kinder verschlang, ihr drastisches Bild erhielt. Das ist die Zeit der Melancholie, die saturnische Zeit, der Goya eines seiner

schönsten dunklen Bilder gewidmet hat. Wir können chronisch leiden, aber unmöglich chronisch leidenschaftlich sein.

Weil diese Vorstellung von der Zeit alleine unerträglich wurde, ließen die Erfinder der Gottheit sich zwei weitere Zeitfiguren zu Chronos hinzugesellen: Aion und Kairos. Aion – das ist die von uns nicht zu begreifende Zeit, die Ewigkeit, die heilige Zeit. Suspekt ist sie uns nicht zuletzt deshalb, weil sie als reine Zeit gedacht wurde. Kairos ist der Gegenentwurf zu ihr. Kairos ist der Gott des günstigen Augenblicks, einmalig, nicht wiederholbar. Er rhythmisiert die Zeit als Chronos und ist transitorisch. In der Geschichte vom trojanischen Krieger Pandaros, der Menelaos mit einem Pfeil töten soll, bezeichnet Kairos die verwundbare Stelle am Körper des zu Treffenden. Aber Kairos heißt auch jene dreieckige Öffnung, »die beim Weben entsteht, wenn die Kettenfäden gehoben bzw. gesenkt werden, um den Schußfaden oder Einschlag hindurchzulassen.«<sup>10</sup> Diese Öffnung ist sehr klein und im Webprozess verschwindet sie rasch, »um sich sofort wieder aufzubauen.«<sup>11</sup>

Die Stimme hat eine Existenz, die zwischen diesen drei Zeitfiguren schwebt. Indem sie sich artikuliert, wird sie auch chronologisch, in ihrem pneumatischen Hauch schwingt Aion mit. Aber in der Kunst und im Lebendigen hat

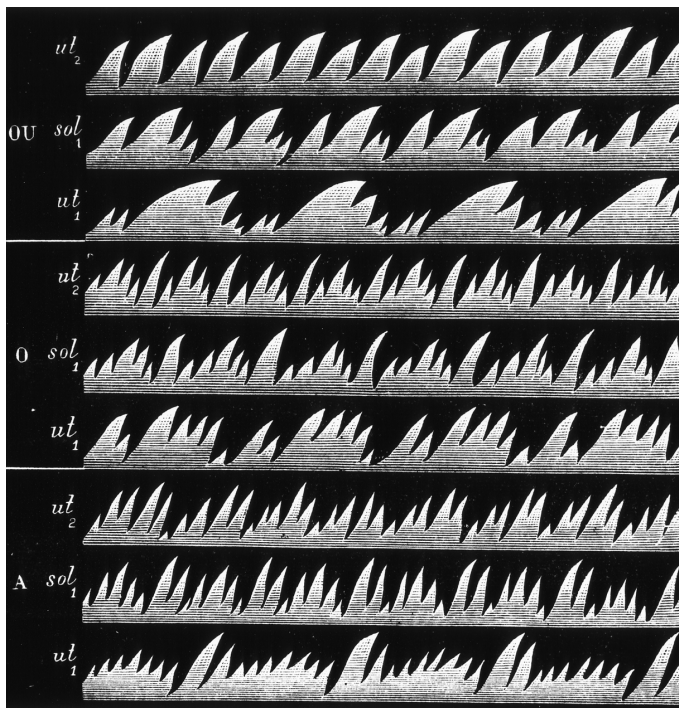


Abb. 6

Fig. 343. Analyse du timbre des différentes voyelles par les flammes manométriques.



Abb. 7

Naturalphabet (1667)

sie ihren stärksten Verbündeten in Kairos, »im gelungenen Wägen auf des Messers Schneide«,<sup>12</sup> als Gelegenheit, die es beim Schopfe zu greifen gilt, als Umschlagstelle, wie die Null, die in den Codices der Mayas in unendlichen Varianten in der Form eines geöffneten Mundes gezeichnet wurde. Wenn die Stimme zum Sprechen anhebt, dann ist dies ein Moment der Entscheidung über Zu- oder Abneigung, über Wahrheit und Lüge, über Lust und Unlust, über Attraktion und Repulsion.

Das Anheben der Stimme zum Sprechen (oder zum Singen), die Art, wie ein Ton eingesetzt wird – das betrifft den aktivsten Teil des Stimmgebrauchs, darin drückt sich die besondere Einmaligkeit und damit auch das Kairotische der Stimme am deutlichsten aus. Beim Ansetzen eines Vokals zum Beispiel wird das Tor des Kehlkopfes regelrecht gesprengt; es verursacht im Inneren ein kurzes explosionsartiges Geräusch. Beim leisen Vokaleinsatz nähert man die Stimmbänder gerade nur so weit einander und spannt sie an, als es zur Bildung des beabsichtigten Stimmtones nötig ist, die Stimmritze wird also nicht fest verschlossen. Der Stimme geht in diesem Falle ein ganz leises Geräusch voraus, das der Prager Physiologe Purkyně 1836 in Breslau erstmals experimentell entdeckte und das von ihm als »leiser Hauch« (*lekki chuch*) bezeichnet wurde.

Anthony Moore / Siegfried Zielinski

292

In der Geschichte der Medien kam die Aufzeichnung der Stimme vor der des sichtbaren beweglichen Körpers (während hingegen der Fotoapparat vor dem Hörgerät erfunden wurde). Der Wunsch nach dem Unsterblichmachen der klingenden Stimme war gewaltig. Anthony Moore sagt: Aufzeichnungsmaschinen wurden erfunden, um die Stimme einzufrieren, und zu keinem anderen Zweck.

