

Musik und Sprache. Überlegungen zu ihrer Struktur und Funktionsweise

Manfred Bierwisch

Die unmittelbare Wirklichkeit des Gedankens ist die Sprache.

Karl Marx

*Die Logik der Kunst ist, paradox nach den Regeln der anderen,
ein Schlußverfahren ohne Begriff und Urteil.*

Theodor W. Adorno

Was gezeigt werden kann, kann nicht gesagt werden.

Ludwig Wittgenstein

1. Einleitende Bemerkungen*

Sprache und Musik in ihrer Wirkungsweise und in ihrem Aufbau etwas genauer zu verstehen, indem man sie miteinander vergleicht: das ist das Ziel der Überlegungen, die ich im folgenden anstellen will. Gesichtspunkte für diesen Vergleich entnehme ich vor allem der modernen Sprachwissenschaft. Das bedeutet allerdings keineswegs, daß ich, wie dies gelegentlich geschieht, die Musik als eine Sprache besonderer Art betrachten und nun untersuchen will, was mit sprachwissenschaftlichen Mitteln über sie herauszufinden ist. Ich werde vielmehr neben den charakteristischen Gemeinsamkeiten auch die entscheidenden Unterschiede zum Vorschein zu bringen versuchen. Um dabei möglichst deutlich zu sein, will ich unter „Sprache“ stets die natürliche, verbale Sprache in ihrer gesprochenen (und nur sekundär in ihrer aufgeschriebenen) Form verstehen und abgeleitete oder metaphorische Verwendungen wie Sprache der Blumen, Sprache der Bienen oder Sprache der Mimik und Gestik vermeiden. Genauer: Ich werde Erscheinungen der letztgenannten Art unter den allgemeineren Begriff der Zeichensysteme fassen, unter den dann die natürliche Sprache als ein ausgezeichneter Spezialfall gehört.¹ Der Vergleich von Sprache und Musik wird damit vor allem zur Charakterisierung verschiedener Arten von Zeichensystemen und ihrer Funktionsweise führen. Als entscheidende Differenz wird sich dabei die Unterscheidung zwischen begrifflicher und anschaulicher Mitteilung, kurz, zwischen Sagen und Zeigen herausstellen.

Zwei Hinweise noch vorab. Erstens: Ich werde mich bei der Erörterung und Illustration konkreter Erscheinungen vornehmlich auf die deutsche

Sprache und die europäische Musiktradition beziehen. Für die sprachliche Seite bedeutet das keine Einschränkung, da die hier interessierenden Probleme am Deutschen so gut wie an jeder anderen Sprache belegt werden können. Die zweite Einschränkung ist problematischer, da andere Musikulturen und die zahlreichen Befunde der Musikethnologie nicht ohne weiteres aus dem Blickwinkel der europäischen Kunstmusik interpretiert werden dürfen. Ich bin zu dieser Einschränkung außer durch Mangel an entsprechenden Detailkenntnissen auch durch die Vielzahl der andernfalls zu berücksichtigenden Sonderprobleme gezwungen. Ich halte sie aber aus zwei Gründen für gerechtfertigt. Zum einen sind in der europäischen Musiktradition die musikalischen Struktur- und Funktionsprinzipien unstrittig zu mindestens dem gleichen Komplexitätsgrad entfaltet wie in irgendeiner anderen Musikkultur, so daß sich hier kaum unangemessene Verkürzungen des Problems ergeben dürften. Zum anderen treten die zahlreichen Erscheinungen, die sich durch die Verwendung musikalischer Mittel in kultischen, militärischen, juristischen und anderen sozialen Zeremoniellen (vgl. dazu etwa Stockmann) sowie im Arbeitsprozeß ergeben, in der europäischen Musiktradition zwar in den Hintergrund. Das beeinträchtigt aber die Erörterung unseres Themas nicht, denn diese Erscheinungen müssen gerade durch das Zusammenwirken musikalischer und anderer Prozesse erfaßt werden. Anders ausgedrückt: Der Versuch, Struktur und Funktionsweise der Musik zu charakterisieren, muß zwar diese Möglichkeit der angedeuteten Wechselwirkungen im Auge behalten, darf die Musik aber nicht darauf reduzieren. Andernfalls würde gerade das Musikalische an der Musik verfehlt, so wie man die Funktionsweise der Sprache nicht darauf zurückführen kann, daß man Wörter und Sätze auch singen kann, oder daß sie zu militärischen Kommandos fixiert werden können. Kurz: Das, worauf es mir ankommt, ist im wesentlichen ohne Verzerrung in der europäischen Kunstmusik zu belegen, wenn die damit gemachte Einschränkung nur als solche berücksichtigt wird.

Und zweitens: Ich werde den Versuch, Musik und Sprache als Zeichensysteme verschiedener Art miteinander zu vergleichen, so weit wie möglich verfolgen. Damit unterstelle ich jedoch nicht, daß sie sich durch den Rückgriff auf ihren Zeichencharakter in gleicher Weise erschließen lassen. Legt man einigermaßen verbindlich einzuhaltende Bestimmungen zu Grunde, läßt sich vielmehr zeigen, welche Momente musikalischer Wirkungsweisen sich dem Bereich der Zeichenfunktion entziehen.

2. Musik, Sprache, Literatur

Ein erster Punkt meiner Überlegungen gilt der ebenso offensichtlichen wie folgenreichen Tatsache, daß es üblich ist, zwischen Sprache und Literatur zu unterscheiden, einen parallelen Unterschied im Bereich der Musik aber nicht zu machen. Diese Asymmetrie findet sich wieder in der Gliederung der Wis-

senschaftsdisziplinen, die sich mit Sprache und Musik befassen. Zwar ist die definitive Trennung von Sprach- und Literaturwissenschaft noch gar nicht so alt und die Grenzziehung keineswegs in jeder Hinsicht selbstverständlich. Ihre Motivation aber ist hinreichend deutlich: Es werden, grob gesprochen, einerseits die Eigenschaften des Mediums Sprache, andererseits die Bedingungen seiner ästhetischen Funktion oder Verwendung zum Gegenstand der Betrachtung gemacht. Die Musikwissenschaft dagegen untersucht ihren Gegenstandsbereich, ohne die beiden Gesichtspunkte auf separate Disziplinen zu verteilen.

Fragt man sich, was es mit der ungleichen Gliederung der beiden Gebiete auf sich hat, so stößt man auf eine Reihe von Gründen, die offenbar in den Sachen selbst liegen. Ein erster Grund liegt in der unterschiedlichen Vielfalt der Verwendungszwecke oder Funktionen von Sprache und Musik. Die Sprache dient den verschiedensten Kommunikations- und Erkenntnisprozessen. Praktisch kein Bereich des menschlichen Lebens kommt ohne sie aus. Auch bei einem weitgefaßten Literaturverständnis ist nur ein kleiner Teil dieser Prozesse für die Literaturwissenschaft von Interesse. Die Musik dagegen tritt vorwiegend, wenn auch nicht ausschließlich, in Zusammenhängen auf, die in den traditionellen Bereich der Musikwissenschaft in jenem engeren Sinn gehören, in dem sie der Literaturwissenschaft an die Seite zu stellen ist. Nur wenig vergrößernd läßt sich sagen: Musikalische Gebilde sind immer zugleich Gegenstand der Ästhetik, sprachliche Gebilde hingegen nicht. Was ich meine, wird auf unverfängliche Weise deutlich, wenn man sich die unterschiedlichen Funktionen vor Augen führt, die sprachliche und musikalische Äußerungen im Rahmen anderer Verhaltensabläufe wie Arbeit, Spiel, Kampf, Therapie oder Erziehung übernehmen können. Beide können die verschiedenen Prozesse begleiten, gliedern, auslösen, beenden, kurz: organisieren. Aber nur mit sprachlichen, nicht mit musikalischen Mitteln kann man solche Prozesse auch beschreiben, erklären, in bezug auf ihre Mittel und Ziele erläutern und organisieren. Oder auf andere Weise zugespitzt: Man kann mit Sprache sowohl Kunst machen wie über Kunst sprechen, mit Musik ist nur das erstere möglich.

Als zweiten Grund könnte man die unterschiedliche Spezifik der Inhalte nennen: Mit sprachlichen Mitteln läßt sich jeder denkbare Sachverhalt, jeder mögliche Gegenstand erfassen und charakterisieren, und zwar innerhalb bestimmter Grenzen mit beliebiger Genauigkeit.² Es ist fraglich, ob in bezug auf musikalische Mittel eine vergleichbare Feststellung überhaupt sinnvoll ist. In jedem Fall wird man die Klasse der identifizierbaren Sachverhalte grundsätzlich einschränken müssen. Vor allem aber ist die musikalische Charakterisierung eines Sachverhalts von grundsätzlich anderer Art. Sie benennt ihn nicht, sondern führt ihn vor – eine Unterscheidung, die später genauer zu bestimmen ist. Vorerst genügt es, die Tatsache zu umschreiben, daß die Mittel der Sprache offensichtlich deutlicher von ihren Inhalten ablösbar sind als die Mittel der Musik. Dies hängt mit der begrifflich strukturierten Be-

deutung sprachlicher Äußerungen zusammen, und es ist zugleich die entscheidende Bedingung für die zuvor erwähnte Vielfalt der Verwendungsweisen sprachlicher Mittel.

Ein dritter Grund ist schließlich die unterschiedliche Ausprägung der Mittel. Der einfachste Beleg für diesen Punkt ist die Tatsache, daß es zahlreiche verschiedene, untereinander äquivalente Sprachen gibt, eine Feststellung, die im Bereich der Musik keinen Sinn macht. Zwar gibt es historisch und geographisch verschiedene Systeme musikalischer Mittel, so wie es historisch und geographisch verschiedene Sprachen gibt. Aber sie sind nicht untereinander äquivalente Systeme in dem Sinn, daß Äußerungen aus einem System in ein anderes übersetzt werden können. Während unmittelbar klar ist, inwiefern z. B. die Äußerungen *Er kommt nicht*, *He doesn't come* und *On ne pridjot* Entsprechungen voneinander in der deutschen, englischen und russischen Sprache sind, ist es unsinnig, parallele Entsprechungen etwa in der japanischen, indischen und französischen Musik angeben zu wollen. Mit anderen Worten, im Bereich der Sprache gilt es, die Gesetzmäßigkeiten zu bestimmen, die den Unterschieden und Gemeinsamkeiten verschiedener Sprachsysteme zugrunde liegen unter einem Gesichtspunkt, der auf die verschiedenen Systeme musikalischer Mittel gar nicht anwendbar ist. Auch dies hängt, wie man sich klarmachen kann, mit der Tatsache zusammen, daß musikalische Mittel sich weniger von ihren Inhalten ablösen lassen als sprachliche. Wir werden dieser Tatsache noch in verschiedener Gestalt begegnen und in ihr einen zentralen Punkt unseres Problems erkennen.

Ich habe drei voneinander nicht unabhängige Gründe angeführt, mit denen die Unterscheidung von Sprache und Literatur einerseits, die Zusammenfassung der entsprechenden Aspekte der Musik andererseits zusammenhängt. Diese Aufzählung ist unvollständig und ganz provisorisch, aber sie liefert Anhaltspunkte, die einen sinnvollen Vergleich von Musik und Sprache leiten können. Diese Anhaltspunkte besagen, daß die ungleiche Aufteilung der Gebiete nicht ein bloß äußerlicher Zufall ist, sondern daß sie im Charakter der Mittel und Funktionsweise von Sprache und Musik selbst begründet ist.

Nun ließe sich einwenden, daß die Musikwissenschaft seit langem die Unterscheidung von Musikgeschichte und systematischer Musiktheorie kennt, und daß diese Einteilung genau der Trennung von Literatur- und Sprachwissenschaft entspricht. Die systematische Musikwissenschaft, von der Analyse der akustischen Grundlagen bis zu den traditionellen Kompositionslehren wie etwa der von Riemann (1902 ff.), ist nicht auf die Wertung und historische Einordnung gegebener Werke aus, sondern auf die Bestimmung der zu verwendenden Mittel. Sie scheint damit durchaus der Linguistik vergleichbar. Doch diese Schlußfolgerung täuscht. Linguistik und Literaturwissenschaft haben nämlich ihrerseits jeweils einen systematischen und einen historischen Aspekt: Die deskriptive und theoretische Linguistik befaßt sich mit der Struktur gegebener Sprachen und den allgemeinen Prinzipien und

Mechanismen, die ihnen zugrunde liegen, die historische Linguistik oder Sprachgeschichte dagegen analysiert den Wandel sprachlicher Erscheinungen und ordnet sie in historische Bezüge ein. Und in der Literaturwissenschaft hat sich neben der Literaturgeschichte eine systematische Literaturtheorie entwickelt – ihre Vorläufer reichen bis in die Rhetorik und Poetik der Antike zurück –, die sich mit der Struktur und Funktionsweise spezifisch literarischer Mittel und Formen befaßt. Metrik und Versbau, Struktur von Erzählformen und Dramen werden dabei in ganz ähnlichem Sinn analysiert wie die musikalischen Mittel in der systematischen Musikwissenschaft.³ Mit anderen Worten, Literatur- und Musikgeschichte haben es beide mit der historischen Einordnung und Wertung überlieferter Gebilde zu tun. Dabei gehört zur Literaturgeschichte auch die Entwicklung literarischer Formen und Mittel (jedoch nicht die Sprachgeschichte) und zur Musikgeschichte auch die der Formen und musikalischen Mittel. Systematische Literatur- und Musikwissenschaft haben es entsprechend mit den Strukturen und Mechanismen zu tun, auf denen die in der Literatur- und Musikgeschichte behandelten Erscheinungen beruhen.

Die in diesem Abschnitt diskutierten Beziehungen lassen sich in folgendem Schema zusammenfassen:

	sprachlicher Bereich		
systematische Disziplinen	theoretische Linguistik	systematische Literaturtheorie	systematische Musikwissenschaft
historische Disziplinen	historische Sprachwissenschaft	Literaturgeschichte	Musikgeschichte
	ästhetischer Bereich		

Bei aller Vereinfachung und terminologischen Unsicherheit gibt das Schema nicht nur die sachlichen Unterscheidungen, sondern auch die methodologischen und inhaltlichen Zusammenhänge im Umfeld unserer Problematik einigermaßen sinngemäß wieder.⁴ Meine weiteren Überlegungen werden sich auf den systematischen Aspekt konzentrieren, der freilich sowenig seiner historischen Dimension beraubt werden kann, wie der historische Aspekt ohne die Voraussetzung systematischer Grundlagen auskommt. Was das Schema nicht richtig deutlich werden läßt, ist die Möglichkeit, Musik und

Sprache direkt, nämlich als Zeichensysteme, und nicht nur unter Zwischenschaltung der Literatur zu vergleichen. Man muß sich das Schema deshalb eigentlich dreidimensional denken. Richtig bleibt indessen der Hinweis, daß der Vergleich von Musik und Sprache unter linguistischem Gesichtspunkt nur in dem Maß sinnvoll und klärend ist, wie dabei zugleich deutlich wird, warum musikalische Mittel so funktionieren, daß sie stets Musik erzeugen, sprachliche Mittel aber nicht notwendig Literatur. Der Blick vor allem auf die Poetik wird das an gegebener Stelle verdeutlichen.

3. Sprache und Musik als akustische Kommunikationsformen

Nachdem wir voreiligen Parallelisierungen zwischen Sprache und Musik vorgebeugt haben, ist es an der Zeit, die wichtigsten Gemeinsamkeiten zu benennen, auf die ein Vergleich sich stützen muß. Die allgemeinste Kennzeichnung dieser Gemeinsamkeit ist offenbar, daß Musik und Sprache akustische Kommunikationsprozesse vermitteln. Beide haben überdies auch schriftliche Darstellungsmittel ausgebildet, die eine abgeleitete Repräsentation der zur Kommunikation benutzten Zeichen ermöglichen. Deren wichtigste Funktion ist, daß sie ein externes Gedächtnis bereitstellen, in dem Äußerungen über den flüchtigen Moment festgehalten werden. Die Konsequenzen dieser Möglichkeit haben tief in die Verfahren der Kommunikation und auch in die Struktur der Mittel eingegriffen, letzteres vor allem in der Musik. Dennoch bleiben Sprache und Musik primär akustische Systeme, ihre schriftliche Fixierung ist sowohl historisch wie systematisch sekundär.

Es ist übrigens aufschlußreich, sich klar zu machen, daß die primär akustische Realisierung für die Musik eine wesentlichere Bedingung ist als für die Sprache. Zugespielt gesagt: Musik, die nur geschrieben wird, ist keine Musik; eine Sprache, die nur geschrieben wird, könnte aber noch immer eine Sprache sein. Natürlich kann man, bei hinreichender Übung, Musik auch „lesen“, so wie man sprachliche Texte liest. Der entscheidende Punkt ist, daß man beim Lesen von Musik die akustische Realisierung kennen, sie gewissermaßen in der Vorstellung hören muß, was für Sprache keineswegs erforderlich ist. Einen sprachlichen Text könnte man vollauf verstehen, auch wenn man keine oder nur falsche Vorstellungen davon hat, wie er auszusprechen ist.⁵ Auszunehmen sind hier bezeichnenderweise nur solche sprachlichen Gebilde, bei denen gerade der Lautwert entscheidend ist, also etwa Gedichte oder Epen. Ich habe diese Asymmetrie verdeutlicht, weil sie wiederum eine Facette der Tatsache ist, daß in der Musik der Sinn weniger von den formalen Mitteln ablösbar ist als in der Sprache – wobei die Dichtung im Interferenzbereich sprachlicher und musikalischer Bedingungen steht. Das Gesagte ließe sich im einzelnen belegen durch eine genauere Analyse der formalen Mittel der schriftlichen Fixierung, die aber ein Thema für sich darstellt.⁶

Ich kehre nach diesem Exkurs zur primären Erscheinungsform zurück und mache zunächst einige Anmerkungen zu den beiden Bestimmungsstücken „akustisch“ und „Kommunikationsprozeß“.

Die primär akustische Erscheinungsform bedingt, daß Musik und Sprache gleichermaßen in der Zeit ablaufen, daß die Zeit also eine Grunddimension ihrer Strukturbildung sein muß. Sie unterscheiden sich damit von Kommunikationsmedien, die nicht zeitlich organisiert sind, insbesondere also von denen der sogenannten bildenden Kunst. Die grundsätzlich akustische Form der Kommunikation unterscheidet Sprache und Musik weiterhin von Medien, die auch an die Zeit gebunden sind, aber andere Sinnesmodalitäten nutzen, also etwa vom Stummfilm oder vom Tanz. Wie aber die akustische Gliederung der Zeit aufgebaut und funktionell belegt wird, darin unterscheiden Musik und Sprache sich entscheidend. Zu diesen Unterschieden gehört auch die Rolle, die die Einbettung der akustischen Vorgänge in begleitende Vorgänge spielt, insbesondere ihre Beziehung zu motorischen Prozessen sowohl bei der Hervorbringung wie bei der Wahrnehmung und Verarbeitung der akustischen Signale. Denn der Gehörsinn und die Motorik der Schallerzeugung sind keine isoliert funktionierenden Teile menschlicher Aktivität, sondern untereinander und mit anderen Komponenten des Gesamtverhaltens auf vielfältige Weise verbunden. Der wechselnde emotionale Status, die Synchronisierung verschiedener Aktivitätsmuster wirken so mit den Kommunikationsabläufen zusammen. Während sie aber für die Sprache vorwiegend instrumentellen oder aber begleitenden Charakter haben, tragen sie bei der Musik entscheidend zur Konstituierung ihres Inhalts bei. Dies wird noch näher zu begründen sein.

Ein Kommunikationsprozeß besteht darin, daß jemand jemandem etwas mitteilt. Diese sehr allgemeine Formulierung steckt einen Rahmen ab, in den musikalische und sprachliche Prozesse grundsätzlich eingeordnet werden können. Grundsätzlich, aber nicht einfach und vielleicht auch nicht restlos. Denn wenn jemand seine Gedanken monologisierend klärt oder auf einem Instrument vor sich hin improvisiert, dann ist da kein Adressat, also geht eigentlich keine Kommunikation vor sich, und doch wird Sprache bzw. Musik realisiert. Ich halte es für einen Trick, hier den Mitteilenden mit dem Empfänger der Mitteilung zu identifizieren, um das Grundschema zu retten. Es ist ehrlicher, zu sagen, daß Sprache und Musik in Kommunikation nicht voll aufgehen. Wenn ich das Monologisieren, entgegen der mythischen Urvorstellung von Musik, dem Flöte blasenden Faun, dennoch für abgeleitet halte, dann eher mit dem Argument von Marx, daß der Mensch nur in der Gesellschaft sich vereinzeln kann. Mit anderen Worten, der Monolog ist nicht Kommunikation, aber er setzt Kommunikation voraus. Dies ist eine kompliziertere Ableitung.

Aber auch, wenn wir von solchen abgeleiteten Erscheinungen absehen, bleibt die Anwendung der genannten allgemeinen Formel schwierig genug. Zu klären sind vier Bestimmungsstücke: *Wer* teilt mit, *wem* wird etwas mit-

geteilt, *was* wird mitgeteilt, und vor allem, *was* heißt *mitteilen*? Es ist üblich, ihren Zusammenhang durch folgendes Schema darzustellen:



Dabei ist der Empfänger in unserem Fall der Hörer, der Sender der Sprecher oder der Musizierende, und die Nachricht sind die Schallwellen. So plausibel dieses Schema aussieht, so ungeeignet ist es, den Kommunikationsprozess zu erfassen. Es ist nicht nur unvollständig, sondern irreführend. Betrachten wir die Bestimmungsstücke genauer.