6.

With the work *Cellula*,<sup>13</sup> we sought to define in sound, a small, confining space within a larger space and to examine questions of inside/outside, of location and identity, the paradox of skin which consists of that which it contains as well as that which contains it, a drop in the ocean. And whenever I hear this, I always imagine the drop suspended somewhere mid-way between the ocean floor and the surface, somehow existing in the middle of the depths of the sea, separate and identical. And if space contains other spaces how are we aware of them, can an emptiness contain other emptinesses, and what is it that allows them to maintain (in the case of the sound of words for example) their autonomous existence? This idea of differing spaces, made up of the same material, existing inside each other, is to ask what separates, what is the skin of a sound that makes it distinguishable from

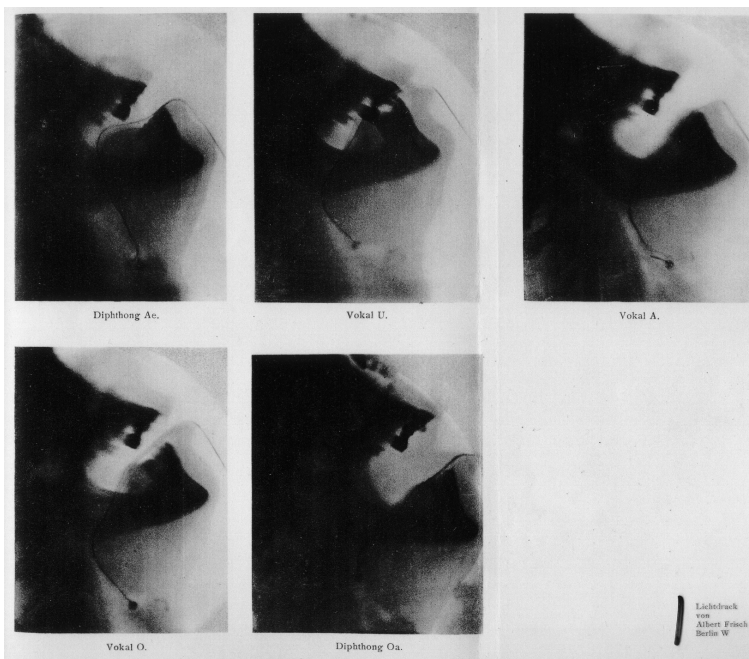


Abb. 8

Archiv für  
Laryngologie

the surrounding soundfield which is more often than not, a cacophonous accompaniment. Why can we understand a conversation against a noisy background, how is it that sounds can retain their individuality, and when do they lose their distinctive qualities and merge and dissolve into each other like paint, forming new colours, greens, where the blues and yellows are no longer distinguishable?

For the very reason that our ears are ever open, even when we sleep, we create sound-screens, psychosomatic ear-lids that close at the point where the signals reach the brain. We have been taught to do this by our bodies. Perhaps, even before birth, we have already learnt to screen out the incessant roar of blood streaming through our veins, by far the loudest sound we could ever hear and yet silent to us. And even if one is unaware, one is still hearing. Just think of that vertiginous feeling which comes with the sudden absence of a continuous sound you had ceased to notice. But if you are aware of listening, then the sound you are listening to will disappear (this, according to some audiologists and a new approach towards a cure for tinnitus based on the idea of the brain turning off anything that goes on and on – »if music be the food of love«). Previously it was thought the more one dwelt on the sounds, the more present and irksome they would become. Now the belief is that it is possible to bring about a kind of psycho-cancellation, as if super-awareness is the penultimate stage towards complete non-awareness, that attention or consciousness doesn't fade in and out but rather switches on and off quite suddenly. This process of habituation is like a time-shift that locks into opposite phase and consequently produces nothing, silence; which forces a kind of forgetting that inescapably asserts itself by the very awareness of its happening. For only the briefest space of time something new makes itself known to us.

The way in which a sound arrives, its envelope-shape, contains as much of its identity as the attendant harmonics or overtones which give us the timbral information needed to distinguish between two instruments sounding the same note. The speed of forgetting is a function of the envelope too and it may be that memories are more likely to arrive suddenly than forgetfulness, which may fade. One switches and is switched between states of awareness and unawareness. The shape of the envelope affects the perception of the state-change. If you have a gradual attack and a slow release then you are not exactly sure when you ceased to be aware.

7.

In Walter Benjamins Hörstück *Lichtenberg*<sup>14</sup> verfügt das auf dem Mond angesiedelte Komitee für Erdforschung über drei Apparate: Mit dem *Spectrophon* kann al-

Anthony Moore / Siegfried Zielinski

294

les, was auf der Erde passiert, gesehen und gehört werden; mit dem *Oneiroskop* können die Mondbewohner die Träume der Irdischen beobachten; mit dem *Parlaronium* können die akustisch durch die Sphärenmusik sehr verwöhnten Mondbewohner die oft lästige Menschenrede in wohlklingende Musik übersetzen.

An den extremen Polen erweisen sich die Unterscheidungen zwischen Fern- und Nahsinn als problematisch. Vilém Flusser, der von sich selbst sagte, dass er fast taub war und sein Hörgerät dazu benutzte, sich in die Welt einzuschalten oder sich aus ihr auszuschalten, spielte in einem unveröffentlichten Manuskript mit der seltsamen Beziehung zwischen dem Hören und dem Sehen und ihrer beider Grenzen:

Glauben Sie bitte nicht, dass ich etwa besser als Sie sähe, nur, weil ich schlechter höre. Das ist ein weit verbreiteter und verhängnisvoller Irrtum. Etwa zu meinen, die Griechen hätten schlechter gehört als die Juden, weil sie weiter gesehen hätten, und die Juden hätten besser gehört als die Griechen, weil sie weniger weit sahen. Oder Homer sei blind gewesen, weil er die Stimme der Muse hörte, und dass daher anzunehmen sei, die großen Seher der Menschheit seien taub gewesen. Eine derartige Kompensation von Auge und Ohr, von ›Theorie‹ und ›Schema Israel‹ ist ein Irrtum. Im Gegenteil: das Auge und das Ohr verstärken einander gegenseitig. [...] Je besser man die Stimme des Daimonion hört, desto besser kann man theoretisieren. Die griechischen Theorien waren kurzsichtig,

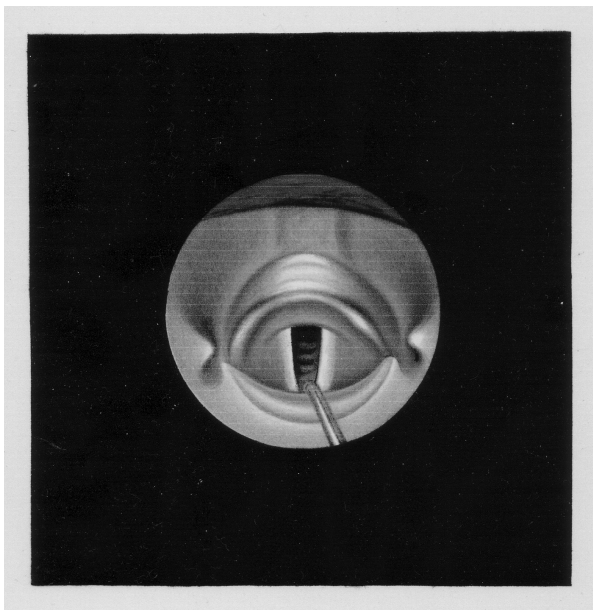


Abb. 9

Archiv für Laryngologie

weil die Griechen schwerhörig waren. Die Stimme, die in den jüdischen Schriften zu Wort kommt, reichte nicht weit, weil die Juden kurzsichtig waren. Meine Taubheit lässt mich weniger sehen. Erwarten sie also nicht, dass ich Ihnen etwa tiefe Einsichten ins Hören bieten kann.<sup>15</sup>

Das Gehör hat eminent politischen Charakter. Die Phrase von den politischen Ansichten müsste nach Flusser zumindest ergänzt werden durch die von der politischen Hörigkeit. Durch das Gehör sind wir ständig in Gesellschaft eingetaucht, unfreiwillig. *Silence is sexy*, nannte Blixa Bargeld die 2000er CD der Restgruppe der *Einstürzenden Neubauten*. Stille ist Luxus, Verschwendung, die höchste Form des Genusses. Darum können Politiker keine Stille dulden. Sie ist privat. Politik ist lärmendes Engagement. Deshalb hat sie die Abstimmung erfunden. Sie will das Konzert und nicht den Konsensus.

1952 veröffentlichte Gottfried Benn drei kurze dramatische Texte mit dem Titel: *Die Stimme hinter dem Vorhang*.<sup>16</sup> Die Stimme aus dem *off* hat durch die bizarren Texte hindurch die Rolle der Provokateurin, der Mahnerin, der Verunsicherung. »Was soll denn sein?«, diese drängende Frage schiebt sie immer wieder zwischen die Äußerungen der Protagonisten, selbsternannte Experten des Alltags allesamt. Gegen Ende des zweiten Textes, *Die Sonntagszeitung*, lässt sie Gottfried Benn einmal aus sich herausgehen, was so viel heißt wie, dass sie nicht mehr bei sich ist. Die Stimme brüllt:

Was soll denn sein, wer seid ihr denn, geschaffen und geworden, die Augen unter Tränen und die Herzen hin und her? Schluckt doch endlich euer