Zunächst zu den Kommunikationspartnern. Wenn ich zu meiner Frau sage: *Deine Mütze liegt auf dem Stuhl*, dann sind Sender und Empfänger leicht identifizierbar. Wie aber verhält es sich, wenn ein Orchester spielt, wenn ein Redner in einer Versammlung spricht oder – noch komplizierter – Liebhaber Hausmusik ohne Zuhörer machen? Zunächst ist offensichtlich, daß Sender und Empfänger auch Kollektive sein können, und zwar mit einem wichtigen Unterschied: Während in einer Gruppe von Hörern jeder einzelne ein kompletter Empfänger ist, gilt das für einen kollektiven Sender, etwa einen Sprechchor oder ein Orchester, normalerweise nicht. Zwar kommen auch beim kollektiven Hören oft Zusatzeffekte zustande, die beim Einzelhörer allein nicht auftreten, vor allem, aber nicht nur dann, wenn die Hörer, etwa beim Tanzen oder Marschieren, in einer gemeinsamen Tätigkeit befangen sind, die durch den Kommunikationsprozess unmittelbar strukturiert wird. Aber diese Effekte modifizieren nur die Aufnahme der Mitteilung, sie konstituieren den Empfang nicht und sind nicht mit den Bedingungen vergleichbar, die z. B. einen Orchestermusiker zum Teil eines kollektiven Senders machen. Hier ist der einzelne sehr oft kein in sich vollständiger Sender, er produziert vielmehr nur Teile der Mitteilung, die jeder einzelne Hörer als Ganzes aufnimmt. Wenn wir also den Kommunikationsprozess so elementar wie möglich bestimmen, dann ist der Empfänger immer der einzelne Hörer – der gegebenenfalls im Rahmen einer Gruppe die gleiche Mitteilung aufnimmt –, der Sender dagegen ist je nach Art des Prozesses ein einzelner oder ein Kollektiv, das als Ganzes fungiert.⁷

Die Möglichkeit des kollektiven Senders bringt eine weitere Komplizierung ins Spiel. Wenn die Mitteilung durch einen kollektiven Sender erzeugt wird (aber nicht nur dann), gehen dem hier betrachteten Kommunikationsprozess eine Reihe vorbereitender Phasen voraus. Planung, Aufbau und Aussendung einer Mitteilung werden nur im einfachsten Fall gewissermaßen in einem Schritt und von einer Person vollzogen. Anders ausgedrückt, Kommunikation kann ein vielstufiger Prozess sein, von dem wir uns hier auf die letzte Phase konzentrieren müssen, die Phase, in der die Mitteilung den

Empfänger erreicht, in der, einfach gesagt, die Musik erklingt, die Sprache hörbar wird. Vor dieser Phase kann eine Reihe verwickelter Prozesse liegen, an denen verschiedene Personen beteiligt sein können, die auf unterschiedliche Weise am Zustandekommen der Mitteilung beteiligt sind: als Komponist oder Autor, als Arrangeur, Dirigent, Regisseur oder Übersetzer zum Beispiel. Viele dieser Prozesse schließen selbst Kommunikationsvorgänge ein, die im Prinzip den gleichen Überlegungen zu unterwerfen sind, die wir hier für die entscheidende letzte Phase anstellen. Die Ausweitung des einfachen Falls, daß eine einzelne Person einem Hörer oder einem Hörerkreis etwas mitteilt, kommt nicht nur bei musikalischen, sondern auch bei sprachlichen Äußerungen vor – man denke nur an den Text eines Schauspielers in einem Bühnenstück –, aber sie spielt eine ungleich größere Rolle in der Musik als in der Sprache: Während die direkte Kommunikation zwischen einem Mitteilenden und einem Hörer für die Sprache den Grundtypus darstellt, ist sie im Bereich der Musik eher der Sonderfall, auch wenn man nicht den modernen Konzertbetrieb im Sinn hat. Der Grund dafür liegt wesentlich in der bereits erörterten Unterschiedlichkeit der Zwecke oder Funktionen von Musik und Sprache, die wiederum von der Struktur und den Inhalten der beiden Medien abhängen.

Daß wir uns im folgenden auf die letzte Phase (bzw. den einfachsten Fall) des Kommunikationsprozesses konzentrieren, hat zwei Gründe. Erstens und vor allem ergibt sich aus dieser Phase der Sinn auch der vorgeschalteten Prozesse, was wiederum heißt, daß Struktur und Funktion der Zeichensysteme, um die es uns geht, aus dieser und nur aus dieser Phase zu entnehmen sind. Unter diesem Gesichtspunkt können wir die Gesamtheit der vorbereitenden Prozesse, von der Konzipierung eines Werkes bis zu den Dispositionen der Interpreten, als interne Vorgänge in einem kollektiven, gegebenenfalls vielstufigen Sender auffassen. Und zweitens sind die verschiedenen vorgeschalteten Phasen durch die Orientierung auf die Endphase determiniert und setzen deren Verständnis immer schon voraus.⁸

Nach dieser Verständigung über die Kommunikationspartner im primären Kommunikationsakt bleibt das zu klären, was in dem oben wiedergegebenen Schema am unzulänglichsten erfaßt wird: der Charakter des Mitteilens und der Mitteilung. Entgegen dem ersten Anschein ist das eine ziemlich verwickelte Aufgabe, der ich einen eigenen Abschnitt widmen will.

4. Grundbedingungen der Kommunikation

Zweifellos gilt für Sprache und Musik gemeinsam der einfache Grundsatz: Ohne Schallereignis keine Mitteilung, was sich etwas genauer durch folgende Bedingung ausdrücken läßt:

(B 1) Ein Sender *S* produziert ein akustisches Ereignis *E* und der Hörer *H* nimmt das Ereignis *E* wahr.

Das ist sicher eine vernünftige Interpretation des angeführten Schemas, wenn

wir *E* als die Nachricht ansehen, zeigt aber zugleich, daß damit nicht erklärt ist, wie eine Mitteilung zustande kommt. Denn zahlreiche Vorgänge, die gewiß keine Kommunikation sind, erfüllen diese Bedingung, so z. B. der folgende:

(A) *S* ist ein Gast in einem Restaurant, *H* ein Gast am Nebentisch, der das gelegentliche Klappern der Bestecke von *S* hört, dem Geräusch aber keine Beachtung schenkt.

Ich will jetzt zeigen, daß noch drei weitere Bedingungen nötig sind, um Kommunikation hinreichend zu bestimmen. Ich adaptiere dabei eine Analyse, die, ausgehend von Grice (1957), vor allem im Hinblick auf sprachliche Kommunikation entwickelt worden ist. Die Unterscheidungen, die dabei zu treffen sind und die ich zunächst an Hand einfacher Beispiele begründe, die weder musikalischer noch sprachlicher Natur sind, mögen manchem als Haarspalterei erscheinen. Sie werden aber für unser Problem wesentliche Merkmale ergeben.

Zunächst schließen wir Fälle vom Typ (A) aus, indem wir ausdrücklich verlangen, daß der Hörer das Signal beachtet.

(B 2) Die Wahrnehmung des Ereignisses *E* erzeugt eine bestimmte Wirkung *W* im internen Zustand des Hörers *H*.

Die Wirkung *W* soll sich auf den inneren Zustand im weitesten Sinne, auf die Kenntnisse, Einstellungen, Absichten, Erwartungen, Bewertungen und Emotionen des Hörers beziehen, jedoch nicht notwendig Konsequenzen im äußeren Verhalten zur Folge haben. Ein aufmerksamer, aber regloser Zuhörer im Konzert oder bei einem Gespräch würde sonst nicht als Kommunikationspartner gelten. Betrachten wir nun als nächstes Fälle folgender Art:

(B) *S* schlägt nachts mit dem Hammer in seiner Wohnung einen Nagel in die Wand. *H* ist sein Nachbar, der sich darüber ärgert.

Offensichtlich sind hier (B 1) und (B 2) erfüllt, und doch wird man nicht von Kommunikation sprechen, weil nämlich die Wirkung beim Hörer nicht im Plan des Senders enthalten ist. Man muß also fordern, daß die Signal-erzeugung durch den Sender und die beim Hörer eintretende Wirkung nicht zufällig zusammentreffen.

(B 3) Daß das Ereignis *E* im Hörer *H* normalerweise die Wirkung *W* erzeugt, ist der Grund dafür, daß der Sender *S* das Ereignis *E* produziert.

Da im Fall (B) das Hämmern nicht stattfindet, um den Nachbarn zu stören, liegt gemäß (B 3) hier keine Mitteilung vor. Für die Formen der Kommunikation, die uns hier interessieren, können wir den in (B 3) genannten ‚Grund‘ übrigens einfach als Intention oder Absicht des Senders bestimmen, wenn wir Absicht in genügend weitem Sinn, also nicht notwendig bewußt überlegte Absicht, auffassen. (In der Tierkommunikation hingegen sind Signalproduktion und Wirkung häufig durch genetisch erworbene Kausalmechanismen verknüpft.) Ich ersetze (B 3) also durch die etwas einfachere Bedingung (B 3'):

(B 3') Der Sender *S* produziert das Ereignis *E*, um durch *E* im Hörer *H* die Wirkung *W* zu erzeugen.

Genau diese Bedingung wird nun aber erfüllt, wenn man sich den Fall (B) folgendermaßen abgewandelt denkt:

(C) *S* will seinen Nachbarn ärgern und hämmert deshalb nachts einen Nagel in die Wand. Der Nachbar *H* ist über den Lärm verärgert.

Die Situation ist nicht eindeutig. Da der Störenfried nicht nur rücksichtslos, sondern mit der Absicht zu stören hämmert, ist die Wirkung bei der Signalproduktion geplant. Dennoch liegt keine Kommunikation vor, solange der Nachbar keine Absicht in der Störung erkennt. Zur Verdeutlichung stelle ich eine echte Mitteilungssituation gegenüber:

(D) *S* klopft an die Wand, um den Hörer *H* zu mahnen, sein Radio leiser zu stellen.

Gleichgültig ob *H* die Mahnung befolgt oder nicht: Wenn er das Klopfen als Mahnung verstanden hat – und nur dann –, ist eine echte Mitteilung zustande gekommen. Den Unterschied zwischen (C) und (D) drückt folgende Bedingung aus:

(B 4) Der Sender *S* produziert das Ereignis *E* in der Absicht, im Hörer *H* die Wirkung *W* dadurch zu erzeugen, daß *H* eben diese Absicht bemerkt.

Damit ist der Kreis geschlossen: Sender und Empfänger sehen beide das Signal als Auslöser der Wirkung *W* an, wegen dieser Wirkung wird es produziert, und es wirkt auch nur, weil es um dieser Wirkung willen produziert wird. Damit ist ein kommunikatives Signal von allen anderen Signalen hinlänglich unterschieden. Wir können nun die Wirkung *W*, auf die es der Sender abgesehen hat, und die – im Fall einer gelungenen Kommunikation – im Hörer auch zustandekommt, die Mitteilung nennen. Die Bedingung (B 4) sichert dabei, daß nicht irgendwelche, sondern eben nur die als solche beabsichtigten Wirkungen zur Mitteilung gehören.⁹ Obwohl das letztere eine grundsätzlich berechtigte und sogar notwendige Festlegung ist, läßt ihre Anwendung im einzelnen interessante Spielräume offen. Wir kommen darauf zurück.

Die Erklärung, daß die Mitteilung die absichtlich im Empfänger hervorgebrachte Wirkung *W* ist, ist natürlich noch viel zu unbestimmt, vor allem sagt sie nichts über die Spezifik von Musik und Sprache. Ehe ich dieses Problem weiter verfolge, komme ich noch auf einen wichtigen Zusammenhang zwischen Signal und Mitteilung, also zwischen dem Ereignis *E* und der Wirkung *W*, zurück. Betrachten wir dazu nochmals den Fall (C) des böswilligen Nachbarn und modifizieren ihn so:

(E) *S* hämmert nachts einen Nagel in die Wand, um seinen Nachbarn zu ärgern und in der festen Annahme, daß dieser die Absicht erkennt. Der Nachbar *H* tut das auch und ärgert sich über den Lärm und die Absicht.

Da dieser Fall die Bedingung (B 4) erfüllt, müssen wir ihn als Kommunika-

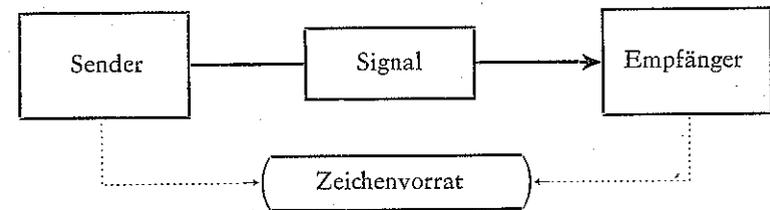
tionsakt ansehen, und ich denke, zu Recht. Zumindest, was den zweiten Ärger angeht, nämlich den über die Absichtlichkeit der Störung. Worauf es mir ankommt, ist, daß hier zwei Wirkungen durch das Ereignis E ausgelöst werden: Auch wenn die beiden Nachbarn nicht voneinander wüßten, daß der eine den andern ärgern will, würde H sich über den Lärm ärgern. Das Ereignis E hat also eine extrakommunikative Wirkung und eine, die durch seine kommunikative Interpretation zustandekommt. Ich will die erste mit $W(n)$, die zweite mit $W(k)$ abkürzen. Generell hängt ~~man~~ $W(n)$ enger mit dem Charakter des Signals zusammen als $W(k)$.⁴⁰ Interessant ist nun der Zusammenhang von $W(k)$ mit $W(n)$ (und auf diesem Weg mit dem Charakter des Signals). Im Beispiel (E) motiviert offenbar die eine Wirkung die andere: Die Mitteilung ‚ich störe jetzt‘ geschieht durch ein Signal, das schon an sich als Störung wirkt. Solche „Motivation“ von $W(k)$ durch $W(n)$ kann vielfältiger Art sein. Es kann z. B. einfach eine nichtkommunikative Handlung in eine kommunikative umfunktioniert werden, so etwa, wenn jemand nicht erzählt, sondern vormacht, wie ein Ereignis, etwa ein Diebstahl, sich abgespielt hat. Ein motivierender Zusammenhang zwischen $W(k)$ und $W(n)$ ist keineswegs die Regel, häufig hat $W(n)$ mit der eigentlichen Mitteilung gar nichts zu tun, kann sogar praktisch gleich null sein. Zum letzteren Typ gehören weitgehend sprachliche Kommunikationsakte: Die außerkommunikative Wirkung der Äußerung *Komm mit* hat mit der durch sie vermittelten Aufforderung so gut wie nichts zu tun.

Meine These, die ich im weiteren begründen will, ist nun, daß in der Musik (und in vielen anderen Bereichen der Kunst) der motivierende Zusammenhang zwischen $W(n)$ und $W(k)$ eine entscheidende Rolle spielt, daß die zweite ohne die erste praktisch nicht zustandekommt. In der Sprache spielt die außerkommunikative Wirkung des Signals dagegen eine ganz marginale Rolle. Ein einfaches Beispiel: Die Mitteilungen, die mit dem gleichen Satz gemacht werden, wenn er einmal langsam, ein andermal schnell gesprochen wird, unterscheiden sich viel weniger als die, die mit einer einmal langsam, einmal schnell vorgetragenen musikalischen Äußerung gemacht werden. Ähnliches gilt für unterschiedliche Lautstärke.⁴¹ In diesem Verhältnis von „natürlicher“ und „kommunikativer“ Wirkung liegt der Schlüssel zu der mehrfach erwähnten Tatsache, daß sprachliche Mitteilungen weniger eng an ihre Mittel gebunden sind als musikalische.

An die Unterscheidung der beiden Wirkungen $W(n)$ und $W(k)$ schließt sich automatisch die Beobachtung an, daß auch die kommunikative Wirkung $W(k)$ aus zwei (oder noch mehr) Komponenten zusammengesetzt sein kann, die auf verschiedene Aspekte des Signals zu beziehen sind. Ein Lehrer, der heftig mit dem Bleistift auf den Tisch klopft, teilt mit dem Klopfen etwa eine Mahnung zur Aufmerksamkeit mit und durch die Heftigkeit des Klopfens die Dringlichkeit der Mahnung. Hierher gehört nun natürlich vor allem gesungene Sprache, also das Zusammenwirken einer sprachlichen und einer musikalischen Mitteilung. Noch weiter in dieser Richtung geht das Zusam-

menspiel verschiedener Mitteilungsformen im Musiktheater. Ehe komplexe Erscheinungen dieser Art ins Auge zu fassen sind, ist es jedoch nötig, über die Mitteilung mehr zu sagen, als daß sie die beabsichtigte Wirkung $W(k)$ ist.

Die Hauptfrage ist, wie die für das Zustandekommen der Mitteilung nötige Abstimmung der Absichten und (internen) Reaktionen zustandekommt, die in der Bedingung (B 4) gefordert wird, und welchen Charakter diese Grundlage der Mitteilung hat. Für unsere Beispiele (D) und (E) können wir dabei auf das Wissen verweisen, das jeder der beiden Nachbarn über den anderen hat. Was aber ist das nötige Hintergrundwissen im allgemeinen, also vor allem bei Partnern, die sich nicht kennen, und was verleiht ihm die nötige Stabilität? Die übliche Antwort auf diese Frage ist die Annahme, daß Sender und Empfänger sich auf einen gemeinsamen Vorrat an Zeichen beziehen, wobei dann die Signalereignisse als Zeichenträger, die Bedeutungen der Zeichen als Wirkungen, also als Mitteilungen, angesehen werden. In diesem Sinn wird dann das Kommunikationsschema meist folgendermaßen erweitert:



Da Zeichen in der Tat eine wesentliche Bedingung der Kommunikation sind, ist dieses Schema zwar besser als das auf Seite 16 wiedergegebene. Dennoch gibt es Anlaß für gravierende Mißverständnisse. Zwar können die Bedingungen (B 1) bis (B 4) als Leitfaden für die Interpretation des neuen Schemas dienen, falls wir den Zeichenvorrat als Kenntnisse sowohl des Senders wie des Empfängers ansehen (und ihn also konsequenterweise – genauso wie auch die Mitteilung – im Inneren der Kommunikationspartner lokalisieren). Aber das Schema beantwortet trotzdem einige Fragen gar nicht oder aber falsch.

Falsch ist vor allem die unterstellte Gleichsetzung von Zeichenbedeutung und Mitteilung. Am besten ist das an Hand der sprachlichen Kommunikation zu verdeutlichen. Ich muß dabei im Augenblick annehmen, daß wir bereits wissen, was die Bedeutung sprachlicher Zeichen, also z. B. eines Satzes wie *Ich warte auf eine Antwort* ist.

Die Äußerung eines Satzes kann bei gleicher Bedeutung unter verschiedenen Umständen mit ganz unterschiedlichen Mitteilungen verbunden sein. Für

unseren Beispielsatz ergäbe sich als Mitteilung z. B. eine Feststellung, die mit der Bedeutung des Satzes beinahe identisch ist, wenn er ein Gespräch abschließt, in dem der Sprecher dem Hörer von einem Brief an einen Freund erzählt hat. Hat dagegen der Sprecher dem Hörer eine unangenehme Frage gestellt und äußert dann nach längerem Schweigen den fraglichen Satz, so ist das eine Aufforderung, sich zu der Frage zu äußern. Der mit diesem simplen Beispiel angedeutete Spielraum zwischen Bedeutung und Mitteilung kann außerordentlich groß sein und umfaßt sehr verschiedenartige Erscheinungen. Die sogenannte übertragene oder metaphorische Verwendung sprachlicher Ausdrücke gehört ebenso dazu wie das Auffinden der nötigen Bezugspunkte für Wörter wie *dort* und *es*, etwa in der Äußerung von *Dort liegt es*.¹²

Wenn nun Zeichen die Basis dafür sein sollen, daß kommunikativ verwendete Signale eine bestimmte Wirkung $W(k)$ haben, dann fragt sich, wie der Zusammenhang zwischen der Bedeutung B eines bestimmten Zeichens und dieser Mitteilung $W(k)$ zustandekommt. Naheliegend ist folgende Hypothese:

(H 1) Für jedes Zeichen Z aus dem Zeichenvorrat gibt es eine Klasse von Primärsituationen, in denen die Bedeutung B von Z mit der durch die Äußerung von Z gemachten Mitteilung $W(k)$ identisch ist.

Die Klasse von Primärsituationen wären dann die Fälle, in denen sich die Bedeutung des Zeichens herausbildet, in denen sie also erworben wird. Für alle anderen, nicht-primären Situationen wäre dann folgende zweite Hypothese anzunehmen:

(H 2) Es gibt zu jedem Zeichenvorrat ein allgemeines System von Prinzipien, auf das sich Sender und Empfänger stützen, um auf Grund der Bedeutung B eines Zeichens Z die situationsspezifische Mitteilung $W(k)$ zu bestimmen.

Die in (H 2) genannten Prinzipien machen das aus, was in bezug auf sprachliche Zeichensysteme allgemein als Pragmatik aufgefaßt wird.

Die in (H 2) angenommenen „pragmatischen“ Prinzipien sind in dem angegebenen Kommunikationsschema, auch in seiner erweiterten Form, aber gar nicht vorgesehen, weshalb ich es für falsch erklärt habe. Hinzu kommt, daß wir die Hypothese (H 1) fallenlassen müssen. Das ergibt sich bereits aus der Tatsache, daß es chamäleonhafte Zeichen gibt, deren Anteil an der Mitteilung von Situation zu Situation wechselt, für die es also gar keine wirkliche „Primärsituation“ gibt, weil ihre Bedeutung gewissermaßen situationsdurchlässig ist. *Dort, damals, solche* und viele andere sogenannte deiktische Elemente gehören hierher. Verwickelter sind Probleme von komplexen Zeichen, die ohne Hinzunahme der in (H 2) genannten Prinzipien gar keine Mitteilung machen, weil sie wörtlich genommen keine Bedeutung haben, so daß auch keine Primärsituationen für sie existieren können. Man versuche, sich über die wörtliche Bedeutung des Satzes *Aus Sümpfen sollte man keine Konsequenzen ziehen* von Jerzy Lec Rechenschaft zu geben. Wir

sind hier bei der folgenreichen Tatsache angelangt, daß durch die Bezugnahme auf einen Zeichenvorrat mehrere Stufen des ‚Verstehens‘ einer kommunikativen Äußerung entstehen. Zu unterscheiden ist mindestens das Erfassen der Bedeutung der verwendeten Zeichen vom Erfassen der intendierten Mitteilung. Im Grenzfall kann beides identisch sein.¹³ Verstehen heißt im ersten Sinn, wissen, was die verwendeten Zeichen bedeuten, und im zweiten Sinn, erfassen, was mit der Äußerung vermittelt werden soll. (*Ich verstehe die Erläuterungen nicht, denn ich kann kein Spanisch bzw. denn ich habe keine Ahnung von Wirtschaftsrecht* sind Beispiele für die beiden Arten von Verstehen.)

Was ich damit für sprachliche Kommunikation deutlich zu machen versucht habe, gilt ebenso für musikalische Äußerungen, ist dort aber schwerer zu illustrieren, vor allem, weil es kein ähnlich sicheres Vorverständnis dafür gibt, was in diesem Bereich „Bedeutung“ ist. (Nicht zufällig heißt *Ich verstehe diese Musik nicht* praktisch immer nur „Ich weiß nicht, was sie vermitteln soll“, bezieht sich also auf die Mitteilung, nicht die Bedeutung.) Um wenigstens einen ersten Anhaltspunkt zu haben, vergleiche man z. B. die intendierte Mitteilung, wenn das Adagietto aus der *Fünften Sinfonie* von Mahler einmal im Rahmen dieser Sinfonie, ein andermal als Begleitung in Viscontis Verfilmung von *Tod in Venedig* gespielt wird.¹⁴ Was sich als Bedeutung der musikalischen Zeichen herauschälen läßt, ist in beiden Fällen gleich, nicht jedoch die ausgelöste (beabsichtigte) Wirkung. Diese Behauptung erhält einen klaren Sinn erst, wenn wir uns über den Charakter musikalischer Zeichen und ihrer Bedeutung ein besseres Bild verschafft haben. Sie wird sich dann als sehr wesentlich für das Verständnis der Wirkungsweise von Musik erweisen.