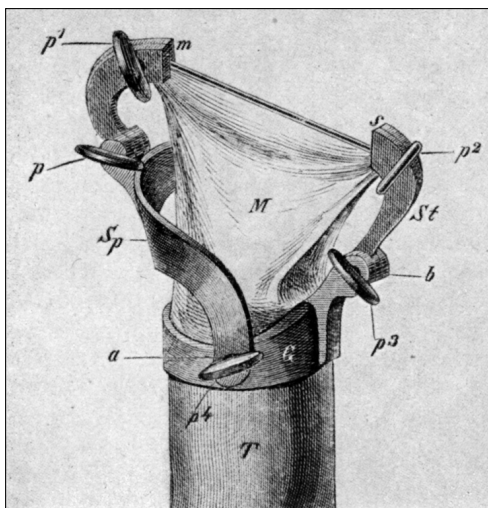


Abb. 10

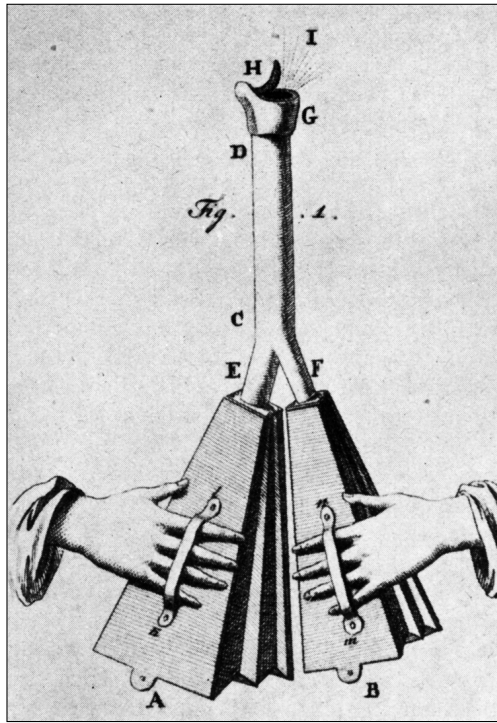
Künstlicher Kehlkopf, wie ihn der Wiener Laryngologe Johann Nepomuk Czermak (1828–1873) zu Demonstrationszwecken verwendete.

Anthony Moore / Siegfried Zielinski

296

Abb. 11

Modell zum Vergleich der Stimmbildung mit der Tonerzeugung in einer Zungenpfeife. Das Pfeifenrohr entspricht der Luftröhre, die Zunge den Stimmlippen, die damals als häutige Gebilde aufgefasst wurden; die Funktion der Lunge ist durch Blasebälge angedeutet.



Inneres und eure Beine und haltet die Schnauze über eure Kaldaunen oder wie der Psalmist sagt: meine Seele ist stille zu Gott. Was soll denn sein, ihr nennt schon heilig, was ich nur doof nenne, ihr billiger Krimskrams mit Gemüts- und Blasenpanik – immer schnell aufs Töpfchen! [...] Betreibt nur weiter euren Ausverkauf – [...] Alles Freiplatzschnorrer, wiehern auf der Galerie, wenn in der Szene die Herzen sterben – was soll denn sein, wieviel Nächte habt ihr denn alleine durchgestanden, welche Trauer ohne Geschwätz getragen – das Nichts, euer Nichts war immer noch durchklimpert von Gebetsmühlen und Schuhschnallen – wenn ich euch nun sagte: im Dunkel leben, im Dunkel tun, was wir können – das soll sein?<sup>17</sup>

8.

There is this model of the listening self isolated at the single point of the centre. How to achieve the necessary escape velocity to float out onto the edges of your soundfield? Singing with someone else is the answer because unlike particles, the coherence of waves allows for a simultaneous and separate existence. Waves can



co-exist in the same place at the same time. You won't displace them. If you climb into a bath of sound waves they won't spill over onto the floor. Just as with the simultaneity of harmonic or inharmonic spectra, occupying the same place at the same time allows us to escape from the sublunary solitude of being fixed at the core of our hearing. The howl for air (or the art of slapping newborns). The howl for air, the expanding sphere, the first feedback loop, the yell to the ears. If we can accept the anthropomorphic notion of a sound that listens to itself, a sound with ears, coupled with the model of the negentropic soliton, then that first sound you made and heard, the howl for air, becomes an expanding bubble inside which all other proceeding sounds, heard or made, continue to be contained. The permeable and elastic membrane that surrounds the expanding, spherical after-life of a sound is pervious to carrier waves of information flowing in. These waves are themselves other sounds moving out from their sources and passing through each other. This is simultaneous autonomy and permeability. The limitation might be time, an earlier sound may not be overtaken by a later one. But if all sounds are contained inside each other, by those that came before them, and they all travel at the same speed, and a more recent sound fades out before an earlier one, then what happens is a contained dispersal, a more entropic event within a lesser one. So fading is the dissolution of the skin, perhaps the most important organ we have.

There is a transcript from Charles Babbage called *The Ninth Bridgewater Treatise* from 1837. We end our text with a fragment out of this fragment.

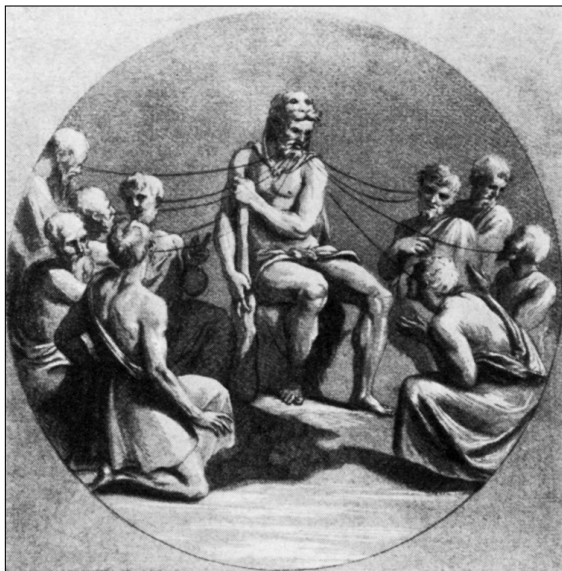


Abb. 12

Gallischer Herkules. Allegorie auf die Beredsamkeit. Clairobscur nach einer Zeichnung von Raffael, in Kupfer gestochen von Charles-Nicolas Cochin d.J. (1715–1790), und in Holz geschnitten von Vincent Le Sueur (1668–1743).