Die These, daß Bedeutung und Mitteilung nicht identifiziert werden dürfen, erklärt natürlich weder, was Bedeutungen noch was Mitteilungen sind. Ich habe Mitteilungen bisher sozusagen von außen, aus dem Kommunikationsakt heraus, als Wirkung $W(k)$ bestimmt, ohne $W(k)$ weiter zu qualifizieren. Der Weg, den ich zu diesem Ziel einschlagen will, ist nun der, die Zeichensysteme genauer zu analysieren und dadurch zur Einsicht in die Natur der Mitteilungen zu gelangen, die mit ihnen gemacht werden können. Ein zentrales Problem ist dabei die Tatsache, daß Sprache und Musik nicht nur erlauben, mit dem gleichen Zeichen gegebenenfalls viele verschiedene Mitteilungen zu kommunizieren, sondern vor allem auch nach Bedarf immer neue komplexe Zeichen zu bilden. Mit dem Ausdruck „Zeichenvorrat“ ist diese Tatsache eher verdeckt als berücksichtigt. Zu klären ist also, was Zeichen sind, wie sie zustandekommen, wie neue komplexe Zeichen funktionieren und wie sich daraus die Spezifik sprachlicher und musikalischer Mitteilungen herleiten läßt.

5. Zeichen und Zeichensysteme

In diesem Abschnitt will ich einige Festlegungen zum Zeichenbegriff treffen, die notwendigerweise sehr allgemein gehalten sind und später mit Substanz gefüllt werden müssen.¹⁵

Als erstes ist auf Grund der vorangehenden Überlegungen festzuhalten, daß Zeichen eine besondere Form von Kenntnissen sind, also auf dem Gedächtnisbesitz der Kommunikationspartner beruhen. Zeichen beziehen sich zwar auf physikalische Signalereignisse, sind aber nicht mit ihnen identifizierbar. Ihr Verhältnis zu den Ereignissen ist abstrakt in genau dem Sinn, in dem z. B. ein Begriff wie „Baum“ abstrakt ist im Verhältnis zu den wirklichen Bäumen.

Weiterhin wissen wir, daß ein Zeichen dazu geeignet sein muß, eine Beziehung zwischen einem Signalereignis E und einer Mitteilung $W(k)$ herzustellen, wengleich noch offen ist, wie diese Beziehung zustandekommt. Jedenfalls also muß ein Zeichen zwei Seiten oder Aspekte haben. Die eine will ich die formale Seite, kurz Form F des Zeichens nennen, die andere habe ich bereits als seine Bedeutung B eingeführt. Damit haben wir ein Zeichen Z als eine durch das Gedächtnis getragene Verbindung einer Bedeutung B mit einer Form F bestimmt.¹⁶

Die Form F eines Zeichens ist nun auf zweierlei Weise mit dem Signalereignis E verbunden. Für den Sender einer Mitteilung dient sie als Muster, das die Produktion des Ereignisses E steuert und reguliert. Sie organisiert also den Artikulationsprozeß beim Sprechen und Singen, aber auch die Motorik beim instrumentalen Musizieren. Für den Empfänger ist sie ein Muster, das aus dem Signal entnommen, besser noch: durch das das Signal in bestimmter Weise identifiziert werden muß. Das Signal muß als Realisierung einer bestimmten Zeichenform wahrgenommen werden, so wie ein Baum als Realisierung des Typs Baum wahrgenommen werden muß. Die Beziehung zwischen Ereignis und Zeichenform kommt für den Hörer also auf Grund der Mechanismen zustande, die allgemein die ordnende Wahrnehmung der Umwelt tragen. Mit einiger Vorsicht können wir die Form F eines Zeichens als die Invariante aller seiner möglichen Realisierungen durch verschiedene Signalereignisse E betrachten, und zwar die Invariante gegenüber den Prozessen der Signalproduktion wie auch der Signalwahrnehmung. Diese Tatsache macht es möglich, die Form eines Zeichens weitgehend mit Hilfe der relevanten physikalischen Eigenschaften oder auch der Modalität seiner Produktion zu charakterisieren.

Die Bedeutung B eines Zeichens muß für Sender und Empfänger mit der beabsichtigten Mitteilung $W(k)$ verbunden sein. Für diese Verbindung habe ich in der Hypothese (H 2) nicht näher ausgewiesene (pragmatische) Prinzipien verantwortlich gemacht. Sie sind insofern das Pendant zu den eben besprochenen Produktions- und Perzeptionsmechanismen, als sie die Zeichenbedeutung situationsspezifisch konkretisieren. Für besondere Fälle

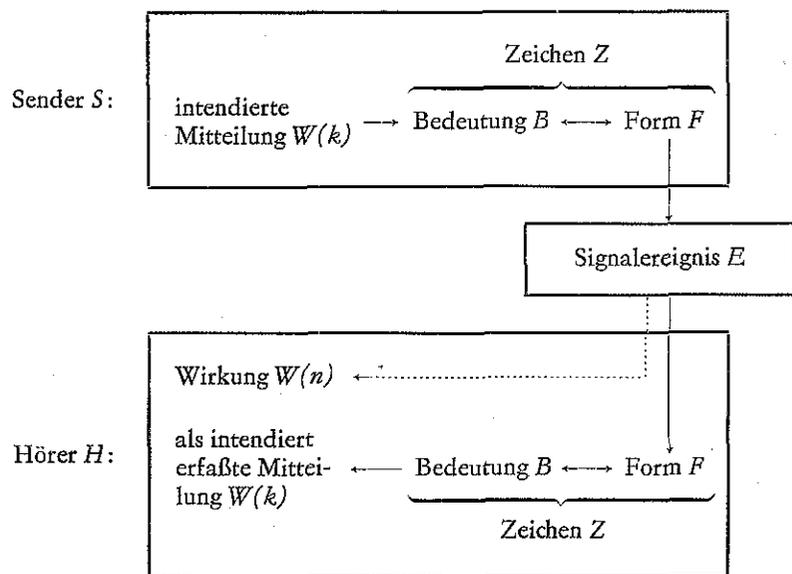
läßt sich die Parallele sogar weiterführen zu der Feststellung, daß die Bedeutung B eines Zeichens die Invariante aller möglichen Mitteilungen $W(k)$ ist, die mit dem Zeichen gemacht werden können. Invarianz besteht dann gegenüber den pragmatischen Prinzipien aus (H 2). Für kompliziertere Fälle allerdings wird diese Feststellung sinnlos. Wir werden sehen, daß die verschiedenen Arten der Varianz für Sprache und Musik eine entscheidende Rolle spielen. (Ich habe oben ja gerade verschiedene Formen des Verstehens mit dem Erfassen der Bedeutung und der Mitteilung in Zusammenhang gebracht.)

Betrachten wir zur Illustration des bisher Gesagten zwei einfache Beispiele: Das optische Muster „rote Kreisfläche mit weißem Querstrich“ ist die Form des Verkehrszeichens, dessen Bedeutung gewöhnlich mit „Einfahrt verboten“ wiedergegeben wird. Die konkreten Ereignisse sind hier die einzelnen Schilder, die an entsprechenden Stellen angebracht werden; die Mitteilung, die durch wirklich aufgestellte Schilder gemacht wird, legt über die Bedeutung hinaus noch den Ort fest, auf den sich das Verbot bezieht. Die Varianz des Signals E und der Mitteilung $W(k)$ ist hier aus offensichtlichen Gründen nur sehr gering gegenüber der Form F und der Bedeutung B . Anders bei dem zweiten Beispiel, einem Wort der natürlichen Sprache. Das mit grün transkribierte Lautmuster ist die Form des Wortes, dessen Bedeutung ein Farbbegriff ist. Die Ereignisse sind hier alle realen Äußerungen des Wortes, die Mitteilung ist (unter anderem) die jeweils gemeinte Farbwahrnehmung oder Farbvorstellung. Innerhalb einer Äußerung wie *Der Apfel ist noch ganz grün!* kann die Mitteilung aber auch Unreife oder Geschmackswerte einschließen. Die Variation von $W(k)$, aber auch von E ist hier unvergleich größer. (Man denke nur an die Aussprache des gleichen Satzes durch einen nuschelnden Mann und eine deutlich artikulierende Frau.)

Ich fasse die erläuterten Komponenten und Beziehungen noch einmal schematisch zusammen. Zu der zeichenvermittelten Mitteilung $W(k)$ füge ich dabei noch die mögliche extrakommunikative Wirkung $W(n)$ des Ereignisses E hinzu (siehe Schema auf der folgenden Seite).

Natürlich ist auch dieses Schema nur so gut wie die Erläuterungen, die seine Interpretation festlegen. (In diesem Sinn ist zu bedenken, daß auch die Wirkung $W(n)$ durch entsprechende vermittelnde Mechanismen zustandekommt, die die verschiedensten psychischen und physischen Komponenten einschließen können.)

Als nächstes haben wir uns zu vergegenwärtigen, daß Zeichen nicht isoliert existieren, sondern zu Zeichensystemen zusammengeschlossen sind. Innerhalb ihres Systems bestimmen die Zeichen sich wechselweise, erst dadurch legen sie ihren Wert, ihre speziellen Grenzen fest. Sowohl die Bedeutungen wie die Formen werden damit zu Elementen jeweils eines besonderen, mehr oder weniger komplex strukturierten Bereichs. Für jedes Zeichensystem Z^* muß sich demnach ein Bedeutungsbereich B^* und ein Formenbereich F^* identifizieren lassen. Zwischen diesen beiden Bereichen muß eine Zuord-



nung bestehen, die jeweils die einzelnen Bedeutungen aus dem Bereich B^* mit den einzelnen Formen aus dem Bereich F^* verbindet. Ein Zeichensystem Z^* ist mithin generell durch die Festlegung von B^* und F^* mit ihren jeweiligen Strukturen und eine Zuordnung zwischen diesen beiden Bereichen bestimmt. Diese Zuordnung schafft das charakteristische „Quidproquo“ der Zeichen, also die Möglichkeit, daß die Form (und ihre Manifestation im Signal) für die Bedeutung steht, auf sie verweist.

Eine Zuordnung der hier geforderten Art, die zwei strukturierte Mengen so aufeinander abbildet, daß die Elemente der einen die der anderen repräsentieren können, heißt üblicherweise ein Code. Ein jedem geläufiges Beispiel ist der Morse-Code. Die Formen bestehen hier aus Folgen von Längen und Kürzen, die Bedeutungen sind die Buchstaben und Ziffern. Die Zuordnungsfunktion wird durch eine Liste angegeben. Den Morse-Code kennen, heißt, diese Codierungsliste im Gedächtnis gespeichert zu haben.

Wenn natürliche Sprachen und musikalische Systeme im eben angegebenen Sinn als Codes betrachtet werden sollen, dann wird sofort deutlich, daß die Codierung hier anderen Charakter haben muß als beim Morse-Code, weil sie nämlich nicht durch eine abgeschlossene Liste von Zeichen angegeben und im Gedächtnis gespeichert werden kann. Dies bringt uns zum letzten wichtigen Punkt dieses Abschnitts: Sprachliche und musikalische Zeichen-

systeme müssen durch Verfahren oder Regeln gekennzeichnet sein, mit deren Hilfe beliebige komplexe Formen gebildet und ihren Bedeutungen zugeordnet werden können. Nur so wird verständlich, wieso immer wieder neue Zeichen gebildet und verstanden werden können. Dabei sind komplexe Zeichen keineswegs nur die Summe ihrer Bestandteile, wie schon ein einfaches Beispiel zeigt: *Er kommt um* wird nicht nur erweitert, sondern ändert seine ganze Bedeutung, wenn es zu dem komplexeren Zeichen *Er kommt um den Verstand* oder *Er kommt um neun Uhr* ergänzt wird. Für die Musik gilt ähnliches, schon weil hier nicht einmal klar ist, was die einfachen, isolierbaren Zeichen des Codes eigentlich sind. Kurz, die Zeichensysteme der Sprache und der Musik sind nicht feste Zeichenvorräte, sondern weit eher Systeme zur Zeichenbildung. Was das heißt, und worin die ganz unterschiedlichen Verfahren bestehen, ist zu klären, indem wir uns für Musik und Sprache systematisch die folgenden drei Fragen vorlegen, an denen, wie oben gesagt, die Bestimmung von Zeichensystemen hängt:

- Was macht den Bereich F^* aus, aus dem die Formen der Zeichen stammen, wie ist er strukturiert?
- Was macht den Bereich B^* der Zeichenbedeutungen aus und wie kommt seine Struktur zustande?
- Wie werden die beiden Bereiche konstruktiv aufeinander bezogen?

Die Beantwortung dieser Fragen zwingt uns dazu, die globale Redeweise von Bedeutungen B und Formen F aufzugeben und deren innere Struktur ins Auge zu fassen. Auf diese Weise werden wir schließlich auch den Charakter der möglichen Mitteilungen und ihren Bezug zu den verwendeten Zeichen genauer bestimmen können. Als Vororientierung gebe ich einige provisorische Anhaltspunkte für die gesuchten Antworten.

Für Sprache und Musik besteht der Formenbereich F^* aus gegliederten Lautmustern. Das ist nach dem Gesagten offensichtlich. Wesentlich sind die unterschiedlichen Merkmale, Dimensionen und Strukturbildungen, die Sprache und Musik in diesem Bereich verwenden.

Der Bedeutungsbereich B^* besteht für die Sprache aus den Strukturen, die das Denken aus der Realität abstrahiert, aus Begriffen für Dinge, Eigenschaften, Sachverhalte und Operationen mit ihnen. Um eine ähnlich allgemeine Angabe für die Musik zu machen, greife ich vor auf einen noch zu erläuternden Begriff, den der Struktur von Gesten. Er steht hier nur als Hinweis auf genauer zu Erklärendes.

Die dritte Frage ist für die Sprache durch die Angabe von zwei Komponenten zu beantworten: ein festes, wengleich sehr kompliziertes System von Grundzeichen und eine bedeutungsbildende Syntax, die die Zeichenverknüpfung regelt. Für die Musik hat die Antwort einen grundsätzlich anderen Charakter. Während Form und Bedeutung eines sprachlichen Zeichens prinzipiell verschieden strukturiert sind und deshalb eine besondere Bezugsetzung verlangen, sind sie in der Musik in der Tendenz isomorph, das heißt, sie haben einen ganz analogen Aufbau. Die Zuordnung der Bedeu-

tion ist deshalb mit dem Aufbau der Form bereits mitgesetzt.¹⁷ Zugespitzt führt das zu der These, daß Form und Bedeutung in der Musik nicht zwei unterschiedlich strukturierte Bereiche bilden, sondern gewissermaßen zwei Interpretationsbereiche der gleichen Struktur. In der Tat halte ich das für die strukturelle Seite des mehrfach erwähnten Befunds, daß in der Musik die Bedeutung nicht wirklich von ihren formalen Mitteln abgehoben werden kann. Der Sinn dieser These und ihre Begründung sollen im folgenden deutlich werden.

6. Kommunikative, kognitive und emotive Strukturen

Wie sind die eben formulierten Fragen etwas verbindlicher zu beantworten? Um mit der Charakterisierung von Sprache und Musik als komplexe Codesysteme nicht zu kurz zu greifen, müssen sie wenigstens in Umrissen in die Gesamtheit des menschlichen Verhaltens eingeordnet werden.

Erwerb und Gebrauch der natürlichen Sprache gehören zu den unabdingbaren Komponenten des menschlichen Sozialisationsprozesses, die Fähigkeit dazu ist eine artspezifische Disposition des Menschen. Vermutlich gilt etwas Analoges für den Erwerb und die Ausübung fundamentaler musikalischer Verfahren.¹⁸ Insofern Musik und Sprache akustische Kommunikationsprozesse organisieren, liegt es nahe, ihre gemeinsame Wurzel in der vormenschlichen Lautgebung zu sehen, die Ausdruck interner Zustände war und so zugleich der Verhaltensregulation zwischen den Individuen diente. Auch wenn diese Vermutung zutreffen sollte, darf sie nicht zu geradlinig interpretiert werden. Die Wandlungen der Funktion und der organismischen Grundlagen der akustischen Kommunikation, die sich bei der Herausbildung des Menschen abgespielt haben, sind bislang nicht in einer zusammenfassenden Theorie darstellbar.¹⁹ Aufzuklären ist vor allem die Wechselwirkung akustischer und optischer Systeme, also der Lautgebung und der Gebärden und Mimik. Als Resultat der Entwicklung ist jedenfalls konstaterbar, daß die Produktion und Perzeption sprachlicher Signale auf eigenen, spezialisierten Mechanismen beruht, deren Wirkungsweise sich von denen für andere akustische Signale deutlich unterscheiden läßt.²⁰ Der gemeinsame Ursprung von Sprache und Musik hat sich jedenfalls in verschiedene Verarbeitungssysteme der Lautform aufgelöst, denen auch unterschiedliche Arten der Strukturbildung zugeordnet sind.

Wichtiger als dieser Punkt ist aber die Verschiedenartigkeit der Bedeutungsbereiche. Betrachten wir zuerst die Sprache. Sie läßt sich ganz generell charakterisieren als Verbindung von kommunikativen und kognitiven Strukturen. Die Muster der Lautbildung sind in ihr verkoppelt mit den Strukturen der kognitiven Klassifizierung und Verarbeitung von Umwelterfahrungen. Kurz, die Sprache verbindet begriffliches Denken mit akustischer Kommunikation. Aus der Verzahnung dieser beiden Systeme, deren Vorformen im

Tierreich unabhängig nebeneinander stehen, ist sie phylogenetisch mit der Entwicklung des Menschen hervorgegangen in einem einzelnen noch klärungsbedürftigen Zusammenspiel biologischer und sozialer Prozesse. Die Verkoppelung hat beide Systeme zu prinzipiell neuer Leistungsfähigkeit geführt, wobei zwei Tendenzen zusammenwirken: Einerseits werden die beiden Systeme von ihrer strikten Bindung an feste Bedürfnisstrukturen, an motivationale Muster entkoppelt. Die Lautmuster drücken nicht mehr unmittelbar Angst, Drohung, Werbung, Warnung, Revierbehauptung usw. aus. Und die Klassifikationsstrukturen werden durch ihre zusätzliche, lautliche Repräsentanz losgelöst von unmittelbaren Situationsbedingungen und so auch außerhalb von Aktionen direkter Bedürfnisbefriedigung frei verfügbar. Zum anderen differenzieren und bereichern die Strukturen beider Systeme sich auf Grund dieser Verkoppelung wechselweise im Hinblick auf neue, vermittelte Situationszusammenhänge.

Objekte oder Individuen, Eigenschaften, Vorgänge, Beziehungen und schließlich Sachverhalte als gegliederte Verbindungen dieser Bestandteile werden damit als Kategorien der kognitiven Strukturbildung unterscheidbar, und als Bedeutung von Namen, Prädikaten und Propositionen sprachlich codiert. Die Art dieser Codierung in Form von Wörtern und syntaktisch komplexen Strukturen wird uns noch beschäftigen.²¹ Zunächst ist nur festzuhalten: Die Fähigkeit, Strukturen der angedeuteten Art auszubilden und in fixierter Weise mit Lautmustern zu belegen, macht die Etablierung sprachlicher Zeichensysteme möglich. Diese stellen den Zusammenhang zwischen akustischen Signalen und situationsgebundenen Mitteilungen her, ohne daß die Bedeutung der Zeichen stets mit der intendierten Mitteilung identisch sein muß. Die Codierung kognitiver Strukturen durch Lautmuster gibt vielmehr auch den Spielraum für die oben in der Hypothese (H 2) postulierten pragmatischen Prinzipien und die Differenz zwischen Bedeutung und kommunikativer Wirkung einer Äußerung frei.

Bleibt hinzuzufügen, daß der Gebrauch sprachlicher Zeichen sich natürlich keineswegs außerhalb motivationaler Zusammenhänge abspielt, sondern durch Absichten und Bedürfnisse gelenkt wird und von Emotionen und Affekten begleitet sein kann. Letztere drücken sich, wie jedem bewußt ist, in der Art der Signalproduktion, etwa in heftigem, schleppendem oder zurückhaltendem Sprechen aus. Die akustischen Eigenschaften, die diese Wirkung vermitteln – und sie kann durchaus kommunikativ beabsichtigt werden –, sind aber nicht sprachlich codiert, sie wirken auf eine Weise, die sogleich bei der Musik noch näher zu betrachten ist. Das hier Gemeinte ist strikt zu unterscheiden von der Möglichkeit, mit sprachlichen Mitteln über Emotionales zu reden: Es ist etwas anderes, ob jemand in sachlichem Ton sagt *Ich bin empört über Ihre Behauptung* oder ob er in aufgebrachtem Ton sagt *Das können Sie nicht behaupten!* Im ersten Fall wird eine Emotion begrifflich codiert, und nur der zweite Fall macht sie direkt sichtbar. (Die Mitteilung als Ganzes kann dabei in beiden Fällen ungefähr die gleiche sein.)

Man sieht, die Mitteilung, die mit einer sprachlichen Äußerung gemacht wird, ist auch aus diesem Grund nicht immer voll durch die verwendeten sprachlichen Zeichen und ihre Bedeutung festgelegt.

Nun zum musikalischen Code. Ich beginne mit einer negativen, aber zunächst kaum strittigen Feststellung: Musik codiert keine begrifflichen Strukturen, sie ist kein Zeichensystem für kognitive Bedeutungen. Da ich diese Feststellung für zentral halte und ganz prinzipiell festhalten werde, will ich sie nach zwei Seiten hin erläutern.

Sie besagt einerseits, daß Musik als Zeichensystem im Charakter und in der Struktur ihrer Bedeutungen sich grundsätzlich von der Sprache unterscheidet. Sie ist demnach weder eine „weniger scharfe“, ungenauere oder allgemeinere Codierung kognitiv-begrifflicher Strukturen noch ein auch nur gelegentlich oder zusätzlich mit solchen Bedeutungen versehenes Zeichensystem. Wo musikalische Mittel solche Bedeutungen zu haben scheinen – am deutlichsten etwa in militärischen Signalen, aber auch bei sehr vielen textgebundenen Melodien oder Themen –, da handelt es sich gerade nicht um musikalische Bedeutungen. Es liegt vielmehr ein gewissermaßen umgekehrter Fall der Erscheinung vor, die ich soeben für sprachliche Äußerungen erörtert habe: Die mit einer musikalischen Äußerung gemachte Mitteilung ist durch die Bedeutung der musikalischen Mittel nicht voll determiniert, es wirken an ihr weitere Komponenten mit, so z. B. die Kenntnis eines (bei anderen Gelegenheiten) mit dem musikalischen Zeichen verbundenen Textes.²²

Andererseits besagt die Feststellung, daß musikalische Zeichen keine kognitive Bedeutung haben, keineswegs, daß musikalische Äußerungen keine kognitiven Wirkungen haben. Ich werde vielmehr zeigen, daß gerade der Übergang von Bedeutungen B zu Wirkungen $W(k)$ in der Musik ein komplizierter kognitiver Prozeß sein kann. In ihm liegt ein großer Teil dessen, was zumeist mit dem Ausdruck „Musikverständnis“ gemeint wird.

Es genügt nicht zu sagen, was musikalische Bedeutungen nicht sind. Die positive Bestimmung kann bei der Überlegung ansetzen, daß Musik und Sprache sich (gewiß über verwickelte Zwischenstufen) aus vormenschlicher Kommunikation herausdifferenziert haben. Wenn nun die Sprache dabei zur spezifischen Repräsentation kognitiver Strukturen geführt hat, liegt folgende Annahme nahe: Die Herausbildung musikalischer Codes macht parallel dazu die Komponenten verfügbar, die durch die Sprache von der festen Bindung an klassifikatorische Merkmale freigesetzt werden. Es sind dies vor allem die affektiven und emotionalen Strukturen, die Einstellungen und Erwartungen, die ursprünglich an bestimmte Situationstypen und die in ihnen wirkenden Bedürfnisse und Motivationen gebunden waren. Die Musik verbindet in diesem Sinn Lautmuster mit emotiven Strukturen. Und wie die Sprache die kognitiven Strukturen durch ihre spezifische Codierung differenziert und ausgestaltet, so macht die Musik die Differenzierung emotiver Strukturen möglich. Die Entstehung sprachlicher und musikalischer Codes erscheint so als Herausbildung einander ergänzender Aspekte im Verlauf

der Anthropogenese, die einerseits kognitive, andererseits emotive Prozesse sozial zugänglich und also gesellschaftlich wirksam werden lassen.

Die Charakterisierung des Bedeutungsbereichs musikalischer Zeichen mag unbefriedigend erscheinen im Vergleich zu der der sprachlichen Bedeutungen. Zwei Schwierigkeiten sind hier in Rechnung zu stellen. Erstens können wir zwar mit sprachlichen Mitteln über Emotionen, Affekte, Einstellungen und die von ihnen begleiteten Motivationen sprechen. Aber diese Mittel wie auch unsere Begriffe und geläufigen Kenntnisse sind hier viel weniger artikuliert als im Bereich kognitiver Strukturen. Und zweitens ist auch die wissenschaftliche Analyse dieses Bereichs weit weniger ausgearbeitet als die der kognitiven Strukturen und Prozesse. Diese Schwierigkeiten dürfen nicht verdecken, daß wir hier einen komplex strukturierten Bereich interner Zustände und Prozesse zu identifizieren haben, der in ganz verbindlichem Sinn durch musikalische Lautmuster codiert werden kann und also deren Bedeutungen bereitstellt. Die Art dieser Codierung wird noch genauer zu bestimmen sein.

Die globale Charakterisierung der Bedeutungsbereiche von Sprache und Musik, die ich gegeben habe, führt nun leicht zu der Vorstellung: Sprache teilt Gedanken mit, Musik vermittelt Gefühle. In dieser Verkürzung aber verfehlt sie das Ziel, und wir können den Grund benennen: Kognitive und emotionale Strukturen bilden (im wesentlichen) die Bedeutungsbereiche sprachlicher bzw. musikalischer Zeichen, legen damit aber nicht vollständig die mit diesen Zeichen vermittelbaren Wirkungen fest. Dies ist zugleich der Grund dafür, daß die Wirkung musikalischer Äußerungen nicht durch den Zeichencharakter von Musik ausgeschöpft werden kann.

7. Sprachkenntnis und Musikkenntnis

Der vorläufigen Umschreibung der Form- und Bedeutungsbereiche will ich nun drei Erläuterungen zum Charakter der Kenntnisse folgen lassen, auf denen sprachliche und musikalische Codes beruhen. (Sprach- und Musikkenntnis meint dabei nicht sprach- oder musikgeschichtliches Wissen, die Kenntnis einzelner Werke oder die Fähigkeit Noten zu lesen oder ein Instrument zu spielen. Die Frage ist vielmehr: Auf welcher Kenntnis beruht das Verstehen sprachlicher und musikalischer Zeichen.)

Erstens wissen wir bereits, daß Sprache und Musik nicht durch abgeschlossene Zeichenvorräte, sondern nur durch Zeichenbildungsverfahren zu charakterisieren sind. Im Gedächtnis sind nicht alle möglichen komplexen Zeichen gespeichert, sondern nur bestimmte Grundelemente und auf sie anwendbare kombinatorische Operationen oder Regeln.²³ Das schließt keineswegs aus, daß man eine große Zahl ganzer Kompositionen oder literarischer Texte vollständig im Gedächtnis speichern kann. Es geht jetzt aber nicht um das Behalten oder Wiedererkennen bereits „bekannter“ Texte oder Kompositionen, sondern um das Erfassen immer wieder neuer Gebilde. Und um das

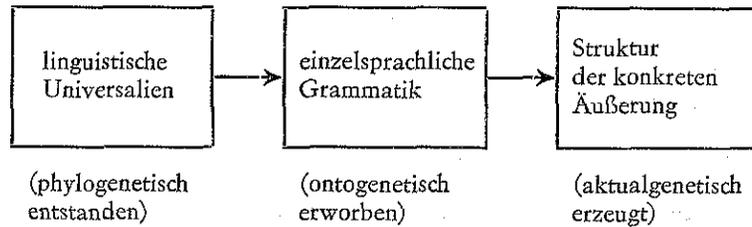
zu ermöglichen, müssen Musik und Sprache nach einem zunächst nur auf die Sprache bezogenen Wort Wilhelm von Humboldts „*von endlichen Mitteln einen unendlichen Gebrauch machen*“. Für eine gegebene Sprache wird die Gesamtheit der Mittel, die die Form und Bedeutung aller möglichen Ausdrücke zugänglich machen, durch die Grammatik dieser Sprache beschrieben. (Grammatik hat dabei einen weiten, aber genau bestimmten Sinn. Zu ihr gehört insbesondere der Aufbau des gesamten Wortschatzes.) Ein Ziel meiner Überlegungen ist es, die Notwendigkeit eines parallelen Begriffs im Bereich der Musik zu begründen und Gesichtspunkte zu seiner Formulierung zu diskutieren. Da dies nicht durch kurzschlüssige Übertragungen linguistischer Begriffe in die Musikanalyse geschehen darf, behalte ich den provisorischen Terminus „*musikalischer Code*“ für die Charakterisierung der Musikkenntnisse bei. Grammatiken und musikalische Codes beschreiben dann in expliziter Form jeweils besondere, komplizierte Systeme von Kenntnissen, die als solche meist impliziter, nicht bewußt kontrollierter Natur sind: Es handelt sich um Strukturen und Funktionen des Gedächtnisses, die weitgehend automatisch wirksam werden.

Die zweite Bemerkung betrifft die Grundlagen oder Herkunft dieses Gedächtnisbesitzes. Zwei Quellen sind hier zu unterscheiden. Die erste ist die phylogenetisch entstandene Ausstattung des Organismus, also die gattungsgeschichtlich erworbene und erblich fixierte Anlage des Menschen. Sie gibt gewissermaßen den Rahmen, das allgemeine Gerüst der hier erörterten Kenntnisse vor. Hierher gehört die Tatsache, daß bestimmte Lautstrukturen für sprachliche Zeichen in Betracht kommen und, wie erwähnt, auf spezifische Weise verarbeitet werden; sodann vor allem die Basis der komplexen konzeptuellen Strukturen, die das begriffliche Erfassen der Umwelt möglich machen; weiterhin die allgemeinen Funktionsmechanismen emotiver Prozesse mit ihrer Bindung an entsprechende physiologische Vorgänge; schließlich die Grundschemata kombinatorischer Operationen über Gedächtnisstrukturen, oder besser: die zerebrale Grundlage solcher Operationen. Die zweite Quelle ist die individuelle Erfahrung. Durch sie formen sich die Kenntnisse, die den konkreten sprachlichen bzw. musikalischen Gebilden entsprechen, mit denen ein Mensch in seiner sozialen Umwelt konfrontiert wird. Auf Grund der individuellen Erfahrungen und Lernprozesse werden die eben skizzierten allgemeinen Vorgaben aktiviert, mit Inhalt versorgt und zur Ausbildung der eigentlichen Gedächtnisstrukturen veranlaßt. Mit anderen Worten, das Zusammenspiel der biologischen Grundlagen und der individuellen Erfahrungen führt in einem verwickelten Prozeß zur Entstehung der Strukturen und Funktionen, die mit Sprach- und Musikkenntnis gemeint sind. Ausführlicher untersucht, wenn auch keineswegs völlig geklärt, ist dieser Prozeß für den Spracherwerb. Außer dem kommunikativen Umgang mit Sprache spielen dabei praktische Situationserfahrung, allgemeine kognitive Entwicklungsprozesse sowie extrakommunikative Sprachspiele eine wichtige Rolle. Es liegt auf der Hand, daß im Prinzip der Erwerb musikalischer

Kenntnisse ein dem Spracherwerb⁸ paralleles, wengleich im einzelnen durchaus andersartiges Untersuchungsfeld darstellt.

Drittens ergibt sich aus dem Gesagten, daß alles das, was durch individuelle Lernprozesse erworben und auf diesem Wege auch weitergegeben wird, Musik und Sprache zu gesellschaftlichen Erscheinungen macht. Das zu Lernende ist das Ergebnis gesellschaftlicher Tradition, die wiederum zur Geschichte der Gemeinschaft gehört, die die Kommunikationsprozesse ermöglicht und trägt. In diese Tradition müssen sich die Kommunikationspartner lernend einleben. In der Sprachkenntnis steckt mithin jeweils das Resultat historischer Entwicklungen, verfestigt in sprachlichen Elementen, Regeln und Strukturen. Und die Musikkenntnis enthält entsprechend das Ergebnis der Geschichte musikalischer Strukturbildungsmittel. Mit anderen Worten, eine Geschichte haben nicht nur musikalische oder literarische Werke, sondern vorab bereits die ihnen zugrundeliegenden Kenntnissysteme. Aus dieser Geschichte geht die Verschiedenartigkeit der Idiome hervor, während die genetisch verankerten Vorgaben ihre gemeinsamen Grundzüge bedingen.

Die generellen Grundlagen, die die Sprache betreffen, werden in der allgemeinen Linguistik analysiert und durch sogenannte linguistische Universalien theoretisch erfaßt. Das Ziel ist dabei, die für alle Sprachen gleichermaßen geltenden Strukturprinzipien zu bestimmen. Die Kenntnis einer Einzelsprache, beschrieben durch eine entsprechende Grammatik, ist dann jeweils eine besondere Ausprägung in dem durch die Universalien festgelegten Rahmen. Mit anderen Worten: Die linguistischen Universalien charakterisieren die Klasse aller auf natürliche Weise lernbaren Sprachen. Sie geben die Möglichkeiten an, denen gemäß jede Einzelsprache ihre Elemente, Regeln und Strukturen ausbildet. Die individuellen Erfahrungen, die der einzelne in der sozialen Umwelt macht, bestimmen dann den Inhalt der tatsächlich erworbenen Sprachkenntnisse, also die einzelsprachliche Grammatik. Und diese schließlich determiniert die Zeichenstruktur der einzelnen Äußerungen, die in jeweils konkreten Situationen gebildet und verstanden werden können. Damit haben wir drei Arten sprachlicher Strukturbildung unterschieden und aufeinander bezogen: Phylogenetisch haben sich die Grundlagen der Sprache insgesamt herausgebildet. Den strukturellen Aspekt dieser Disposition beschreiben die linguistischen Universalien. Ontogenetisch wird auf dieser Grundlage die Kenntnis jeweils konkreter Einzelsprachen erworben. Die Struktur dieser Kenntnis beschreibt die Grammatik. Aktualgenetisch werden auf Grund dieser Kenntnis konkrete Äußerungen gebildet und verstanden. Deren Struktur ist die Form und Bedeutung der erzeugten Zeichen. Abkürzend, das heißt unter Auslassung zahlreicher beteiligter Faktoren, läßt sich das in folgendem dreistufigen Schema darstellen:



Die gleichen Stufen der Strukturbildung sind prinzipiell auch in der Musik zu unterscheiden: Die phylogenetisch entstandenen Grundlagen möglicher musikalischer Systeme müssen durch musikalische Universalien beschreibbar sein. Auf diesen müssen die ontogenetisch zu erwerbenden musikalischen Codes beruhen, aus denen sich dann die Struktur der einzelnen musikalischen Gebilde ergibt.

Die Unterscheidungen, die ich zu erläutern versucht habe, haben sich als wesentlich für das theoretische Verständnis der Sprachstruktur und Sprachkenntnis erwiesen. Ich bin sicher, daß das gleiche für die musikalische Struktur gilt, wenn die Parallelen und die wesentlichen Unterschiede genau bedacht werden. Die Unterschiede zwischen Sprache und Musik betreffen dabei sowohl den Charakter der Kenntnissysteme wie den Anteil, den das individuell zu Erlernende an ihnen hat. Grob gesprochen: Sprachkenntnis setzt einen vielleicht nicht größeren, aber sehr andersartigen Anteil von individuellen Lernprozessen voraus als Musikkenntnis, wie die folgende Überlegung zeigt: Auch beim ersten Hören von exotischer Musik versteht man etwas mehr als bei der ersten Begegnung mit Äußerungen in einer unbekanntem Sprache. Dabei ist gewiß ziemlich unklar, was dieses „etwas mehr verstehen“ heißt. Aber soweit es nicht bloßes Mißverstehen ist, beruht es auf Zusammenhängen, die wirksam werden, ohne daß sie eigens erlernt werden. Damit hängt der unterschiedliche Charakter der Kenntnissysteme selbst zusammen. In ihm liegt unter anderem die schon erwähnte Tatsache begründet, daß sprachliche Äußerungen aus einer Sprache in eine andere übersetzbar sind, was in der Musik keine Parallele hat. Das heißt: Nur die Sprachkenntnis ist ein Mittel, etwas auszudrücken, das in einer anderen Sprache auf andere, aber gleichwertige Weise wiedergegeben werden kann. Es gehört nicht viel Phantasie zu der Überlegung, daß so grundlegende Funktionsunterschiede mit Verschiedenheiten der Struktur der Kenntnissysteme und ihrer universellen Grundlagen verbunden sind.

Ich wende mich nun den Eigenschaften sprachlicher und musikalischer Codesysteme im einzelnen zu. Da es mir dabei um die zusammenfassenden Eigenschaften von Sprache einerseits und Musik andererseits geht, bewegen sich die Überlegungen genaugenommen auf der Ebene der Universalien. Die Beispiele, durch die sie zu belegen sind, gehören jedoch notwendiger-

weise zu jeweils speziellen Codesystemen. Bei der folgenden Erörterung der verschiedenen Aspekte oder Ebenen halte ich mich nur bedingt an die Systematik, die innerhalb der Linguistik begründet und in wichtigen Teilen auch formal ausgearbeitet worden ist,²⁴ weil es mir mehr auf die Gegenüberstellung sprachlicher und musikalischer Erscheinungen als auf die innere Systematik sprachlicher Strukturen ankommt.

8. Zeitmuster

Jeder der Aspekte, die wir jetzt betrachten, bildet eine Strukturebene sprachlicher bzw. musikalischer Zeichen, die zu anderen Ebenen in systematischer Beziehung steht. Dabei ist eine Strukturebene genauer durch folgende Faktoren bestimmt:

- (i) Ein mehr oder weniger umfangreiches Inventar von Grundeinheiten, den Elementen der jeweiligen Ebene.
- (ii) Ein System von Verknüpfungs- und Transformationsoperationen, durch die komplexe Einheiten der jeweiligen Ebene gebildet werden.
- (iii) Ein System von Eigenschaften und Relationen, die die Einheiten der jeweiligen Ebene voneinander unterscheiden und zueinander in Beziehung setzen.

Die unter (iii) genannten Eigenschaften und Beziehungen bilden die Dimensionen, durch die die Bereiche der Formen und Bedeutungen in sich strukturiert sind. Wir werden sehen, daß in der Musik dabei graduell gegliederte Dimensionen eine dominierende Rolle spielen, während in der Sprache die binäre Gliederung, das Auftreten oder Fehlen einer Eigenschaft als Strukturprinzip Vorrang hat.

Die beiden ersten Ebenen betreffen die Organisation der akustischen Substanz, wobei es jedoch nicht um die physikalische Struktur der Signalereignisse geht, sondern um psychologische Erscheinungen, nämlich um die Möglichkeiten, Zeichenformen aufzubauen.

Als erstes betrachten wir die Zeitmuster, die Gliederung des Zeitablaufs. Ihre Grundeinheiten sind Zeitabschnitte unterschiedlicher Dauer, die zueinander in Proportionen gesetzt und gewichtet, also unterschiedlich bewertet werden. Aus ihrer Abfolge entstehen Rhythmus und Takt. Diese zeitliche Rasterung akustischer Signale korrespondiert mit Zeitmustern anderer organischer Abläufe, mit Puls und Atemrhythmus, mit der Zeitstruktur von Körperbewegungen, so daß sich in der Regel Entsprechungen einstellen, die weitgehend automatisch und unabhängig von Lernvorgängen funktionieren.

Insofern zeitlich organisierte Zeichen stets ein Zeitmuster aufweisen müssen, hat auch die Sprache an dieser Strukturierung teil. Sie kommt hier durch die Abfolge langer und kurzer Silben, gewichtet durch die Akzentstufen in Wörtern und Sätzen, zustande. Sie hat jedoch keinen eigenständigen Charakter, sondern ist determiniert durch Bedingungen in anderen Aspekten der

Sprachstruktur, und zwar auf Grund jeweils spezieller Regeln von unterschiedlicher Art. So springt im Deutschen z. B. der Akzent um auf Grund bestimmter Wortbildungserscheinungen in Fällen wie *Musik – musikalisch – Musikalität*, bleibt dagegen konstant in Fällen wie *Witz – witzig – Witzigkeit*, er ändert sich mit der Bedeutung in Sätzen wie *Sie ist auch witzig* gegenüber *Sie ist auch witzig* und unterliegt noch anderen Bedingungen in *Das ist eben so* gegenüber *Das ist ebenso*. Kurz, die rhythmischen Muster der Sprache sind von anderen Ebenen der Struktur abhängig, und zwar auf Grund verwickelter Regeln, die zum großen Teil erworben werden müssen als spezielle Ausprägung allgemeiner Möglichkeiten.

Zeitmuster bilden also keinen eigenständigen Strukturaspekt der Sprache. Sie können aber auf zwei von der Sprache nicht determinierte Arten in sprachlichen Äußerungen wirksam werden. Einmal, indem der sprachliche Rhythmus zusätzlichen Ordnungsbildungen unterworfen wird, die einem eigenen Code, nämlich dem der poetischen Strukturbildung, angehören. In der dadurch entstehenden metrisch gebundenen Sprache unterliegt die Abfolge der gewichteten Elemente der zeitlichen Rasterung besonderer Regeln, die sich auf sprachliche Einheiten beziehen, aber nicht zum sprachlichen Code gehören.²⁵ Zum anderen, indem das Sprechtempo als Symptom oder Ausdruck des allgemeinen Aktivitätszustands, speziell des emotionalen Status, in das Signalereignis eingetragene auf Grund automatisch wirkender psychophysischer Mechanismen.

In der Musik spielen die Zeitmuster eine eigenständige, sogar dominierende Rolle, sie hängen nicht von anderen Ebenen ab, sondern determinieren sie gegebenenfalls. Hand in Hand damit geht die wesentlich größere Differenzierungsmöglichkeit der Zeitmuster. Dies beginnt bereits bei den Grundeinheiten, also den Zeitintervallen, die eine Vielzahl unterscheidbarer Proportionen bilden können. Dabei sind feste, diskrete (ganzzahlige) Verhältnisse ebenso möglich wie kontinuierlich gleitende Werte. Durch ihre sequentielle Verknüpfung wird das zum Grundprinzip, was in der Sprache allenfalls eine Randerscheinung ist: die Unterscheidung eines periodischen Grundrasters – des Takts – von dessen variierender Ausgestaltung – dem Rhythmus. Beide können nach eigenen Regeln organisiert und aufeinander bezogen werden. Schließlich können die Muster insgesamt bestimmten Transformationen unterworfen, gedehnt oder verkürzt, beschleunigt und verlangsamt werden. Das sind Operationen, die kein sprachliches Analogon besitzen. Sofern die genannten Regeln und Operationen lernabhängige Spezialisierungen aufweisen, betreffen sie die Ausgestaltung der Zeitmuster und die Beziehung ihrer Bestandteile untereinander, nicht deren Abhängigkeit von anderen Strukturebenen.

Nach dieser groben Charakterisierung der Zeitmuster in Sprache und Musik noch zwei Bemerkungen zu ihrer Interaktion. Erstens liegt auf der Hand, daß sie eine fundamentale Rolle spielt bei der Verkoppelung von Musik und Sprache im Gesang, das heißt, bei Signalereignissen, die sowohl eine

sprachliche wie eine musikalische Codierung tragen. Die lokale Ausgestaltung musikalischer Zeitmuster wird dann vom sprachlichen Rhythmus mitbestimmt, dieser durch die musikalischen Muster konturiert, verstärkt oder auch maskiert. Das musikalische Zeitmuster absorbiert dabei gewissermaßen den oben erwähnten Ausdruck des emotionalen Status und organisiert ihn nach musikalischen Regeln. Zweitens zeigt sich, daß metrisch gebundene Sprache gerade dadurch poetisch wird, daß sie ein musikalisches Prinzip aufgreift, nämlich die eigenständige Durchorganisation von Takt und Rhythmus.