The principle of the equality of action and reaction, when traced through all its consequences, opens views which will appear to many persons, most unexpected. The pulsations of the air, once set in motion by the human voice, cease not to exist with the sounds to which they gave rise. [...] Thus considered, what a strange chaos is this wide atmosphere we breathe! Every atom, impressed with good and with ill, retains at once the motions which philosophers and sages have imparted to it, mixed and combined in ten thousand ways with all that is worthless and base. The air itself is one vast library, on whose pages are forever written all that man has ever said or woman whispered.<sup>18</sup>

- 1 Die Autoren nennen das Format, das sie zusammen entwickelt und an verschiedenen Themen experimentell erprobt haben, *expanded lecture*. Neben den Textelementen, die hier wiedergegeben werden können, bestand dieser Vortrag aus zehn Tonfragmenten, die während des 3-tägigen Festivals *PerSon* 1999 in der Trinitatiskirche in Köln aufgenommen wurden. Das waren stimmliche Performanzen u. a. von John Berger, Sainkho Namchylak, Nicole Tibbels und Raiili Viljakainen sowie ein Ausschnitt aus einer Kontrabass-Performance von Peter Kowald. Von den visuellen Elementen, die parallel zu den Musikfragmenten montiert wurden, ist hier nur eine kleine Auswahl enthalten.
- 2 Paul Grützner: *Physiologie der Stimme und Sprache*. Handbuch der Bewegungsapparate, Zweiter Theil, Leipzig 1879, S. 108.
- 3 Carl Ludwig Schleich: *Vom Schaltwerk der Gedanken*. Neue Einsichten und Betrachtungen über die Seele, Berlin 1916, S. 31.
- 4 Friedrich Leipoldt: *Stimme und Sexualität*. Das Problem der inneren Zusammenhänge von Stimme und Sexualität in gemeinverständlicher Form dargestellt, Leipzig 1926, S. 15 f.
- 5 Leipoldt: *Stimme* (Anm. 4), S. 60.
- 6 Ebd., S. 31.
- 7 Empedokles-Fragmente, in: Jaap Mansfeld (Hg.): *Die Vorsokratiker*. Band 2, Stuttgart 1986.
- 8 Laura Otis: *Membranes*. Metaphors of Invasion in Nineteenth-Century Literature, Science and Politics, Baltimore 2000, S. 23.
- 9 Ebd., S. 23.
- 10 Manfred Kerkhoff: Zum antiken Begriff des Kairos, in: *Zeitschrift für philosophische Forschung* 27 (1973), S. 259
- 11 Ebd., S. 259. Kerkhoff bezieht sich hier wiederum auf das berühmte Buch R. B. Onians zu den Ursprüngen des europäischen Denkens: *Richard Broxton Onians: The Origins of European Thought about the Body, the Mind, the Soul, the World, Time, and Fate*. New Interpretations of Greek, Roman and Kindred Evidence, Cambridge 1954.
- 12 Kerkhoff: *Kairos* (Anm. 10), S. 259.
- 13 This was a work conceived in 1998 for the cavernous space of the old railway station in Firenze. Rather than filling the space with sound the idea was to create a small and invisible cell in the centre. At this time ›soundbeam‹ technology was in its infancy but we were already thinking about edges and borders of sound.
- 14 Walter Benjamin: *Lichtenberg*. Ein Querschnitt, in: ders.: *Drei Hörmodelle*, Frankfurt/M. 1971.
- 15 Das unveröffentlichte Manuskript befindet sich im Vilém-Flusser-Archiv an der Kunsthochschule für Medien Köln. Vielen Dank an dessen Leiterin Silvia Wagnermaier.
- 16 Gottfried Benn: *Die Stimme hinter dem Vorhang* [1952], Wiesbaden 1955.
- 17 Ebd., S. 41 f.
- 18 Charles Babbage: *The Ninth Bridgewater Treatise*. A Fragment, London 1837. Vielen Dank an Friedrich Kittler.

## BILDNACHWEISE

### CORNELIA EPPING-JÄGER: »EINE EINZIGE JUBELNDE STIMME«

- Abb. 1 Heinrich Hoffmann (Hg.): Hitler über Deutschland, Text v. Josef Berchtold, München 1932, S. 69.  
Abb. 2 Ebd., S. 60.  
Abb. 3 Hans Wendt: Der Tag der Nationalen Arbeit. Die Feier des 1. Mai 1933«, Berlin 1933, zwischen S. 26 und S. 27.  
Abb. 4 Ebd., zwischen S. 18 und S. 19.  
Abb. 5 Rudolf Herz: Hoffmann & Hitler. Fotografie als Medium des Führer-Mythos, München 1994, S. 234.  
Abb. 6 Ebd., S. 224.

### MICHEL CHION: MABUSE – MAGIE UND KRÄFTE DES »ACOUSMÈTRE«

- Abb. 1 Michel Chion: The Voice in Cinema, edited and translated by Claudia Gorbman, New York 1999, S. 16.  
Abb. 2 Ebd., S. 20.  
Abb. 3a/3b Ebd., S. 48.  
Abb. 4 Michel Chion: La voix au cinéma, Paris 1982, S. 37.  
Abb. 5 Ebd., S. 41.  
Abb. 6 Ebd., S. 42.

### INES STEINER: THE VOICE BEHIND THE CURTAIN

- Abb. 1–16 Reproduziert aus der DVD des MGM-Musicals SINGIN' IN THE RAIN (Stanley Donen / Gene Kelly, 1952).

### DANIEL GETHMANN: DAS ZITTERN DER LUFT

- Abb. 1 Denis Dodart: Mémoire sur les causes de la voix de l'homme, et de ses différens tons, in: Mémoires de l'Académie royale des sciences (1700), S. 238–287.  
Abb. 2 Wolfgang von Kempelen: Mechanismus der menschlichen Sprache nebst der Beschreibung seiner sprechenden Maschine [1791], Stuttgart-Bad Cannstadt 1970, S. 80.  
Abb. 3 Christian Gottlieb Kratzenstein: Sur la naissance & la formation des voyelles, in: Journal de Physique, Bd. 21 (1782), S. 358–380.  
Abb. 4 Alfred Chapuis & Édouard Gélis: Le monde des automates, Bd. 2, Paris, S. 205.

## Bildnachweise

300

- Abb. 5 Kempelen: Mechanismus der menschlichen Sprache (Abb. 1), S. 400.  
 Abb. 6 Ebd., S. 439.  
 Abb. 7 The Journal of the Acoustical Society of America, Vol. 22, Nr. 2 (1950), S. 164.  
 Abb. 8 Hermann Gutzmann: Physiologie der Stimme und Sprache. Braunschweig 1909, S. 70.  
 Abb. 9 Illustrated London News, Jg. 9 (1846), S. 96.  
 Abb. 10 La Nature, 1908, S. 7.

## JIN HYUN KIM: DIE SINGSTIMME ALS AUSDRUCKSZEICHEN

- Abb. 1 Johan Sundberg: Die Singstimme, in: Klaus Winkler: Die Physik der Musikinstrumente, Heidelberg 1992, S. 14–21 (hier: S. 15).  
 Abb. 2 Manfred Clynes: Sentic. The Touch of Emotions, London 1977, S. 29.  
 Abb. 3–5 Eliezer Rapoport: Emotional Expression Code in Opera and Lied Singing, in: Journal of New Music Research 25/2 (1996), S. 109–149 (hier: S. 116).  
 Abb. 6–11/  
     13-1 Ergebnisse eigener Analyse  
 Abb. 12 Luciano Berio: sequenza III per voce femminile [1966], London: universal edition 1968, S. 1.  
 Abb. 13-2 Ebd., S. 2.

## ANTHONY MOORE / SIEGFRIED ZIELINSKI: VOX PER → SONARE

- Abb. 1–2 Robert Willis 1829, entnommen aus der Zeitschrift *Nature*, 1908.  
 Abb. 3 Johannes Müller: Über die Compensation der physischen Kräfte am menschlichen Stimmorgan, Berlin 1839.  
 Abb. 4 Mercuriali: De arte gymnastica libri sex, 2. Aufl., Venedig 1573.  
 Abb. 5 Handbuch der Kinderheilkunde, Bd. VI, Leipzig 1912.  
 Abb. 6 Aus Karl Rudolf Königs (1832–1901) Analyse des Timbres verschiedener Vokale durch manometrische Flammen, Étienne-Jules Marey: Le mouvement, Paris 1894.  
 Abb. 6 Aus Karl Rudolf Königs (1832–1901) Analyse des Timbres verschiedener Vokale durch manometrische Flammen, Étienne-Jules Marey: Le mouvement, Paris 1894.  
 Abb. 7 F. M. B. Hellmont: Kurtzer Entwurf des eigentlichen Naturalphabets der heiligen Sprache, Abraham Lichtenthaler: Sultzbach 1667.  
 Abb. 8 Archiv für Laryngologie. Bd. XIX, Tafel 16, Lichtdruck von Albert Frisch, Berlin.  
 Abb. 9 Archiv für Laryngologie. Bd. XLX, Taf. XIII, Fig. 5.  
 Abb. 10 Johann Nepomuk Czermak: Gesammelten Schriften Bd. 2, Leipzig 1879.  
 Abb. 11 W. v. Kempelen (1734 bis 1804): Mechanismus der menschlichen Sprache, Wien 1791.

## AUTORENVERZEICHNIS

**Heike Behrend** ist Professorin am Institut für Afrikanistik der Universität zu Köln und Teilprojektleiterin am Forschungskolleg (SFB/FK 427) *Medien und Kulturelle Kommunikation*. Ihre Arbeitsschwerpunkte sind Religionsethnologie und Medienanthropologie. Letzte Veröffentlichungen: »Flugapparate ins Jenseits«. Geisterarchive im interkulturellen Vergleich, in: Hedwig Pompe/Leander Scholz (Hg.): Archivprozesse. Die Kommunikation der Aufbewahrung, Köln 2002; »Call and Kill«: Zur Verzauberung und Entzauberung westlicher technischer Medien in Afrika, in: Erhard Schüttpelz/Albert Kümmel: Signale der Störung, München 2003.

**Gereon Blaseio, M. A.**, ist Wissenschaftlicher Mitarbeiter an der Universität zu Köln. Er arbeitet in einem Forschungsprojekt zur »Synchronisation in gender-spezifisierter Perspektive« unter der Leitung von Claudia Liebrand. Letzte Veröffentlichungen: Genre und Gender: Zur Interdependenz zweier Leitkonzepte der Filmwissenschaft und: Heaven and earth. Vietnam Film zwischen Male Melodrama und Women's Film, beide in: Claudia Liebrand/Ines Steiner (Hg.): Hollywood hybrid. Genre und Gender im Mainstream-Film, Marburg (im Druck).