9. Segmentstrukturen

Natürlich sind Zeitmuster nur in Gedanken abzuheben von dem, was sie gliedern. Die „Ausfüllung“ der Zeitmuster bildet unsere zweite Strukturebene. Sie ordnet das Feld der akustischen Möglichkeiten in Elemente und Muster. Vorgegeben sind zunächst drei Bedingungsgefüge: die Physik der Schallschwingungen, der menschliche Artikulationsapparat, im Bereich der Musik erweitert durch Instrumente, und der Gehörsinn mit den an ihn anschließenden Verarbeitungsstufen. Die Grundeinheiten dieser Strukturebene will ich Segmente nennen. Sie sind gewissermaßen die kleinsten abtrennbaren Bündel von Merkmalen, auf denen die Form möglicher Zeichen beruht. In grober Näherung werden sprachliche Segmente durch Buchstaben, musikalische durch Noten wiedergegeben. Dabei besteht in der Musik eine starke Affinität zwischen den Segmenten und den Grundelementen der Zeitmuster: In der Regel besetzt ein Segment gerade eine Stelle in einem Zeitmuster (weshalb Noten meist gleichzeitig das segmentale Merkmal der Tonhöhe wie den Stellenwert im Zeitmuster darstellen). In der Sprache besteht diese Koppelung nicht: Silben sind die Grundeinheiten der Zeitmuster, bestehen aber in der Regel aus mehreren Segmenten, nämlich einem vokalischen Kern und mehreren möglichen nichtvokalischen Satelliten.

Die Charakterisierung von Segmenten als kleinsten Merkmalsbündeln hebt in gewissem Sinn ihren elementaren Charakter wieder auf. Sie sind zwar Grundeinheiten, insofern sie gewissermaßen kleinste wahrnehmbare Signalereignisse determinieren, nämlich sprachliche Laute und musikalische Töne oder Geräusche, aber sie sind in sich zusammengesetzt aus Merkmalen wie Stärke, Tonhöhe, Artikulationsart, Klangfarbe etc. Ich verdeutliche diesen einfachen, aber folgenreichen Gedankenschritt an einem sehr vereinfachten Beispiel. Wenn in einer Sprache, wie z. B. im Deutschen, die Segmente *b*, *p*, *d*, *t* unterschieden werden, dann lassen sich ihre Beziehungen zueinander durch zwei Dimensionen angeben: Die Stimmbeteiligung faßt *b* und *d* gegenüber *p* und *t* zusammen, die verschiedene Artikulationsstelle faßt *b* und *p* gegenüber *d* und *t* zusammen. Ziehen wir weiterhin die Segmente *s* und *f* in Betracht, dann kommt zu den beiden genannten Dimensionen eine dritte

hinzu: die Artikulationsart, die Verschuß von Engebildung unterscheidet. Auf diese Weise können die Segmente auf Merkmalsdimensionen bezogen werden, was sich folgendermaßen darstellen läßt:

	b	p	f	d	t	s	g	k	ch
stimmhaft	+	-	-	+	-	-	+	-	-
verschußbildend	+	+	-	+	+	-	+	+	-
labial	+	+	+	-	-	-	-	-	-
velar	-	-	-	-	-	-	+	+	+

Dieser Ausschnitt aus dem Segmentsystem des Deutschen ordnet neun Segmente in vier Dimensionen ein, auf denen jeweils zwei Ausprägungsgrade, nämlich + und -, unterschieden sind. Die Unterscheidungsmöglichkeiten in einem solchen System können vergrößert werden, indem entweder die Anzahl der Dimensionen oder die Anzahl der Werte auf einer Dimension vergrößert wird. Während die letztere Möglichkeit in der Sprache eine untergeordnete Rolle spielt, ist sie für die Musik von zentraler Bedeutung: Die Tonhöhe ist im hier erläuterten Sinn eine Dimension mit einer großen Zahl verschiedener (aber wiederum systematisierter) Werte. Prinzipiell aber ist einsehbar, daß musikalische Segmente auf die gleiche Weise durch ein System von Dimensionen mit entsprechenden Werten bestimmt werden können, wie ich das in dem angegebenen Schema für sprachliche Segmente angedeutet habe. Dimensionen sind dabei u. a. Tonhöhe, Lautstärke, (relative) Dauer, Klangfarbe. Mit anderen Worten: jedes Segment wird durch seine Merkmale in verschiedene Dimensionen eingeordnet und so zu anderen Segmenten in Beziehung gesetzt. So gesehen sind die Segmente Knotenpunkte jeweils verschiedener Dimensionen, die ihre spezifischen Eigenschaften ausmachen. Die wirklich elementaren Faktoren, mit denen wir es zu tun haben, sind diese Dimensionen und die unterschiedlichen Werte, die in ihnen unterschieden werden können. Als eigentliche Basis der Segmentstruktur erhalten wir damit, abstrakt gesprochen, einen Merkmalsraum, dessen Dimensionen - Tonhöhe, verschiedene Artikulationsarten, Tonstärke etc. - die Eigenschaften möglicher Segmente definieren. Und die Segmente sind nur Bündelungen, die diese Eigenschaften auf die sequentiell organisierten Plätze der Zeitmuster verteilen.²⁶

Die Sprache besetzt nun in dem Bereich, der durch die Gesamtheit der akustisch wahrnehmbaren Merkmalsdimensionen gebildet wird, einen sehr speziellen Ausschnitt. Etwa ein Dutzend hauptsächlich binärer, also nicht gradueller Merkmale legen die Möglichkeit fest, sprachliche Segmente voneinander zu unterscheiden. Die akustische Charakteristik dieser Merkmale ist zum Teil äußerst komplex, liegt aber in einem vergleichsweise engen Bereich des wahrnehmbaren Frequenz- und Lautstärkepektrums. Die Sprache nutzt also für die Bildung von Lautmustern einen nicht sehr großen, aber stark differenzierten, vorwiegend binär gegliederten Bereich aus. Bedingt

durch die Entwicklung der Artikulationsorgane und der besonderen Funktion der Sprachwahrnehmung, ist dieser Bereich stammesgeschichtlich mit der menschlichen Sprachfähigkeit gekoppelt worden. In sich ist dieser Bereich weiterhin abgestuft in zentralere und periphere Distinktionen: Jede Sprache unterscheidet Vokale und Konsonanten, aber nicht jede Sprache unterscheidet z. B. palatalisierte und unpalatalisierte Konsonanten: Die erste Distinktion ist zentral, die zweite peripher. Mit anderen Worten, die verschiedenen Einzelsprachen treffen aus dem Arsenal der überhaupt unterscheidbaren Segmente historisch bedingte, dem Sprachwandel unterliegende Auswahlen, die jedoch nicht völlig beliebig sind, sondern bestimmten Bedingungen gehorchen. So können z. B. Reibelaute nur dann in stimmhafte und stimmlose unterteilt sein (etwa in *reißen* gegenüber *reisen*), wenn auch die entsprechenden Verschußlaute diese Unterteilung aufweisen (etwa in *Seite* gegenüber *Seide*). Kurz, jedes einzelsprachliche Segmentsystem ist in sich nach allgemeinen Gesetzen strukturiert, die einen relativ gut erforschten Teil der linguistischen Universalien ausmachen.

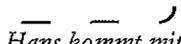
Der Ausschnitt, der für musikalische Lautmuster zur Disposition steht, ist anders aufgebaut. Es scheint, daß im Prinzip der gesamte Bereich des Frequenz- und Lautstärkepektrums in die musikalische Strukturbildung einbezogen werden kann. Er wird primär organisiert durch zwei Dimensionen: Tonhöhe und Lautstärke, zu denen Klangfarbe und Geräuschanteile hinzutreten. Die beiden Hauptdimensionen sind eindeutig graduell gegliedert, und zwar entweder in diskreten Stufen oder in gleitenden Übergängen. Für die Musik steht demnach ein sehr viel größerer Bereich zur Verfügung, der zwar nur durch eine geringe Zahl fester Dimensionen, jedoch eine große Zahl gradueller Merkmalswerte strukturiert ist. Die Menge der möglichen Distinktionen und der durch sie unterscheidbaren Segmente liegt damit um Größenordnungen über der der sprachlichen Segmente.²⁷ Hinzu kommt, daß Segmente mit veränderlichen Merkmalswerten - zu- oder abnehmende Lautstärke und gleitende Tonhöhe - gebildet werden können, eine Erscheinung, die über die Segmentgrenze hinaus eine für die Musik charakteristische Rolle spielen kann. Wie in der Sprache werden in diesem Bereich historisch bedingte Systeme gebildet. Pentatonische, Naturton- und temperierte Skalen, die Kirchentonarten und die Dur-Moll-Geschlechter mit ihrem jeweiligen Beziehungsgefüge sind Beispiele. Aber auch an unterschiedliche Systematisierung der Dynamik, der Klangfarben- und Geräuschmerkmale ist zu denken: verschiedene Musikkulturen und historische Epochen besitzen hier sehr verschiedene und keineswegs zufällige Inventare. Gesetzmäßigkeiten dieser Systembildung sind nur für den Tonhöhenbereich analysiert worden. Sie beruhen auf Proportionen zwischen Grund- und Obertonfrequenzen, genauer: auf deren perzeptiver Verarbeitung.

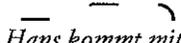
Umfang, Struktur und Ausnutzung der Merkmalsbereiche sind also in Sprache und Musik prinzipiell verschieden. Diese offensichtliche Tatsache wird noch deutlicher, wenn wir die Segmente innerhalb der aus ihnen ge-

bildeten Muster, also der eigentlichen Zeichenformen betrachten. Die bereits in der internen Charakteristik der Segmente angelegte Verschiedenartigkeit breitet sich bei ihrer Verknüpfung sozusagen über die Lautmuster als Ganzes aus. Auf eine sehr vereinfachte Formel gebracht: Die Merkmale sprachlicher Segmente haben fast ausschließlich eine lokal distinktive Funktion, die musikalischen Merkmalsdimensionen dagegen spielen eine Zusammenhang stiftende Rolle.

Der Aufbau von Mustern aus Segmenten geschieht in der Sprache in mehreren Schritten: Wie erwähnt, werden Segmente zu Silben verbunden. Aus diesen werden dann Wörter gebildet. Ein weiterer Verknüpfungsprozeß fügt Wörter zu komplex organisierten Folgen zusammen. Dieses Grundschema und bestimmte Prinzipien seiner Ausgestaltung gelten universell, innerhalb dieser Prinzipien bildet jede Einzelsprache ihr besonderes Regelsystem aus. Solche Systeme enthalten zwei Typen von Regeln: phonologische und syntaktische Regeln. Die ersteren steuern die Segmentkombinatorik innerhalb der Wörter. So kann im Deutschen z. B. am Silben- oder Wortende kein stimmhafter Geräuschlaut wie *b*, *d*, *g*, *w* stehen. Deshalb wird das *b* des Wortes *bleiben* stimmlos gesprochen in der Form *bleib*, und ein englisches Wort wie *Code* nimmt im Deutschen den gleichen Auslaut an wie *tot*. Ein anderes Beispiel für die oft verwickelten Regeln dieses Typs bilden die in vielen Sprachen geltenden Formen der „Vokalharmonie“, die innerhalb eines Wortes nur bestimmte Vokalkombinationen zulassen.

Die syntaktischen Regeln, die die Verknüpfung von Wörtern zu Phrasen und deren Integration zu größeren Komplexen determinieren, bilden ihrem Charakter nach bereits eine neue Strukturebene, in der gesondert zu erörternde Faktoren ausschlaggebend sind. Die Merkmale der Segmente spielen dabei keine Rolle mehr, die auf sie bezogenen Regeln reichen nicht über die Wortgrenze hinaus. Eine Zwischenstellung nimmt lediglich die Intonation ein, die das Tonhöhen- und zum Teil das Betonungsmerkmal direkt mit syntaktischen Bedingungen koppelt, so z. B. in der Unterscheidung von Frage und Aussage:

(a)  *Hans kommt mit?*

(b)  *Hans kommt mit.*

Für die Bildung musikalischer Segmentfolgen ist der zuletzt genannte Punkt, die Nutzung des Tonhöhenverlaufs, so zentral, daß man ihn leicht für ihr eigentliches Kernstück hält: die Bildung von melodischen Gestalten auf Grund des Tonhöhenmerkmals der Segmente. Die Tonhöhenwerte werden – im Grenzfall bei Konstanz aller anderen Merkmale – auf die Plätze des Zeitmusters verteilt und stellen eine entsprechende Beziehung zwischen ihnen her. Ein Beispiel wie der Beginn von Bartóks *Musik für Saiteninstrumente, Schlagzeug und Celesta* (Takt 1–4) läßt weiterhin erkennen, daß dabei wie

in der Sprache stufenweise Einheiten höherer Ordnung gebildet werden können. Vier Teilfolgen bilden hier eine größere Einheit, ein Thema, das aus vier (verwandten) Motiven besteht. Dieser Aufbau ist jedoch kein generelles Prinzip, und allgemeiner gesagt: Es ist nicht möglich, für die Musik ein Stufenschema zu konstatieren, das der sprachlichen Rangreihe Segment – Silbe – Wort – Phrase – Satz entspricht. Daß der Schritt vom Segment zur Silbe für die Sprache charakteristisch ist, hatte ich bereits erwähnt. Entscheidender ist, daß es kein musikalisches Analogon zum Wort und folglich auch nicht zu seinen Verknüpfungsbedingungen gibt. Eine Analogie zwischen Wörtern und dem, was gemeinhin „Motiv“ genannt wird, ist grundsätzlich irreführend. Denn während alle sprachlichen Lautfolgen systematisch in Wörter zerlegbar sind, bestehen musikalische Strukturen keineswegs aus klar abtrennbaren, aneinandergereihten Motiven. Ein Blick auf nahezu beliebige musikalische Gebilde macht das deutlich: Techniken des „Fortspinnens“, der Übergangsbildung usw. ergeben keine Zerlegung eines Ablaufs in Motive oder Themen. Dies hängt, wie im weiteren zu erörtern sein wird, fundamental mit den unterschiedlichen Codierungsprinzipien von Sprache und Musik zusammen, und es führt zu den ganz verschiedenartigen Funktionen der syntaktischen Regeln in Musik und Sprache.

Wie das Tonhöhenmerkmal, so bildet auch die Dimension der Lautstärke über die Segmente hinausgreifende Zusammenhänge, und zwar einerseits durch Zusammenfassung von Segmentfolgen unter einem festgehaltenen oder aber zu- bzw. abnehmenden Lautstärkewert und andererseits durch lokale Akzente, die einzelne Segmente zur Umgebung in Kontrast setzen. Schließlich bilden auch Klangfarbe und Geräuschmerkmale Zusammenhänge, die Segmentfolgen binden oder gliedern.²⁸ Kurz, die musikalischen Merkmalsdimensionen stellen verschiedene, aber aufeinander bezogene Zusammenhänge her zwischen den Segmenten, die die Zeitmuster ausfüllen. Die Veränderungen von Segment zu Segment können dabei kontinuierlich oder diskret sein.

Zu den Kombinationsmöglichkeiten kommt nun für die Musik noch die Gleichzeitigkeit als eigenes Prinzip hinzu. Dies führt zu simultan verlaufenden Segmentfolgen, also zur Mehrstimmigkeit. Werden dabei bestimmte regelmäßige Beziehungen zwischen den Tonhöhen simultaner Segmente hergestellt, ergibt sich als eigenständige Kombinationsdimension die Harmonik.²⁹ Beide Erscheinungen, Mehrstimmigkeit und Harmonik, haben in der Sprache keinerlei Analogon: Mehrere Segmente, Silben oder Wörter können nicht simultan, sondern nur sequentiell zu einem sprachlichen Zeichen verknüpft werden. Gleichzeitig gesprochene Wörter ergeben Durcheinander, nicht eine komplexere Ordnung.³⁰ Es bleibt die Frage, welche universellen Bedingungen den Aufbau musikalischer Lautmuster begrenzen, welche Arten von Regeln in jeweils speziellen Musiksystemen ausgebildet werden können. Was ich hier rekapituliert habe, gibt nur lückenhafte Anhaltspunkte für eine genauere Antwort auf diese Frage. Erkennbar ist mindestens dies: Mu-

sikalische Segmente werden zu Folgen verknüpft, in denen Teilfolgen als Motive oder Themen (ich will mit einem möglichst neutralen Terminus von „Gestalten“ sprechen) ausgrenzbar sein können, aber nicht müssen. Die Segmentmerkmale stellen in diesen Folgen dimensionsgebundene Zusammenhänge her. Zusätzlich zur sequentiellen Verknüpfung ist simultane Ordnungsbildung möglich. Zweifellos sind über diese Feststellungen hinaus interessante universelle Prinzipien formulierbar. Ihre grundsätzliche Verschiedenheit von den Prinzipien der sprachlichen Lautmuster dürfte bereits aus den gemachten Andeutungen hinreichend deutlich hervorgehen. Der bisher betrachtete Strukturaspekt wird für die Sprache durch die Phonologie erfaßt. Sie ist als zusammenhängende Theorie formuliert und enthält relativ genaue Kenntnisse von den allgemeinen Strukturgesetzen des Bereichs, vom Aufbau und den Veränderungsmöglichkeiten einzelsprachlicher Systeme.³¹ Eine vergleichbare Theorie für die Lautmuster der Musik ist eine zumindest klar formulierbare Aufgabenstellung, die zahlreiche Befunde der systematischen Musikwissenschaft in einen umfassenderen Zusammenhang zu bringen hätte. Wichtig wäre für eine solche Theorie, daß sie, wie die Phonologie, den Zusammenhang der in ihren Bereich fallenden Strukturprinzipien mit denen des übergeordneten Gliederungsaspekts, nämlich der Syntax, klärt. Diesen bisher nur erwähnten Zusammenhang gilt es nun ins Auge zu fassen. Dafür ist es jedoch zuvor nötig, den Zusammenhang zwischen Form und Bedeutung von Zeichen näher zu bestimmen, denn ohne ihn läßt sich der Charakter syntaktischer Strukturen nicht einsichtig machen.

10. Codierungsformen

Ich habe die unterschiedlichen Prinzipien zu kennzeichnen versucht, nach denen sprachliche und musikalische Lautmuster gebildet werden können. Damit diese Lautmuster als Zeichenformen fungieren, müssen ihnen Bedeutungen beigelegt werden, d. h., sie müssen Codierungen von Bedeutungen sein. Um die unterschiedliche Codierungsart von Sprache und Musik bestimmen zu können, führe ich einige Unterscheidungen ein.

Die erste ist die zwischen motivierter und arbiträrer Codierung. Ein Zeichen Z ist motiviert durch ein Prinzip P , wenn seine Bedeutung B auf Grund von P aus seiner Form F abgeleitet werden kann. Andernfalls ist es arbiträr. Mit anderen Worten: motivierte Zeichen beruhen auf einer vorgegebenen Grundlage für den Zusammenhang zwischen Form und Bedeutung. Fehlt eine solche Grundlage, sind die Zeichen unmotiviert oder arbiträr.

Ein Beispiel für unmotivierte Zeichen ist der bereits erwähnte Morse-Code: Es gibt keinerlei Grundlage, aus der abzuleiten wäre, daß $-$ die Bedeutung a hat und nicht d oder f . Arbiträre Zeichen sind auch die Nationalflaggen, die Verkehrszeichen und die Ziffern, um nur einige Systeme zu nennen. Für unmotivierte Zeichen muß durch Konvention festgelegt sein,

welche Form welcher Bedeutung zugeordnet ist. Sie heißen deshalb auch konventionelle Zeichen.

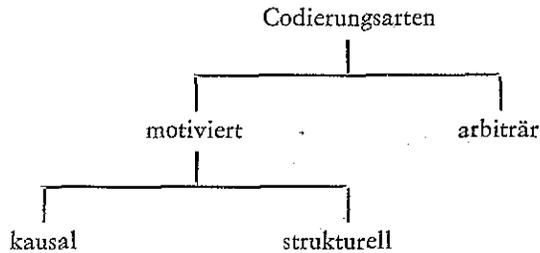
Motivierte Zeichen – damit komme ich zur zweiten Unterscheidung – werden je nach ihrem Motivierungsprinzip P seit Peirce (1932) gewöhnlich in indexikalische und ikonische Zeichen unterteilt.

Indexikalische Zeichen kommen durch einen Kausalzusammenhang zwischen Bedeutung und Form (genauer: Bedeutung und Signalereignis) zustande. Das Signal ist hier eine (kausal verursachte) Folge der Bedeutung. In diesem Sinn verweisen etwa Rötung des Gesichts und erhobene Stimme auf Erregung, Tränen auf Schmerz, aber auch rote Hautflecken auf Masern oder unkontrollierte Bewegungen auf Trunkenheit. Die Signale sind Symptome oder Anzeichen dessen, was sie bedeuten. Das Motivierungsprinzip P ist die Gesamtheit der Erfahrungen, die man mit den einschlägigen Kausalverhältnissen gemacht hat. Man sieht sofort, daß es sich hier nicht um rein kommunikative Zeichen handelt, da das Auftreten der Ursache und der entsprechenden Wirkung im allgemeinen nicht den in Abschnitt 4 erörterten Kommunikationsbedingungen unterliegt. (Man bekommt nicht rote Flecken, damit der Empfänger daraus entnimmt, daß man Masern hat.)³² Da aber Symptome auf ihre Ursachen verweisen und damit Verwandtschaft zum Verweisungscharakter von Zeichen besteht und da Symptome außerdem in gewissen Grenzen auch gezielt erzeugt werden können – gehobene Stimme oder unkontrolliert wirkende Bewegungen etwa –, kann Kausalität auch zum Codierungsprinzip kommunikativer Zeichen werden. Eine besondere Rolle kommt ihnen überdies für die Entstehung der oben erwähnten extrakommunikativen Wirkung $W(n)$ eines Signals zu.

Ikonische Zeichen beruhen dagegen auf der Ähnlichkeit zwischen dem Signal und dem, worauf es verweist. Etwas strenger und allgemeiner: Das motivierende Prinzip P besteht in hinreichend allgemeinen Korrespondenzen zwischen strukturellen Eigenschaften der Form und der Bedeutung. Ein wohlbekannter Bereich ikonischer Zeichen ist die zweidimensionale Darstellung räumlicher Objekte und Sachverhalte. Als Motivierungsprinzip fungieren dabei die Regeln der perspektivischen Projektion. Die relative Größe dargestellter Figuren als Codierung ihrer Wichtigkeit, wie sie unter anderem in der romanischen Buchmalerei anzutreffen ist, ist ein anderes Beispiel, dessen Motivierungsprinzip mit dem der Perspektive überdies unvereinbar ist. Ein drittes Beispiel ist die Zeitdarstellung durch die Zeigerstellung der Uhr, deren Bewegung dem Ablauf entsprechender Zeitintervalle korrespondiert. (Im Unterschied dazu beruhen Digitaluhren auf arbiträr-konventioneller Codierung der Zeit durch Zahlen, deren Form keine Ähnlichkeit zu den Zeitintervallen hat.) Ikonische Zeichen im akustischen Bereich sind onomatopoeische Wörter wie *quaken*, *gackern*, *quietschen*, *summen* usw. oder musikalische „Lautmalereien“ wie z. B. die Vogelstimmen in der *Szene am Bach* der *Sechsten Sinfonie* von Beethoven. Ein Wort wie *Kuckuck* zeigt schließlich eine Kombination von zwei Motivierungen: seine Lautstruktur ist ein

ikonisches Zeichen für ein akustisches Ereignis, das seinerseits als Symptom einer Vogelart genommen wird.

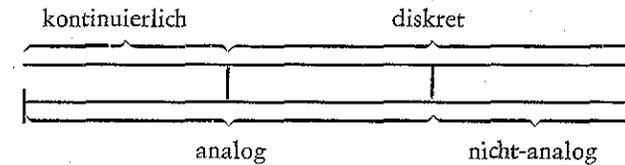
Der entscheidende Punkt bei ähnlichkeitsmotivierten Zeichen ist die Korrespondenz zwischen der Struktur ihrer Form und ihrer Bedeutung. Ich will diese Codierungsart deshalb verallgemeinernd „strukturell motiviert“ nennen und die diskutierten Unterscheidungen in folgendem (vorläufigen) Schema zusammenfassen:



Unter dem Einfluß der Computertechnik hat sich die Gegenüberstellung von analoger und digitaler Codierung eingebürgert. In Digitalrechnern werden Informationen durch diskrete Schaltstellungen, in Analogrechnern durch kontinuierlich variierbare Werte einer physikalischen Größe repräsentiert. In den damit verbundenen Codierungsarten werden zwei Bedingungen kombiniert: Analoge Codierung setzt erstens voraus, daß eine physikalisch realisierbare Formdimension und eine Bedeutungsdimension in funktionale, also motivierte Beziehung gebracht werden können, und zweitens, daß diese Dimensionen kontinuierliche Werte annehmen können. Digitale Codierung beruht dagegen erstens auf einer im Prinzip arbiträren Zuordnung und zweitens auf diskreten Strukturen. Hält man diese beiden Bedingungen auseinander, so ergeben sich zwei Unterscheidungen: analoge gegenüber nicht-analoger und kontinuierliche gegenüber diskreter Codierung.

Analoge Codierung ist ein Spezialfall motivierter Zuordnung, bei der in der Form und der Bedeutung der Zeichen jeweils mindestens eine Merkmalsdimension ausgezeichnet ist, für die gilt: Die Werte der Formdimension sind eine lineare Funktion der Werte in der Bedeutungsdimension. Vereinfacht: Form und Bedeutung von Zeichen stehen in proportionaler Beziehung zueinander. Analog in diesem Sinn sind z. B. fast alle graphischen Darstellungen variierender Größen wie Wachstum, Stimmenverteilung, Planerfüllung usw. durch Kurven oder Diagramme. Auch die Beziehung zwischen Erregungsgrad und Lautstärke der Stimme gehört – mit einiger Vorsicht – hierher: Je größer die Erregung, desto heftiger die Stimme. Analoge Zeichen sind also diejenige Teilklasse der motivierten Zeichen, bei denen die Motiviertheit auf graduierbaren Dimensionen beruht. Alle anderen Zeichen sind nicht-analog codiert.

Kontinuierliche Codierung ist nun ein Spezialfall der analogen, bei dem die Werte der Form- und Bedeutungsdimensionen stetig variieren können. Alle Zeichen, für die das nicht gilt, sind diskret codiert. Statt der Zweiteilung analog/digital erhalten wir damit folgende Dreiteilung:



Um die drei Möglichkeiten mit je einem Beispiel zu belegen, greife ich nochmals auf die Uhren zurück: Analog und kontinuierlich ist die Codierung bei „normalen“ Uhren mit (de facto) stetigem Zeigergang; analog und diskret ist die Codierung bei solchen Uhren, deren Zeiger „springen“, also nur diskrete Werte, etwa ganze Minuten, anzeigen können. Nicht-analog und diskret wird die Zeit in Digitaluhren codiert. Unmöglich ist die Kombination nicht-analog und kontinuierlich, da stetige Werte ja eine feste Dimension, und damit Analogie, voraussetzen.

Zwischen den verschiedenen Zeichenklassen, die mit den erläuterten Codierungsformen unterschieden werden können, ergeben sich nun eine ganze Reihe von Beziehungen, von denen ich die wichtigsten erwähnen will:

Alle kontinuierlich codierten Zeichen sind analog codiert.

Alle analog codierten Zeichen sind motiviert.

Diskrete Zeichen können motiviert oder arbiträr sein.

Motivierte Zeichen können analog oder nicht-analog, diskret oder kontinuierlich codiert sein.

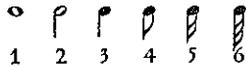
Arbiträre Zeichen sind immer diskret codiert.³³

Ich verzichte darauf, die erörterten Codierungsformen und die Beziehungen zwischen ihnen in einem Gesamtschema zusammenzufassen, und zwar weniger wegen der Kompliziertheit eines solchen Schemas, sondern vor allem, weil es die Tatsache nicht deutlich machen könnte, daß verschiedene Codierungsformen innerhalb eines Zeichensystems und sogar einzelner Zeichen verbunden sein können. Damit komme ich zu einigen zusammenfassenden Überlegungen.

Die erste betrifft die Kombination der Codierungsarten, die zwar auf unterschiedlichen Prinzipien beruhen, sich jedoch keineswegs immer gegenseitig ausschließen. So kann ein indexikalisches, also kausal motiviertes Zeichen zugleich Strukturgemeinsamkeiten von Form und Bedeutung aufweisen, also ikonisch sein. Wenn wir annehmen, daß die Lautstärke eines Aufschreis der Stärke des Schmerzes entspricht, der ihn verursacht, dann ergibt sich eine Klasse indexikalischer Zeichen, die zugleich ikonisch sind, da die Lautstärke die Stärke des Schmerzes abbildet. Auf ähnliche Weise analog

und kausal motiviert ist die Zeitanzeige durch Sand- und Sonnenuhren. Genereller läßt sich sagen, daß die in der analogen Codierung benutzte Entsprechung der Dimensionen immer zugleich kausal begründet sein kann, aber nicht muß. Dieser potentielle Doppelcharakter der Analogcodierung spielt, wie ich zeigen will, eine wichtige Rolle in der Musik.

Eine weitere Kombination von Codierungsprinzipien ergibt sich daraus, daß die Elemente und Merkmalsdimensionen, auf die sich die strukturelle Motivierung bezieht, ihrerseits arbiträr, also konventionell gewählt oder zumindest begrenzt sein können. So ist z. B. die Aneinanderreihung von Buchstaben der Abfolge der durch sie repräsentierten Lautsegmente analog, also motiviert, die Schreibrichtung von links nach rechts (wie in den europäischen Schriften) oder von rechts nach links (wie etwa im Arabischen) dagegen konventionell. Die bereits erwähnten graphischen Darstellungen von Verlaufskurven bieten beliebig weitere Beispiele. Eine besonders aufschlußreiche Verbindung verschiedener konventioneller und analoger Faktoren findet sich in der modernen Notenschrift:



Die Bedeutungsdimensionen dieser Zeichen bilden eine einfache Skala – jede Note bezeichnet den halben Zeitwert der vorangehenden –, die Codierung ist für 1, 2 und 3 rein konventionell, von 4 an folgt sie dagegen einem (theoretisch beliebig weiterführbaren) Analogieprinzip, dessen Formdimension (Zahl der Fähnchen bzw. Balken) aber konventionell gewählt ist. Nimmt man hinzu, daß die Stellung einer Note im Liniensystem zugleich eine analoge Codierung der Tonhöhe ist (bei wiederum konventioneller Wahl der Codierungsdimension), dann ist eine Viertelnote z. B. zugleich durch eine arbiträre und eine analoge Codierung mit konventionell fixierter Dimension determiniert. (All dies betrifft natürlich die Codierungsart der Notenschrift, nicht die der Musik.)

Außer dem gewissermaßen systematischen Zusammenwirken, das ich eben illustriert habe, gibt es schließlich auch ein bloßes Nebeneinanderstehen verschiedener Codierungsformen in einem Zeichensystem. Das System der römischen Zahlen ist ein Beispiel für diese Erscheinung: I, II, III sind ikonisch, V, X, L, C und M sind arbiträr, und die Verknüpfungen in IV, VI, VII usw. sind gemischt. Man sieht, die Unterscheidung der Codierungsprinzipien ist ebenso wichtig wie die Bestimmung ihrer Verbindung, wobei wiederum die Frage von Interesse ist, ob es in einem Zeichensystem mit verschiedenen Codierungsprinzipien ein dominantes Prinzip gibt und wie gegebenenfalls untergeordnete Prinzipien sich dem einordnen.

Die zweite Überlegung betrifft das, was man den Zugriffsbereich der Codierungsformen nennen könnte. Generell gilt, daß die Mannigfaltigkeit der

durch ein Zeichensystem unterscheidbaren Bedeutungen nicht größer sein kann als die Mannigfaltigkeit der unterscheidbaren Formen der Zeichen. Das heißt, daß die Struktur des Formbereichs festlegt, was im Bedeutungsbereich unterschieden werden kann. Bei strukturell motivierten Zeichen müssen darüber hinaus die Dimensionen des Bedeutungsbereichs denen des Formbereichs zugeordnet werden können. (Ein Gemälde kann aus eben diesem Grund keinen Prozeß in der Zeit darstellen.) Diese Bedingung gilt für arbiträre Zeichen nicht. Sie sind darum in gewissem Sinn weniger eingegrenzt, d. h., jeder beliebige, diskret strukturierte Bereich kann zum Feld der Bedeutungen eines arbiträren Zeichensystems werden. Mathematische Symbole sind dafür der extremste Beleg: Sie können Gebilde praktisch beliebiger Komplexität wie imaginäre Zahlen oder überabzählbare Mengen erfassen. Die Tatsache, daß alles überhaupt Denkbare (in einem präzisierungsbedürftigen Sinn) in der natürlichen Sprache ausgedrückt werden kann, hat hier ihre Wurzel. Der Preis liegt in der Diskretheit der Bedeutungen: Arbiträre Zeichen können keine kontinuierlichen Übergänge codieren, ohne ihnen zunächst ein Raster diskontinuierlicher Distinktionen aufzuprägen. Sie sind deshalb in einem ganz speziellen Sinn notwendigerweise abstrakt.

Die dritte Überlegung gilt der Beziehung zwischen Codierungsform und Lernabhängigkeit. Als erste Näherung ließe sich sagen, daß die Bedeutung konventioneller Zeichen gelernt werden muß, die Bedeutung motivierter Zeichen dagegen nicht. Die Überlegungen zur Kombination verschiedener Codierungsarten machen klar, daß diese Formulierung differenziert werden muß. Bei strukturell motivierter Codierung müssen zwar nicht die Bedeutungen der einzelnen Zeichen gesondert gelernt werden. Wohl aber kann lernabhängig sein, auf welche Dimensionen sich die Korrespondenz zwischen Form und Bedeutung bezieht, da diese Korrespondenz ja konventionellen Charakter haben kann: Um eine Fieberkurve lesen zu können, muß man wissen, welche Dimensionen die Zeit, welche die Temperatur repräsentiert und nach welchem Schlüssel. Die Verfügung über das Motivierungsprinzip *P* kann also Lernanteile voraussetzen, sie garantiert aber die Beherrschung einer ganzen Klasse von Zeichen, deren Bedeutung dann nicht mehr gelernt werden muß.

Auch für die andere Seite ist eine Differenzierung notwendig: Wenn für arbiträre Zeichen die Bedeutung Stück für Stück gelernt werden muß, dann können solche Systeme immer nur ein endliches, abgeschlossenes Repertoire umfassen, da ja ohne Lernvorgang die Bedeutung neugebildeter Zeichen nicht bekannt wäre. Das würde wiederum heißen, daß natürliche Sprachen entweder nicht auf konventioneller Codierung beruhen, was mit der offensichtlichen Verschiedenheit der einzelnen Sprachen in Widerspruch stünde, oder aber, daß sie jeweils nur über ein endliches Zeichenrepertoire verfügen, was den oben getroffenen Feststellungen widerspricht. Mit anderen Worten, es muß auch bei arbiträren Zeichensystemen die Möglichkeit geben, die Bedeutung neuer Zeichen auf Grund allgemeiner Verfahren zu erschließen,

die aber nicht zu den bereits besprochenen Codierungsprinzipien gehören. Sie beruhen vielmehr auf konventionellen Regeln, die die Bedeutungszuordnung für komplexe Zeichen bestimmen. Der Übergang von elementaren zu komplexen Zeichen macht den Inhalt der Syntax aus, den wir noch zu erörtern haben. Wir halten hier zunächst nur fest, daß auch bei rein konventioneller Codierung durch das Mittel der regelhaften Kombination Zeichen erzeugt werden können, deren Bedeutung nicht einzeln erlernt werden muß. Im strengen Sinn lernabhängig sind jedoch in jedem Fall die Grundzeichen eines arbiträr codierten Systems.

Was sind nun die Schlußfolgerungen aus all dem für Sprache und Musik?

Zunächst gibt es sowohl in der Sprache wie in der Musik ikonische Zeichen, die motiviert sind durch die Ähnlichkeit zwischen dem von der Zeichenform bedingten Schallereignis und dem, worauf die Bedeutung der Zeichen verweist. Beispiele dafür habe ich erwähnt. Solche lautmalenden Zeichen haben jedoch marginalen Charakter. Sie spielen quantitativ eine sehr geringe Rolle; vor allem aber werden sie den dominierenden Funktionsprinzipien des jeweiligen Systems untergeordnet und kommen nur in diesem Rahmen zum Zuge. Wie aber sind diese dominierenden Prinzipien zu beschreiben?

Die Sätze einer Sprache sind komplexe Zeichen, deren Bedeutung sich aus den Grundzeichen, die sie enthalten, und der Struktur ihrer Verknüpfung ergibt. Diese Grundzeichen – etwas vereinfacht gesprochen, die Wörter – sind rein konventionell.³⁴ Beliebig herausgreifbare Beispiele belegen das: Daß ein Tisch im Deutschen *Tisch*, und nicht wie im Russischen *stol* oder wie im Lateinischen *mensa* heißt, ist durch nichts motiviert als durch Tradition, und damit vollständig lernabhängig. Natürliche Sprachen sind demnach prinzipiell arbiträre Zeichensysteme, für die nur die allerdings bedeutsame Bedingung gilt, daß sie im Rahmen der universellen Vorgaben liegen müssen, die für alle Sprachen die gleichen sind. Durchweg konventionsbedingt sind aber nicht nur die Grundzeichen, sondern auch die aus ihnen gebildeten Komplexe (unbeschadet der Tatsache, daß Wortgruppen und Sätze natürlich nicht einzeln gelernt werden), weil auch die Kombinationsregeln konventionell sind.

Für die Musik gilt nun, daß auch hier traditionsbedingte, also konventionelle Faktoren im Spiel sind – die Verschiedenheit der einzelnen Musikkulturen belegt das –, und daß es, wie in der Sprache, komplexe Zeichen gibt, die durch kombinatorische Verfahren erzeugt werden. Beides hat jedoch einen grundsätzlich anderen Charakter als in der Sprache. Ein Schlüssel-punkt ist dabei, daß es in der Musik keine konventionellen Einheiten gibt, die den Wörtern der Sprache an die Seite zu stellen wären. Man kann zwar Motive, Themen, Melodien lernen und wiedererkennen – eine für das Verstehen komplexer musikalischer Strukturen wichtige Bedingung. Aber man lernt mit einem neuen Thema nicht wie bei einem Wort dessen Bedeutung.³⁵ Diese Feststellung kann nun entweder heißen, daß Motive, Themen, allgemein: musikalische Grundgestalten gar keine Bedeutung haben, also keine

Zeichen sind, oder aber, daß ihre Bedeutung von der Art ist, daß sie mit der Form bereits gegeben ist. Will man Musik als Zeichensystem verstehen, wie ich das bislang getan habe, kann man nur die zweite Möglichkeit akzeptieren. Ich vertrete also die These, daß man mit der Form eines Themas auch seine Bedeutung versteht – vorausgesetzt, man kennt die konventionellen Rahmenbedingungen des Codierungsprinzips, das in einem gegebenen musikalischen Code gilt.

Gegen diese These scheinen verschiedene Einwände möglich. Der einfachste ist der, daß bestimmte Motive oder Themen als militärische Signale, als Erkennungsmelodien oder in anderen ähnlichen Funktionen eine feste, eigens zu erlernende Bedeutung haben. Dies ist unstrittig, liegt aber außerhalb der musikalischen Codierung. Musikalische Mittel werden dabei zu Elementen eines konventionellen Zeichensystems, das selbst nicht zur Musik gehört. Vergleichbar ist die Verwendung von Buchstaben wie *e*, *x*, *y*, *z* als mathematische oder *H*, *N*, *O*, *S* usw. als chemische Symbole, die nicht zu ihrer Funktion als Schriftzeichen für die Codierung gesprochener Sprache gehören. Auf kompliziertere Erscheinungen wie z. B. die Funktion von Leitmotiven komme ich später zurück. Auch sie widersprechen aber der hier vertretenen Auffassung nicht.

Gestützt auf die vorangehenden Erläuterungen, läßt sich diese These so formulieren: Die Musik beruht als Zeichensystem auf dem Prinzip der Analog-Codierung, wobei die Merkmalsdimensionen der Zeichenformen auf bestimmte Weise den Dimensionen des Bedeutungsbereichs zugeordnet sind. Diese Zuordnung ist teils ikonischer, teils indexikalischer Natur und unterliegt gewissen konventionellen Eingrenzungen oder Spezialisierungen.

Das bisher für musikalische Grundgestalten Gesagte gilt nun generell auch für komplexe Strukturen, soweit diese Strukturbildung in den Rahmen des Zeichencharakters gehört. Damit erklärt sich in noch auszuführender Weise zugleich der Unterschied in der Rolle der Syntax in Sprache und Musik. Provisorisch gesagt: In der Musik gilt Analogcodierung sowohl für die elementaren wie für die komplexen Zeichen. Und das ist wiederum der Grund dafür, daß die Grenze zwischen elementaren und komplexen Zeichen fließend ist, oder, wie oben bemerkt, daß ein Ablauf nicht eindeutig in Themen bzw. Motive aufgliederbar sein muß. Denn die Bedeutung komplexer Zeichen kommt auf die gleiche Weise zustande wie die ihrer Bestandteile: durch Analogie. In der Sprache dagegen ist die Identifizierung der Grundzeichen Voraussetzung für die Interpretation ihrer Verknüpfungsfunktion und der davon abhängenden Bedeutung der komplexen Zeichen.

Die aufgestellte These ist natürlich nur sinnvoll und in ihren Konsequenzen überprüfbar, wenn zugleich die Dimensionen der Bedeutung angebar sind, auf die sich die behauptete Analog-Codierung der Musik bezieht. Aus diesem Grund werde ich jetzt die Struktur des Bedeutungsbereichs etwas näher zu skizzieren versuchen. Ich stelle dabei wiederum die Charakteristika von Sprache und Musik einander gegenüber.

11. Logische Form als Bedeutung sprachlicher Zeichen

Kognitive Strukturen und Operationen konstituieren den Bedeutungsbereich der Sprache, emotionale Zustände und Prozesse den der Musik: diese vorläufige Kennzeichnung, die ich oben gegeben habe, ist jetzt soweit zu präzisieren, daß sie auf den strukturellen Aufbau der verschiedenen Zeichensysteme bezogen werden kann. Für die Sprache ist dabei der vor allem von Frege geprägte Begriff der logischen Form zu erläutern. Ohne der Fülle von Problemen, die damit zusammenhängen, auch nur annähernd gerecht zu werden, will ich in vier Schritten ein Bild skizzieren, das es dann erlaubt, für die Musik als analogen Begriff den der gestischen Form einzuführen.

Erstens: Die Gesamtheit der kognitiven Prozesse und Strukturen, durch die wir die Umwelt erfassen und unsere Einwirkung auf sie steuern, bildet ein komplexes System von ineinandergreifenden Mechanismen und Verarbeitungsebenen. Sinneseindrücke werden zu Wahrnehmungsstrukturen organisiert, diese werden Klassifikationen unterworfen, die wiederum mit der Organisation kognitiver und motorischer Operationen zusammenhängen, und dies in ständiger Wechselwirkung mit bereits vorhandenem und als Grundlage für die Erzeugung von neuem Gedächtnisbesitz. Die zusammenwirkenden Teilsysteme, die auf diese Weise den gegenständlichen Bezug zur Umwelt herstellen, haben ihre jeweils spezifischen neurophysiologischen Grundlagen, und sie werden insgesamt durch Motivationen und Bedürfnisse aktiviert. Die Verarbeitungsstufe oder Repräsentationsform, in der die Resultate kognitiver Prozesse sprachlich fixiert werden, bildet nun einen Aspekt dieses komplexen Gesamtsystems, und zwar einen sehr zentralen: die Ebene der begrifflichen Repräsentation. Auf sie beziehen sich logische Zusammenhänge und Operationen, durch die auf rein gedanklichem Weg neue Erkenntnisse gewonnen werden können. Es deutet vieles darauf hin, daß ihr eine spezielle Form des Gedächtnisses entspricht, die unter dem Titel ‚semantisches Gedächtnis‘ ein zentrales Untersuchungsfeld der neueren Psychologie geworden ist. Durch die sprachliche Codierung wird den Strukturen dieser Repräsentationsebene nun eine zweite Repräsentation zugeordnet: die lautliche und grammatische Form sprachlicher Zeichen. Dadurch wird die Sprache zur „unmittelbaren Wirklichkeit des Gedankens“ und die begriffliche Ebene der kognitiven Prozesse zur logischen Form sprachlicher Zeichen.

Zweitens: Die zentrale Kategorie der logischen Form ist die Proposition. Eine Proposition ist die begriffliche Repräsentation eines (realen oder fiktiven) Sachverhalts. Das Allgemeinste, was sich über einen Sachverhalt sagen läßt, ist, daß gewisse Dinge bestimmte Eigenschaften oder Beziehungen aufweisen. Eine Proposition ist demnach im einfachsten Fall eine strukturierte Verbindung der Repräsentationen von Dingen und Eigenschaften oder Beziehungen. Ding-Repräsentationen heißen Terme, Repräsentationen von Eigenschaften und Beziehungen heißen Prädikate. Die drei Kategorien Proposition, Term und Prädikat bestimmen sich gegenseitig und legen ein Ele-

mentarschema fest, auf das sich systematische Differenzierungen und Anreicherungen gründen: Prädikate und Terme können in sich zusammengesetzt sein, Propositionen können in Terme eingebettet werden, so daß Sachverhalte von beliebig komplexer Struktur repräsentiert werden können. Die Kombinatorik, die durch das angedeutete Kategoriensystem konstituiert wird, stützt sich wiederum auf ein System von Grundelementen. Jedes dieser begrifflichen Grundelemente faßt mehr oder weniger komplexe Zustände oder Operationen der perzeptiven und höheren kognitiven Verarbeitungsprozesse zu einer Einheit zusammen.³⁶ Ähnlich wie die Elemente der Lautstruktur sind diese Grundeinheiten in Dimensionen geordnet (die abstrakte Repräsentation von Form, Zweck, Bewertung, Zeit und sozialen Beziehungen sind Beispiele dafür), nur ist die Anzahl der Dimensionen und Merkmale hier wesentlich größer, die Art ihrer Beziehungen zueinander verwickelter und abstrakter. Das Grundraster dieser Strukturbildung ist offenbar organisches fixiert und in bestimmten Ansätzen der Klassifikations- und Verarbeitungsleistungen bis in den vormenschlichen Bereich verfolgbare.