**Michel Chion**, ist professeur associé an der Universität Paris III und unterrichtet an den Filmhochschulen ESEC in Paris und DAVI in Lausanne. Er ist zudem Komponist zahlreicher Werke konkreter Musik, Musik- Klang- und Filmtheoretiker. Letzte Veröffentlichungen: La musique au cinéma, Paris 1995; Le Son, Paris 1998 und La comédie musicale, Paris 2003.

**Michael Cuntz**, Dr. phil., Romanist, wissenschaftlicher Mitarbeiter am kulturwissenschaftlichen Forschungskolleg (SFB/FK 427) *Medien und Kulturelle Kommunikation* an der Universität zu Köln. Arbeitsschwerpunkte: Text-Bild-Verhältnisse bei Dante, Blaise Pascal/Moralistik, Literatur der Moderne und Postmoderne, Medientheorie. In Kürze erscheint: Der göttliche Autor. Zu Prophetie, Apologie und Simulation in Texten Pascals, Stuttgart 2003.

**Cornelia Epping-Jäger**, Dr. phil., ist Wissenschaftliche Mitarbeiterin im Projekt »Laut/Sprecher: Mediendiskurse und Medienpraxen in der Zeit des Nationalsozialismus« am Forschungskolleg (SFB/FK 427) *Medien und Kulturelle Kommunikation* an der Universität zu Köln. Arbeitsschwerpunkte: Mediengeschichte und Medientheorie. Letzte Veröffentlichungen: Laut/Sprecher Hitler. Über ein Dispositiv der Massenkommunikation, Marburg 2003.

nikation der Zeit des Nationalsozialismus, in: Josef Kopperschmidt (Hg.): *Hitler der Redner*, München 2003 (im Druck); *Der Literalisierungsprozess und die Ordnung des Wissens: der deutsche Belial*, in: Ludwig Jäger/Erika Linz (Hg.): *Medialität und Mentalität. Theoretische und empirische Studien zum Verhältnis von Sprache, Subjektivität und Kognition*, München 2003 (im Druck).

**Marcus Erbe**, M. A., ist Wissenschaftlicher Mitarbeiter an der Universität zu Köln, wo er für die Bearbeitung des musikwissenschaftlichen Einzelprojektes »Graphische Transkriptionen Elektronischer Musik« im Rahmen des Forschungskollegs (SFB/FK 427) *Medien und Kulturelle Kommunikation* verantwortlich zeichnet. Derzeitiger Arbeitsschwerpunkt: Musik der Gegenwart unter besonderer Berücksichtigung elektroakustischer Kompositionen.

In Kürze erscheinen: *The Transcription of Electronic Music in The Crying of Lot 49*, in: Pynchon Notes, Sonderausgabe *Site-Specific: From Aachen to Zwölfkinder – Pynchon: Germany* (im Druck); »The hills are alive with the sound of music«: *Der Klangraum in MOULIN ROUGE*, in: Claudia Liebrand/Ines Steiner (Hg.): *Hollywood hybrid. Genre und Gender im zeitgenössischen Mainstream-Film*, Marburg 2003 (im Druck); *Karlheinz Stockhausens TELEMUSIK*, in: Christoph von Blumröder/Imke Misch (Hg.): *Kompositorische Stationen des 20. Jahrhunderts (= Signale aus Köln VII)*, Münster 2003 (im Druck).

**Daniel Gethmann**, Dr. phil., arbeitet am Institut für Medienwissenschaft der Ruhr-Universität Bochum. Arbeitsschwerpunkte: *Auditive Kultur*, *NS-Film*, *Vor- und Frühgeschichte der Bewegtbildmedien*.

Letzte Veröffentlichung: *Zwischen Optophonie und Phonovision. Die technische und künstlerische Synthese von Ton- und Bildspeicherung als Vorgeschichte der Videokunst*, in: Ralf Adelman/Hilde Hoffmann/Rolf Nohr (Hg.): *REC – Video als mediales Phänomen*, Weimar 2002, S. 147–164.

**Jin Hyun Kim**, M. A., ist wissenschaftliche Mitarbeiterin im kulturwissenschaftlichen Forschungskolleg (SFB/FK 427) *Medien und kulturelle Kommunikation* an der Universität zu Köln. Arbeitsschwerpunkte: *Systematische Musikwissenschaft* (insbesondere *Musikästhetik*, *Musikpsychologie* und *Musikinformatik*), *Computermusik* und *Interaktive Medienkunst*.

In Kürze erscheinen: *Musikwissenschaft in der Postmoderne. Zur Legitimationsproblematik von Musikwissenschaft*, Osnabrück: 2003 (in Vorbereitung); *Transkriptive Interaktion*, in: Christa Brüstle/Albrecht Riethmüller (Hg.): *Klang und Bewegung. Beiträge zu einer Grundkonstellation*, Aachen 2003 (im Druck).

**Doris Kolesch** ist Professorin für Theaterwissenschaft an der FU Berlin und Mitglied der Jungen Akademie an der Berlin-Brandenburgischen Akademie der Wissenschaften. Ihre Arbeitsschwerpunkte liegen im Bereich der Theorie und Ästhetik von Theater und Literatur sowie der Kulturgeschichte von Theater und Theatralität. Seit Januar 2002 leitet sie ein Forschungsprojekt zu »Stimmen als Paradigmen des Performativen« im Rahmen des interdisziplinären Sonderforschungsbereichs »Kulturen des Performativen«.