Drittens: Eine Proposition ist demnach in einem sehr abstrakten Sinn ein Abbild eines möglichen Sachverhalts, oder besser: eine komplizierte Anweisung, wie der fragliche Sachverhalt zu identifizieren ist. Im Normalfall – aber nicht immer – ist die Befolgung einer solchen Anweisung ein völlig automatischer, unreflektierter Prozeß, der eine gedankliche Struktur auf die Realität bezieht. Auf einen Sachverhalt wird nun in der Kommunikation in verschiedener Weise Bezug genommen: Er kann als bestehend oder als nur vermutet oder erwünscht dargestellt werden, zu seiner Herstellung kann aufgefordert, sein Bestehen kann erfragt werden. Diesen verschiedenen Bedingungen entsprechend ist eine Proposition jeweils einem bestimmten Typ zuzuordnen, der die Einstellung der Kommunikationspartner zu einer Proposition bestimmt. Die Sätze *Du schläfst nicht*, *Schlaf nicht!* und *Schläfst du nicht?* beziehen sich in diesem Sinn auf den gleichen Sachverhalt, daß nämlich der Angesprochene nicht schläft, aber indem der Sachverhalt einmal konstatiert, einmal gefordert, einmal sein Bestehen erfragt wird. Der Propositionstyp gehört zur logischen Form, weil er die Art bestimmt, in der eine Proposition sich auf einen Sachverhalt bezieht. Die darin beschlossene Möglichkeit, einer Sachverhaltsdarstellung verschiedene Funktionen zu geben, ist für die Sprache wesentlich und wird in entsprechenden Mitteln wie Imperativ und Fragesatz fixiert.

Viertens läßt sich nun die Anatomie der Beziehung zwischen Form und Bedeutung in der Sprache etwas genauer beschreiben. Propositionen sind durch die kognitive Informationsverarbeitung erzeugte abstrakte Abbilder von Sachverhalten. Ihre Struktur ist bedingt einerseits durch das, was sie abbilden, andererseits durch die Form der Verarbeitung des Abzubildenden. Ersteres heißt, daß zwischen den Sachverhalten und den sie repräsentierenden Propositionen Strukturähnlichkeiten bestehen müssen, letzteres, daß diese Ähnlichkeiten abstrakter Natur sind. Denn Propositionen reprä-

sentieren Sachverhalte der verschiedensten Art – räumliche, zeitliche, kausale, soziale, ideelle Zustände oder Vorgänge – auf prinzipiell die gleiche Weise, eben als logische Form. Und die hat mit dem identifizierten Sachverhalt jeweils nur die logische Struktur gemeinsam. Die logische Form ist sozusagen dimensionslos und deshalb auch nur greifbar als Struktur von etwas, in dem sie sich manifestiert. Die uns interessierende Manifestation ist die sprachliche Codierung, die die logische Form zur Bedeutung sprachlicher Zeichen macht, sie also in die Struktur sprachlicher Lautmuster überführt – oder verkleidet. Das geschieht in zwei Schritten, die erstens die Bedeutung der Grundzeichen und zweitens die der komplexen Zeichen determinieren. Die Bedeutung der Grundzeichen einer Sprache ist jeweils der Anteil, den sie zur Konstituierung von vollständigen Propositionen leisten. Die Lautform läßt dabei im allgemeinen nichts von der Struktur dieses Anteils erkennen: *hoch* und *doch* oder *Ader* und *aber* tragen trotz ihrer Lautähnlichkeit auf völlig verschiedene Weise zur logischen Form komplexer Zeichen bei. Umgekehrt codieren *beginnen* und *anfangen* oder sogar *steigen* und das komplexe Zeichen *sich aufwärts bewegen* trotz ihrer lautlichen Verschiedenheit praktisch gleiche Teilstrukturen der logischen Form. Anders der zweite Schritt: Die formal-grammatischen Beziehungen zwischen den Teilen eines komplexen Zeichens entsprechen in geregelter Weise bestimmten logischen Verbindungen zwischen den durch sie gelieferten Anteilen der Bedeutung des komplexen Zeichens: Die verschiedene Verknüpfung der Wörter in *Kleider machen Leute* und *Leute machen Kleider* entspricht den Unterschieden in der logischen Struktur der durch die beiden Sätze ausgedrückten Propositionen. Man könnte demnach im Sinn einer These von Leibniz sagen, daß die Grundzeichen ihre logische Form nicht darstellen, wohl aber die aus ihnen gebildeten Komplexe. Doch gilt für natürliche Sprachen auch bei diesem zweiten Schritt, daß die logische Form eher verhüllt als dargestellt wird: *Kafkas Romane* und *die Romane Kafkas* haben die gleiche logische Form bei unterschiedlicher Verknüpfung der Zeichenformen, *Er kommt aus Halle* und *Er kommt aus Versehen* illustrieren den umgekehrten Fall. Das Verhältnis von logischer und grammatischer Form sprachlicher Zeichen ist deshalb ein wichtiges Thema logischer Sprachanalyse.³⁷ Die „Undurchsichtigkeit“, die dabei aufzuhellen ist, hat zwei Gründe: einmal den generellen Grund, daß eine (dimensionslose) logische Form auf die Struktur von Lautmustern abgebildet werden muß, zum anderen den speziellen Grund, daß das in jeder Sprache auf Grund arbiträrer, konventioneller Codierung geschieht.

12. Gestische Form als Bedeutung musikalischer Zeichen

Während die Charakterisierung der logischen Form als Bedeutung sprachlicher Zeichen sich auf detaillierte Befunde der Psychologie, Logik und Linguistik stützen und deshalb auch leicht weiter präzisiert werden kann, haben

die folgenden Überlegungen einen sehr viel weniger abgesicherten Hintergrund.

Die Struktur, für die ich den Terminus „gestische Form“ vorschlage, verhält sich zur Gesamtheit der emotionalen, affektiven und motivationalen Zustände und Prozesse ähnlich wie die logische Form zu den kognitiven Strukturen und Prozessen. Ich werde deshalb zunächst drei Anmerkungen über Emotionen im allgemeinen machen.

Erstens: Kognitive Prozesse und Emotionen beruhen auf unterschiedlichen organismischen Grundlagen, aber sie bilden eine komplizierte funktionale Einheit: Kognitive Leistungen werden von Motivationen geleitet und von emotionalen Einstellungen begleitet, Emotionen haben kognitiv strukturierte Anlässe, Motive haben begrifflich faßbare Ziele. Im Sinn dieses funktionalen Zusammenhangs beschreibt Sartre (1939) Emotionen phänomenologisch als Modi der Umweltinterpretation, als Arten, in der gegenständliche Sachverhalte geformt oder umgeformt werden im Licht subjektiver Erwartungen oder Bedürfnisse: Die gleiche Landschaft enthüllt sich als unübersichtlich oder interessant oder unheimlich, eine Begegnung als erfreulich oder gleichgültig, je nachdem, welche Intentionen sich auf sie richten. Die emotionale Einstellung macht den Modus aus, in dem Informationen gesucht und verarbeitet werden.³⁸ Mit anderen Worten: Kognitive Prozesse – einschließlich ihrer sprachlichen Artikulation – haben stets einen mehr oder weniger ausgeprägten emotionalen Aspekt, Emotionen ein mehr oder weniger deutliches ‚inhaltliches‘ Gerüst kognitiv vermittelter Umweltbezüge. Dabei kann aber die Verarbeitung des gleichen Sachverhalts durch verschiedene Emotionen grundiert sein, die gleiche Emotion kann sehr verschiedene Sachverhalte begleiten. Kurz: Daß kognitive und emotionale Faktoren einander bedingen, ist ein genereller Befund, aber welche Verbindungen dabei entstehen, ergibt sich aus dem jeweils aktuellen Gesamtzusammenhang, nicht aber aus der logischen Struktur der Gedanken und auch nicht aus dem Charakter der Emotionen. Und die kognitiven und emotionalen Faktoren heben sich im Rahmen dieser Einheit umso deutlicher als eigenständige Komponenten ab, je mehr sie durch jeweils spezifische Zeichensysteme strukturiert und fixiert werden.

Zweitens: Während kognitive Prozesse auf (reale oder fiktive) Sachverhalte der Umwelt gerichtet sind und zu ihrer (gegebenenfalls sprachlichen) Repräsentation führen, organisieren emotionale Faktoren eben diese Gerichtetheit, d. h. die subjektive Einstellung zu den Sachverhalten. In dieser Funktion können sie phylogenetisch weit zurückverfolgt werden zu Grundmustern, die zunächst generelle Verhaltenseinstellungen wie Zu- und Abwendung und speziellere Verhaltensformen wie Werbung, Angriff, Flucht determinieren. Aus diesen Zusammenhängen leiten sich die physiologischen Grundlagen und Manifestationen her, die den Organismus für den jeweiligen Verhaltensmodus disponieren. Dies gilt in zweierlei Hinsicht: einmal für die organismusinternen (vor allem nervalen und hormonellen) Mechanismen,

die den Status, den Aktivitätsmodus und das Aktivitätsniveau regulieren, zum anderen für den nach außen gerichteten Emotionsausdruck, die spezifischen Symptome. Sie bilden artspezifische, biologisch fixierte Muster und Abläufe, die zwar lernabhängig in gewissen Grenzen unterdrückt oder modifiziert werden können, aber zunächst spontan wirksam werden. Durch diese Ausdrucksmuster – vor allem in Gesicht und Stimme – werden Emotionen, die vorerst rein subjektiv erfahren werden, intersubjektiv, also sozial zugänglich. Auf diesem Ausdrucksaspekt beruht eine für Emotionen charakteristische Form der Übermittlung, die als Übertragung oder Ansteckung bezeichnet wird und von der Kommunikation durch intendierten Zeichengebrauch zu unterscheiden ist. Was damit gemeint ist, dürfte intuitiv deutlich sein: Ein wütendes Gesicht, eine gedrückte Stimme zeigt einerseits einen emotionalen Zustand an und kann insofern als indexikalisches Zeichen interpretiert werden. Andererseits kann durch solche Symptome der emotionale Zustand auf einen beteiligten Betrachter oder Hörer übertragen, er kann mit Zorn oder Furcht „infiziert“ werden. Dies ist eine Form der Beeinflussung, in der das zum Zuge kommt, was ich oben mit der „nichtkommunikativen“ Wirkung $W(n)$ eines Signals bezeichnet habe.

Drittens: Der damit in groben Zügen umrissene Bereich der Affekte, Emotionen und Motivationen ist nun strukturell und funktionell analysierbar und systematisch beschreibbar. Die Ausführung eines solchen Unternehmens ist noch in vieler Hinsicht kontrovers,³⁹ doch hat sich zumindest in folgenden drei Prinzipien ein gewisser Consensus gebildet:

- (a) Das Prinzip der Emotionsdifferenzierung: Emotionen sind entsprechend ihrer neurophysiologischen Grundlage auf distinkte Grundfaktoren zurückzuführen. Freude, Betrübniß, Ärger, Furcht sind alltagsprachliche Benennungen solcher distinkter Emotionen. Sie können als Dimensionen eines „emotionalen Raums“ angesehen werden, die jeweils kontinuierliche Ausprägungsstärken aufweisen können.
- (b) Das Prinzip der Wechselwirkung von Emotionen: Auch diskrete, fundamentale Emotionen sind in sich gegliedert, sie weisen, wie wir wissen, eine neurale Grundlage, eine expressive Seite und einen variablen, aber konstitutiven Erfahrungsinhalt auf. Diesen drei Komponenten entsprechend können Emotionen in verschiedener Weise aufeinander einwirken, sich überlagern, ineinander übergehen.
- (c) Das Prinzip der emotionalen Musterbildung: Aus (a) und (b) ergibt sich, daß die emotionalen Grundkomponenten komplexe Kombinationen und Übergänge bilden, die die alltäglichen, oft ambivalenten emotionalen Zustände und Abläufe ergeben. Die konkrete Mannigfaltigkeit emotionaler Erscheinungen gliedert sich damit in emotionale Muster, die aus Grundkomponenten und deren Wechselwirkungen hervorgehen.

Einer der Kernpunkte meiner Überlegungen ist nun die Annahme, daß die musikalische Codierung sich auf den Bereich der Emotionen bezieht, indem sie einen Aspekt fixiert, der als gestische Form zur Bedeutung musikalischer Zeichen wird – so wie die logische Form den sprachlich codierbaren Aspekt kognitiver Prozesse darstellt. Die gestische Form hat damit – wie die logische Form – zwei Bezüge: Sie ist ein ausgezeichneter Aspekt der vielschichtigen Gesamtheit emotionaler Prozesse, und sie bildet die Bedeutung eines Typs von Zeichen, von denen die musikalischen die größte Differenzierung erlauben. Der Proposition als der zentralen Kategorie der logischen Form entspricht hier der Gestus als zentrale Kategorie der gestischen Form.

Den damit umschriebenen Begriff des Gestus hat nicht nur Bertolt Brecht weit über die Anwendung auf Formen der Körperbewegungen hinaus verallgemeinert. Auf dem Hintergrund der vorangehenden Erläuterungen will ich einen Gestus etwas genauer bestimmen als die Struktur eines kohärenten emotionalen Musters. Es entsprechen ihm bestimmte neurophysiologische Grundlagen, charakteristische äußere Symptome sowie variable, aber nicht zufällige Erfahrungsinhalte. Von den äußeren Symptomen sind die Ausdrucksformen des Gesichts die differenziertesten. Für uns am wichtigsten sind jedoch Modalitäten der motorischen Aktivität und der Stimmführung. Die involvierten Grundlagen und Symptome sind jeweils für sich identifizierbar, etwa als Pulsfrequenz, Spannung und Lockerung der Artikulationsmuskulatur und zahlreiche weitere Faktoren, sie sind aber durch ein Emotionsmuster als strukturierte Einheit organisiert. Ein Gestus ist gewissermaßen der emotionale Sinn eines Komplexes physiologischer Zustände oder Prozesse, genauer: der strukturell identifizierbare Aspekt eines solchen Komplexes. Fünf Erläuterungen sollen das damit gewonnene Konzept noch verdeutlichen.

Erstens: Während eine Proposition eine sozusagen ausdehnungslose, logische Struktur aufweist, deren Projektion auf die zeitlich organisierte Form sprachlicher Zeichen in gewissem Sinn äußerlich ist, ist ein Gestus als solcher zeitlicher Natur. Er ist selbst die Struktur eines Zustands oder Verlaufs, er vertritt sie nicht nur. Eine Proposition kann Sachverhalte mit beliebiger zeitlicher Struktur und zeitlicher Einordnung repräsentieren, aber auf indirekte, abstrakte Weise, ein Gestus dagegen ist selbst und direkt ein zeitlich strukturiertes Gebilde.

Aus dieser Tatsache ergeben sich wesentliche Bestimmungen der möglichen internen Struktur eines Gestus, die immer die Dimension der Zeit, genauer der Zeiterfahrung, aufweist. So kann ein Gestus andauernd, stationär – gegebenenfalls mit periodisch sich wiederholenden Substrukturen – oder durch seinen Verlauf zeitlich begrenzt, ja schockhaft kurz sein. Immer aber handelt es sich um echte Zeitintervalle, nicht um ausdehnungslose Momente oder um zeitfreie Permanenz. Und das für die Zeit Gesagte ist weiterhin auszudehnen auf alle anderen Aspekte, die Qualitäten, die psychischen und physischen Charakteristika, die sich in die zeitliche Erstreckung der emotio-

nalen Muster ausdehnen, sie füllen oder ‚einfärben‘. Ein Gestus ist die Struktur eines ängstlichen Verharrens oder eines begeisterten Aufschwungs, eines wütenden Sich-Wiederholens, einer gelassenen Abwendung oder nüchterner Aufmerksamkeit – um das Gemeinte provisorisch durch Beispiele anzudeuten. Kutz: Die gestische Form muß genau die Dimensionen aufweisen, innerhalb deren emotionale Muster strukturiert sein können.

Zweitens: Ein Gestus kann in sich gegliedert sein. So kann etwa ein Grundgestus durch modifizierende Gesten überdeckt oder vorübergehend aufgehoben sein; qualitativ verschiedene Gesten können sich überlagern, ablösen oder unterlaufen: Was zum Beispiel als Aufschwung beginnt, kann schon Trauer in sich haben, oder umgekehrt. Worauf solche intuitiv formulierten Möglichkeiten hinweisen sollen, ist dem nicht unähnlich, was ich für Propositionen gesagt habe: Die Bildung komplexer Strukturen aus ineinander eingebetteten Einheiten. Der wichtige Unterschied ist aber der, daß die „Bestandteile“ eines Gestus, die über- und untergeordneten Teilgesten, sich immer in der Dimension der Zeit arrangieren, während Teilpropositionen sich abstrakt, auf Grund logischer Beziehungen ineinander verschachteln. Teilpropositionen können Raum- oder Zeitbeziehungen zwischen den repräsentierten Sachverhalten festhalten – sie selbst stehen in keinem Zeitverhältnis zueinander, wie ihre sprachliche Codierung deutlich macht: *Alle, die jetzt dafür sind, waren gestern dagegen* ist ein beliebig herausgegriffenes Beispiel. Stark vereinfacht: Propositionen sind logisch, Gesten zeitlich verknüpft. Diese Überlegungen machen Grundrisse einer möglichen Theorie der gestischen Form kenntlich, freilich auch nicht mehr.

Drittens: Das Grundprinzip, das die gestische Form zur Bedeutung musikalischer Zeichen macht, ist offensichtlich das der Analogcodierung. Die Dimensionen der Lautmuster werden dabei denen der gestischen Form zugeordnet mit teils diskreter, teils kontinuierlicher Codierung der Ausprägungsgrade. Die einfachste Korrespondenz ergibt sich in der Zeitdimension, für andere Dimensionen ist die Zuordnung weniger direkt. Intensität erscheint als Lautstärke nur relativ zum Gesamtcharakter eines Gestus, die Tonhöhenbewegung repräsentiert emotionale Bewegungen durch so etwas wie einen abstrakten ‚Bewegungsraum‘, der sich durch synästhetische Bezüge zum physischen Bewegungsfeld mit seinen Bestimmungen durch Schwere, Weite, Höhe greifbar machen läßt. Die Zuordnungsformen über solche intuitiven Umschreibungen hinaus systematisch zu analysieren, ist ein kompliziertes, aber sehr wohl begründbares Programm, das auf zahlreiche unstrittige Einzelbeobachtungen zurückgreifen kann.

Viertens: Die Motivationsbasis für die skizzierte Analogcodierung ist vornehmlich indexikalischer Natur, schließt aber ikonische Faktoren ein. Musikalische Zeitmuster ordnen sich automatisch den Zeitmustern organischer Prozesse zu, deren Charakteristik wiederum kausal verbunden ist mit Komponenten emotionaler Muster: Erregung, Freude, Wut beschleunigen Puls und Atemfrequenz, Niedergeschlagenheit und Trauer verlangsamen sie. Die

Bewegtheit oder Gleichförmigkeit von Tonhöhenabfolgen gibt den unterschiedlichen Spielraum wieder, der die Motorik etwa bei Bedrücktheit und Furcht oder bei Freude und Zorn umschreibt. Der Zusammenschluß kausaler und ikonischer Motivation wird überdies begleitet von der Übertragungs- oder Ansteckungswirkung, von der bereits die Rede war: Rasche Bewegung im musikalischen Lautmuster ist nicht nur Symptom etwa von Erregtheit und ikonische Repräsentation ihrer physischen Effekte, sie löst auch unwillkürlich Erregtheit aus. Diese Wirkung spielt zumindest für den jeweiligen Grundgestus fast stets eine konstitutive Rolle, nicht nur in kultischer oder auf andere Weise „beschwörender“ Musik, sondern auch in differenziert organisierter Kunstmusik: Überschwang und Ausgelassenheit etwa teilen sich im vierten Satz der *Siebenten Sinfonie* von Beethoven in diesem Sinn ganz spontan mit, ebenso wie das unheimliche Stocken in einigen Passagen des letzten Satzes in Mahlers *Lied von der Erde* – um beliebig zwei Beispiele herauszugreifen. Allerdings ist auch diese elementare Motivierungsform bestimmten konventionellen Rahmensetzungen und damit gewissen lernabhängigen Eingrenzungen unterworfen. Das bezieht sich sowohl auf die geschichtlich bedingten Mittel der Lautmusterbildung als auch auf die kulturabhängige Entfaltung und Eingrenzung emotionaler Muster und ihrer Symptome, auf die „Erziehung der Gefühle“. Je differenzierter die gestischen Strukturen werden, die musikalisch zu codieren sind, desto stärker tritt dieser kultur- und lernabhängige Anteil in den Vordergrund.

Fünftens: Ein musikalisch codierter Gestus steht in konstitutiver Weise zu einem variablen, aber nicht beliebigen Erfahrungsinhalt kognitiver Natur in Bezug. Das folgt unmittelbar aus dem oben erörterten Charakter emotionaler Prozesse und Musterbildungen. Diese Tatsache spielt eine fundamentale Rolle für die Beziehung zwischen der Bedeutung musikalischer Zeichen und außermusikalischen Faktoren. Diese Beziehung ist durch zwei scheinbar gegensätzliche Feststellungen umschreibbar: Mit welchen Gedanken oder Vorstellungen Musik sich verbindet, ist praktisch nicht festlegbar, und: die Verbindung musikalischer Gestalten mit außermusikalischen Vorstellungen und Gedanken ist keineswegs beliebig. Ersteres wird z. B. belegt durch die Tatsache, daß ein Text mehrere (adäquate) Vertonungen zuläßt, aber auch die gleiche musikalische Gestalt mehreren Texten zugeordnet werden kann. Die zweite Feststellung wird belegt durch die Beobachtung, daß nicht jede Musik zu einem gegebenen Text „paßt“, daß man mehr und weniger adäquate Vertonungen unterscheiden kann. Etwas vereinfacht, aber nicht irreführend auf den Begriff gebracht, besagen diese simplen Feststellungen: Einer musikalischen Gestalt entspricht nicht ein einzelnes kognitives Korrelat, sondern eine Klasse (im einzelnen schwer zu bestimmender) Korrelate kognitiver Natur. Die Vermittlungsstufen dieser Zuordnung habe ich aufzuschlüsseln versucht: Die Form eines musikalischen Zeichens codiert einen Gestus, der die Struktur eines emotionalen Musters ist, zu dem wiederum ein variabler, aber nicht beliebiger Erfahrungsinhalt gehört. Dieser

Zusammenhang kann sich auf unterschiedliche Stufen kognitiver Repräsentationen beziehen, die ihm jeweils eigene Charakteristika aufprägen:

(a) Der kognitive Bezug ist nicht nur begrifflich artikuliert, sondern auch sprachlich fixiert. Die Beziehung zwischen logischer und gestischer Form ist dann die zwischen Text und Musik, die beide durch das gleiche Signalereignis realisiert werden.⁴⁰ Bei der Konkretisierung dieser Beziehung kann die logische Form des sprachlichen oder die gestische Form des musikalischen Zeichens dominieren, beide können einander ergänzen oder in Spannung zueinander stehen. Die Musik fixiert einen Gestus für den Text.