Letzte Veröffentlichungen: Roland Barthes, Frankfurt/M./New York 1997, Mitherausgeberin von: Räume der literarischen Postmoderne. Gender, Performativität, Globalisierung, Tübingen 2000; Ästhetik der Präsenz: Theater-Stimmen; in: Josef Früchtl/Jörg Zimmermann (Hg.): Ästhetik der Inszenierung, Frankfurt/M. 2001, S. 260–276; Roberts Wilsons *Dantons Tod*: Das nomadische Auge und das Archiv der Geschichte, in: Jutta Eming/Annette Jael Lehmann/Irmgard Maassen: Mediale Performanzen. Historische Konzepte und Perspektiven, Freiburg 2002, S. 237–250.

**Sybille Krämer** ist Professorin für Philosophie an der FU Berlin. Arbeitsschwerpunkte: Philosophie und Mathematik in der frühen Neuzeit; Theorie des Geistes und des Bewusstseins; Zeichen- und Medientheorie; Sprachphilosophie; Grundlagenprobleme der Kulturwissenschaften.

Letzte Veröffentlichungen: (Hg.): Medien Computer Realität. Wirklichkeitsvorstellungen und Neue Medien, Frankfurt/M. 1998; Sprache, Sprechakt, Kommunikation, Frankfurt/M. 2001; mit Ekkehard König (Hg.): Gibt es eine Sprache hinter dem Sprechen? Frankfurt/M. 2002.

**Erika Linz**, Dr. phil., ist wissenschaftliche Mitarbeiterin am Kulturwissenschaftlichen Forschungskolleg (SFB/FK 427) *Medien und kulturelle Kommunikation* der Universitäten Köln, Bonn und Aachen. Arbeitsschwerpunkte sind Kognitive Linguistik, linguistische Medien- und Zeichentheorie.

Letzte Veröffentlichungen: Indiskrete Semantik. Kognitive Linguistik und neurowissenschaftliche Theoriebildung, München 2002; mit Klaudia Grote: The influence of sign language iconicity on semantic conceptualization, in: Wolfgang G. Müller/Olga Fischer (Hg.): From sign to Signing, Amsterdam/Philadelphia 2003; mit Ludwig Jäger (Hg.): Medialität und Mentalität. Theoretische und empirische Studien zum Verhältnis von Sprache, Subjektivität und Kognition, München 2003 (im Druck).

**Leander Scholz**, Dr. phil., ist Schriftsteller und wissenschaftlicher Mitarbeiter am Kulturwissenschaftlichen Forschungskolleg (SFB/FK 427) *Medien und kulturelle Kommunikation* der Universitäten Köln, Bonn und Aachen. Derzeitige Arbeitsschwerpunkte: Philosophische Ästhetik und Medientheorie.

Letzte Veröffentlichungen: Das Archiv der Klugheit. Strategien des Wissens um 1700, Tübingen 2002; mit Hedwig Pompe (Hg.): Archivprozesse: Die Kommunikation der Aufbewahrung, Köln 2002; mit Albert Kümmel/Eckhart Schumacher (Hg.): Einführung in die Geschichte der Medien, München 2003.

**Anthony Moore** ist seit 1993 Professor für »Auditive Gestaltung in den Medien« und seit dem Jahr 2000 Rektor der Kunsthochschule für Medien in Köln. Seinen Arbeitsschwerpunkt bilden digitale Aufnahmeverfahren (Signal Processing und MIDI-Sequencing). Seit 1969 komponierte er zahlreiche Filmmusiken für europäische Independent Filme, von denen viele internationale Auszeichnungen erhielten.

Letzte Veröffentlichungen: Hommage à Pink Floyd, in: Friedrich Kittler u. a. (Hg.): Zwischen Rauschen und Offenbarung. Zur Kultur und Mediengeschichte der Stimme, Berlin 2002, S. 371–377.

**Ines Steiner**, M. A., Film- und Kulturwissenschaftlerin; derzeit wiss. Mitarbeiterin am Forschungskolleg (SFB/FK 427) *Medien und kulturelle Kommunikation* mit einem Forschungsprojekt zum Thema: »Zum Verhältnis von Gender und Genre in slapstick-, romantic-, musical- und screwball-comedy.«

Letzte Veröffentlichung: »Bist kein Held, nur ein Mann, der gefällt«. Gender-Performanzen in Bel Ami, in: Armin Loacker (Hg.): Willi Forst. Ein Filmstil aus Wien, Wien 2003, S. 290–320.

**Bernhard Waldenfels** ist Professor Emeritus für Philosophie an der Ruhr-Universität Bochum. Arbeitsschwerpunkte: Phänomenologie und neuere französische Philosophie; Leiblichkeit, Fremdheit, Responsivität, pathische Erfahrung, Phänomenotechnik.

Letzte Veröffentlichungen: Das leibliche Selbst, Frankfurt/M. 2000; Verfremdung der Moderne, Göttingen 2001; Bruchlinien der Erfahrung, Frankfurt/M. 2002.

**Siegfried Zielinski** ist Professor an der Kunsthochschule für Medien Köln mit dem Lehr- und Forschungsschwerpunkt Archäologie der Medien.

Letzte Veröffentlichungen: Audiovisions – Cinema and Television as Entr'actes in History, Amsterdam 1999 und Archäologie der Medien – Zur Tiefenzeit des technischen Hörens und Sehens, Reinbek 2002.