(b) Der kognitive Bezug ist nicht sprachlich fixiert, aber in logisch strukturierter Form identifizierbar. Musikalische Signalereignisse laufen dann entweder parallel zu anderen Mitteilungsprozessen oder bringen vorab bekannte kognitive Strukturen ins Spiel. Theater- und Filmmusik sind Beispiele der ersten, viele Formen direkter oder indirekter Programmmusik sind Beispiele der zweiten Möglichkeit.

(c) Der kognitive Bezug ist nicht auf der Ebene der logischen Form repräsentiert. Er hat dann den Charakter allgemeiner Einstellungsanlässe, die mit dem jeweiligen Gestus verträglich sind. Sie können, müssen aber nicht auf anschauliche Vorstellungsmerkmale beziehbar sein. Das „Frühlingshafte“ in der *Ersten Sinfonie* Schumanns, das „Glitzernde“ in Vivaldis Violinkonzerten deuten solche Vorstellungsbezüge an. Zwischen diesen Möglichkeiten gibt es Übergangsformen, die von Fall zu Fall zu bestimmen sind. Die Aufzählung sollte nur deutlich machen, daß musikalische Zeichen stets eine inhärente Affinität zu begrifflichen oder Vorstellungsstrukturen haben, die sich auf die eine oder andere Weise konkretisiert, daß sie aber nicht selbst diese Strukturen codieren. Welche der möglichen Konkretisierungen tatsächlich zum Zuge kommen, ist in unterschiedlichem Maße mit lernabhängigen Konventionen verbunden, die einzelne Werke, aber auch ganze Systeme musikalischer Mittel betreffen können.

13. Sagen und Zeigen

Ehe ich die Charakterisierung der Zeichenstrukturen von Musik und Sprache mit der Betrachtung der Syntax abschließe, will ich die bisher verfolgten Fäden zusammenfassen. Die Gegenüberstellung von Sagen und Zeigen bietet sich dafür als knappe, aber deutliche Kennzeichnung an. Mit „Sagen“ und „Zeigen“ lassen sich zwei Arten des Mitteilens unterscheiden, die im Kern durch die Codierungsart der verwendeten Zeichen gekennzeichnet sind.

Ich beginne mit der nur scheinbar trivialen Feststellung, daß sprachliche Zeichen dazu da sind, daß mit ihnen etwas gesagt werden kann. Auch im einfachsten Fall lassen sich an einem Kommunikationsakt, in dem mit sprachlichen Mitteln etwas gesagt wird, drei Stufen unterscheiden: Erstens

werden Dinge, Eigenschaften, Relationen, Zustände und Vorgänge benannt. Zweitens wird auf Grund der Benennungen ein Sachverhalt identifiziert, der aus einer bestimmten Konstellation der benannten Gegebenheiten besteht. Drittens wird gesagt, was es mit dem so identifizierten Sachverhalt auf sich hat, daß er z. B. besteht oder herbeigeführt werden soll oder möglich oder wünschenswert wäre. Etwas sagen heißt demnach, ein sprachliches Zeichen äußern, dessen Bedeutung eine Proposition ist und das durch seinen Propositionstyp die Geltungsart des Gesagten bestimmt.⁴¹ Benennen der Gegenstände und Eigenschaften, Identifizieren oder Beschreiben des Sachverhalts und Sagen, was es mit dem Sachverhalt auf sich hat, sind dabei keine getrennten Vorgänge, sondern werden in einem Zug aufeinander aufbauend realisiert. Wesentlich ist dabei, daß das Benennen und das Beschreiben auf Grund konventioneller Regeln geschieht, die erworben werden und (im Prinzip) auch anders sein könnten. Etwas sagen setzt konventionelle Möglichkeiten des Beschreibens voraus, die wiederum konventionelles Benennen verlangen.

Zeigen ist demgegenüber eine vergleichsweise elementare Form des Mitteilens: Sie setzt nicht notwendig ein festes erworbenes Zeichensystem voraus, und sie geschieht nicht durch mehrere aufeinander aufbauende Stufen, sondern gewissermaßen in einem Schritt. Dabei sind zwei Möglichkeiten zu unterscheiden. Man kann erstens jemandem auf verschiedene Weise z. B. ein Buch, eine Vase, eine Landschaft oder ein Gebäude zeigen. Und man kann zweitens zeigen, wie ein Spiel zu spielen ist, wie jemand sich hinsetzt, oder wie sich ein aufgeregter Mensch verhält. Im ersten Fall verweist die Zeigehandlung auf Dinge, Vorgänge, Zustände, im zweiten Fall stellt sie Vorgänge oder Sachverhalte zeigend dar. Auf diese zweite Form des Zeigens, als deren Prototyp die Pantomime gelten kann, kommt es mir im weiteren an. Sie vergegenwärtigt etwas, indem die Zeigehandlung Strukturgleichsamkeiten mit dem gezeigten Vorgang oder Sachverhalt aufweist, und nur sofern diese Gemeinsamkeiten bestehen, kann sie etwas mitteilen. Diese Gemeinsamkeiten sind nicht arbiträr, bedürfen also auch keiner Konvention. Wo eine gemeinsame Sprache fehlt, hilft deshalb – in Grenzen – Vormachen und Zeigen weiter. Mit anderen Worten, Zeigen beruht auf strukturell motivierter Codierung der Bedeutung im Schema der Zeigehandlung.

Löst man das Zeigen von der Bedingung, daß es auf optischem Wege zu geschehen hat, und interpretiert es nicht im engeren Sinn als „sehen machen“, sondern allgemeiner als „wahrnehmen machen“, dann kann man auf natürliche Weise so etwas wie akustische Zeigehandlungen einführen. Die Musik läßt sich damit als ein Zeichensystem bestimmen, mit dessen Hilfe auf akustischem Wege etwas gezeigt werden kann. Und mit dem Blick auf das im vorigen Abschnitt Erläuterte können wir genauer sagen: Musikalische Zeichen zeigen die in ihnen codierte gestische Form, sie machen emotionale Muster wahrnehmbar, indem sie deren gestische Struktur zeigen.

Die Zusammenfassung der bisherigen Erörterungen kann nun lauten: Sprachliche Äußerungen (insbesondere Sätze) drücken Propositionen aus und sagen damit etwas über Sachverhalte, musikalische Äußerungen drücken Gesten aus und zeigen damit emotionale Muster. Und noch kürzer: Die Sprache sagt, die Musik zeigt, was sie mitteilt.

An diese Formel, die nur zusammen mit den vorangehenden Erläuterungen sinnvoll ist, will ich eine Reihe verdeutlichender Kommentare anschließen.

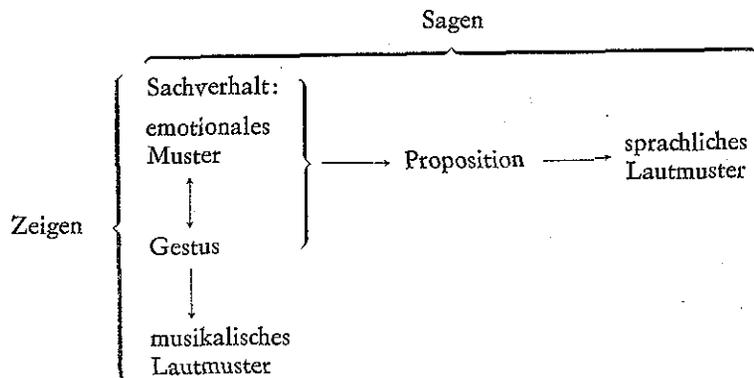
Erstens: Konventionsbedingtes Sagen und strukturell motiviertes Zeigen sind einander bedingende Pole eines Kontinuums. Zwischen diesen Polen liegen Übergangsformen, die durch mehr oder weniger große Anteile von Konventionalität gekennzeichnet sind. Auf lernabhängige, historisch sich wandelnde Faktoren in musikalischen Systemen habe ich bereits hingewiesen. Sie betreffen zunächst die Kenntnis des Grundinventars, das für die Lautmuster genutzt wird. Ohne Kenntnis dieses Rahmens wird man beispielsweise nicht erfassen, daß der erste Satz der letzten g-Moll-Sinfonie von Mozart keineswegs vergnüglich bewegt ist, wie neuerdings eine verpoppte Version suggeriert, sondern einen ganz anderen Grundgestus zeigt. Sie erfassen weitgehend aber auch die Zuordnung gestischer Momente oder Grundkomponenten zu bestimmten Dimensionen oder sogar festen Konfigurationen musikalischer Lautmuster. Wer in die zugehörigen Konventionen nicht eingelebt ist, wird etwa die Art, in der ironisch gebrochene Liebe im Quintett Nr. 6 des ersten Aktes von *Così fan tutte* musikalisch gezeigt wird, nicht wahrnehmen. Oder allgemeiner: Um z. B. japanische Gagaku-Musik durch die formalen akustischen Muster hindurch als Darstellung von sinnkonstituierenden Gesten aufnehmen zu können, ist vorangehende Erfahrung mit diesen Gesten und der Art ihres Zeigens notwendig.

Zweitens: Ein noch tiefer liegender Zusammenhang als die partielle Konventionalisierung des Zeigens verbirgt sich in der Tatsache, daß umgekehrt in jedem konventionellen Zeichensystem eine zeigende Basis enthalten ist, die jenseits aller Konventionalität liegt und das Benennen und Beschreiben erst ermöglicht. Man kann sich das folgendermaßen klarmachen: Das Benennen von Dingen, Eigenschaften usw. verlangt die Identifizierung bzw. Klassifizierung der zu benennenden Gegebenheiten. Sie wird bestimmt durch die Bedeutung der benennenden Zeichen. Die Mannigfaltigkeit der Differenzen zwischen den Bedeutungen muß demnach der Mannigfaltigkeit der Gegenstände, Zustände usw. entsprechen. Weiterhin müssen die Beziehungen zwischen den Bedeutungen der Zeichen den Beziehungen zwischen den Dingen, Eigenschaften usw. entsprechen, soweit sie beim klassifizierenden und identifizierenden Benennen berücksichtigt werden. Um ein einfaches Beispiel zu geben: Die Bedeutung des Wortes *Hund* muß auf bestimmte Weise die Bedeutung des Wortes *Tier* einschließen und die des Wortes *Haus* ausschließen, weil die Klasse der Hunde in der der Tiere enthalten ist und sich mit der Klasse der Häuser nicht überschneidet. Dieses logische Netz, das die Bedeutungen der Zeichen bilden, muß auf abstrakte, aber dennoch struktu-

rell motivierte Weise den Unterscheidungen und Zusammenhängen im Bereich der benannten Gegenstände entsprechen. Ebenso muß die Verknüpfung der Terme und Prädikate in einer Proposition auf abstrakte, aber strukturell motivierte Weise den Zusammenhängen in dem zu beschreibenden Sachverhalt entsprechen. In diesem Sinn zeigt die logische Form eines sprachlichen Zeichens die Struktur der Dinge, Beziehungen und Sachverhalte, die mit ihm benannt oder beschrieben werden. Die logische Form ist das Netz von Zusammenhängen und Unterscheidungen, das die Bedeutung sprachlicher Zeichen mit den benannten und beschriebenen Dingen, Eigenschaften und Sachverhalten gemeinsam hat. Aus dem damit gesetzten Rahmen kann keine Konventionalisierung der Grundzeichen und der Verknüpfungsregeln eines arbiträren Zeichensystems herauspringen. Es ist zwar eine Sache der Konvention, ob es in einer Sprache z. B. ein Zeichen mit der Bedeutung von *schchenken* gibt und welche Lautform es hat. Aber wenn diese Bedeutung in einer Sprache vorkommt, dann muß sie ein dreistelliges Prädikat sein, d. h. eine Beziehung zwischen drei Dingen zeigen. Andersherum gesagt: Wie immer sich ein konventionelles Zeichensystem konstituiert oder verändert, in der logischen Form seiner Bedeutungen ist eine Struktur enthalten, in der sich die Struktur der Sachverhalte zeigt, über die etwas gesagt wird. Nur durch diese Strukturgleichheit, die die Bedeutungen der Zeichen mit den Dingen und Sachverhalten verbindet, kann mit konventionellen Zeichen etwas gesagt werden.⁴² Hinter der Möglichkeit, in konventionsgeregelter Form etwas zu sagen, steht die Bedingung, auf abstrakte Weise die Struktur des Gesagten zu zeigen. (In anderer Ausdrucksweise heißt das: Die Sprache ist ein konventionelles Mittel, Gedanken zu strukturieren und auszudrücken, die ihrerseits die Realität abbilden, also zeigen.) In dem, was man sagt, ist immer auch etwas gezeigt, das selbst nicht gesagt werden kann.

Drittens: Daß das, was sich zeigen läßt, nicht gesagt werden kann, ist eine Tatsache, die zugleich die besonderen, sich der Sprache entziehenden Möglichkeiten der Musik charakterisiert. Um das zu verdeutlichen, machen wir uns zunächst klar, daß ein Gestus im oben beschriebenen Sinn die Struktur eines besonderen Sachverhalts ist, nämlich eines emotionalen Musters mit all seinen Komponenten und Bezügen, ein interner, psychischer Sachverhalt also. Ein solcher Sachverhalt kann gedanklich erfaßt und sprachlich ausgedrückt werden, und zwar auf zwei verschiedene Arten: Zum einen als unmittelbare Erfahrung oder Vorstellung durch Sätze wie *Ich bin ärgerlich*, *Er regt sich auf*, *Sie schwankt zwischen Neugier und Furcht*, aber auch in Aufforderungen oder Fragen wie *Reg dich ab!* und *Haben sie Angst?* Zum anderen als analysierende Vergegenständlichung, in der die internen Zustände in einen theoretischen Rahmen eingeordnet werden. Dies ist der Weg, auf dem ich im vorigen Abschnitt bestimmte Komponenten emotionaler Zustände zu charakterisieren versucht habe.⁴³ In beiden Fällen wird über gestische Strukturen und die mit ihnen verbundenen emotionalen Muster etwas gesagt. Dem steht die Möglichkeit gegenüber, von der die Musik Gebrauch

macht, indem sie Gesten als Sachverhalte zeigt. Was auch an konventioneller Formung in die dabei benutzte Analogcodierung eingehen mag: sie zeigt etwas, was nicht gesagt werden kann. Das Verhältnis entspricht genau dem zwischen der sprachlichen Beschreibung eines Gesichts und seiner Abbildung durch eine Photographie oder ein Gemälde (das ebenfalls konventionalisierte Komponenten enthalten kann): Auch die genaueste Beschreibung kann nicht das sagen, was das Bild zeigt. (Die Ersetzung einer Personenbeschreibung durch ein Phantombild macht das handfest deutlich.) Die Sprache kann alles beschreiben, weil sie es auf seine logische Struktur reduziert, aber eben auch nur, so weit diese Reduktion reicht; die Musik kann nur das zeigen, was den Dimensionen ihrer Zeichenform entspricht, dies aber ohne die Reduktion auf abstrakte Distinktionen und Beziehungen. Mit aller Vorsicht kann man die beiden Möglichkeiten, ein emotionales Muster und seine gestische Struktur durch Lautmuster auszudrücken, in folgendem Schema zusammenfassen:⁴⁴



Viertens: Was gesagt wird, wird gedanklich, d. h. auf Grund seiner logischen Struktur vergegenwärtigt. Was gezeigt wird, wird dagegen sinnlich vergegenwärtigt, insofern die codierenden Eigenschaften des Signals denen des gezeigten Sachverhalts entsprechen. In diesem Sinn kann man sagen: Gezeigte Bewegung, Ruhe, Ängstlichkeit oder Freude *ist* zugleich Bewegung, Ruhe, Ängstlichkeit, Freude. Einfacher und genauer: Ein Lautmuster, das Aufregung zeigen soll, muß aufgeregt sein. Dagegen enthält der Satz *Er ist aufgeregt* nicht mehr und nicht weniger Aufregung als der Satz *Er ist nicht aufgeregt*, obgleich er das Gegenteil sagt. Daß das Gezeigte und nicht das Gesagte sinnlich vergegenwärtigt wird, hat allenthalben strukturelle Konsequenzen. Aufregung oder Gelassenheit kann beliebig lange gezeigt, aber nicht beliebig lange gesagt werden. Divergierende Momente innerhalb eines emotionalen Musters können bzw. müssen simultan gezeigt werden, während bei ihrer Beschreibung ihre zeitliche Beziehung in eine logische umgesetzt

und auf dieser Grundlage codiert wird, etwa: *Er war gelassen, aber die Aufregung klingt noch nach*. Vor allem aber liegt im sinnlichen Wahrnehmbarmachen der Grund für die oben erwähnte Ansteckungswirkung musikalischer Mitteilung: Das Abklingen oder Wachsen von Aufregung, der Gestus schockhafter Angst oder gefasster Trauer überträgt sich, weil seine Symptome gezeigt werden. Dies ist die Grundlage für die in Abschnitt vier formulierte These, daß in der Musik, anders als in der Sprache, die außerkommunikative Wirkung des Signals eine wesentliche (wenn auch keineswegs die einzige) Komponente im Zustandekommen der Gesamtwirkung ist.

Fünftens: Mit diesem Punkt ist auf ganz fundamentale Weise die Tatsache verbunden, daß Musik faktisch immer ästhetische Funktion hat, sprachliche Kommunikation dagegen nur im Spezialfall. Ich will den Grund kurz den „Als-ob-Charakter“ ästhetischer Mitteilung nennen. Der in den verschiedensten Kunsttheorien auf mannigfaltige Weise reflektierte Tatbestand ist sehr einfach: Wenn jemand bei einem Abendspaziergang sagt *Der Mond ist aufgegangen*, so trifft er eine Feststellung, die zutreffen oder ein Irrtum sein kann, jedenfalls auf einen faktischen Sachverhalt zielt. Wer bei der Lektüre oder Rezitation des *Abendlieds* von Matthias Claudius ans Fenster ginge, um eben diese Aussage zu verifizieren, hätte den Sinn künstlerischer Kommunikation verfehlt. Aus dem gleichen Grund wird niemand einem zum Schlag ausholenden Schauspieler auf der Bühne in den Arm fallen, wie er es im entsprechenden Fall auf der Straße täte. Kurz: Kunst beruht wesentlich auf dem Scheincharakter dessen, was sie sagt oder zeigt.⁴⁵ Wesentlich ist nun, daß sprachliche Äußerungen sowohl im faktischen wie im Als-ob-Modus gemacht werden können. Das Beispiel *Der Mond ist aufgegangen* belegt das hinreichend. Musikalische Äußerungen dagegen können ausschließlich im Als-ob-Modus etwas zeigen. Wenn ich gesagt habe, gezeigte Freude *ist* Freude, dann gilt das nur mit dieser zusätzlichen Qualifikation. Gezeigte Freude *ist* Freude in stellvertretender Form, sie ist *wie* Freude. Der Grund ist jetzt deutlich: Weil Musik Gesten in kommunikativ kontrollierter Absicht zeigt, haben die in ihr codierten emotionalen Muster keinen primären, spontanen Charakter, sondern werden stellvertretend gezeigt, aber mit den echten Symptomen und Ansteckungseffekten, soweit die musikalische Codierung genau auf ihnen beruht.⁴⁶ Andersherum: Wer wirklich wütend ist oder Angst hat, äußert das nicht musikalisch. (Diese Einsicht liegt dem naiven Mißverständnis zugrunde, das Oper als Schauspiel mit störender Musik nimmt.) Mit anderen Worten: Mit dem Unterschied zwischen Sagen und Zeigen und seiner strukturellen Verankerung in den Mitteln der Sprache und der Musik ist ursächlich verbunden, daß Sprache im faktischen und im Als-ob-Modus gebraucht werden kann, Musik nur im letzteren.⁴⁷

Sechstens: Die strukturelle Grundlage für den beschriebenen Unterschied ist durch die am Ende von Abschnitt fünf formulierte These zusammenzufassen: Bedeutung und Form sind in der Sprache (im Rahmen universeller Bedingungen) rein konventionell aufeinander bezogen und also entspre-

chend strukturverschieden, in der Musik sind Form und Bedeutung struktur-analog. Geschichtlicher Konvention unterliegt in der Musik nur die Auswahl der Bereiche von Gesten und des Rahmens ihrer Codierungsmittel. Die Bedeutungsstruktur natürlicher Sprachen, der Aufbau der logischen Form, kann deshalb in gewissem Sinn als eigenständiger Gegenstandsbereich unabhängig von der historisch bedingten Form ihrer sprachlichen Codierung analysiert werden. (Dies ist genau die Zielstellung, die die Logik seit Aristoteles verfolgt.) Die Beschreibung der Semantik einer konkreten Sprache besteht dann in der Angabe der Beziehung sprachlicher Zeichenformen zu diesem universellen Bereich möglicher logischer Strukturen. Eine semantische Theorie der Musik im allgemeinen und einzelner musikalischer Systeme im besonderen hat einen deutlich unterschiedlichen Charakter. Nicht nur, daß sie es nicht mit logischen Formen, sondern mit geistlichen Strukturen zu tun hat. Sie hat auch nicht die Zuordnung zweier verschieden strukturierter Bereiche zu charakterisieren, sondern die gleiche Struktur unter zwei Aspekten. Zugespitzt: Eine semantische Analyse in dem Sinn, in dem sie als Teil einer Theorie der natürlichen Sprache entwickelt wird (vgl. dazu etwa Katz 1972), wird in einer Theorie der Musik nicht enthalten sein. Das ist der theoretische Ausdruck dafür, daß man zwar einen chinesischen Satz ins Englische oder Deutsche übersetzen kann, aber chinesische Musik so wenig ins Europäische übersetzt werden kann wie ein chinesisches Bild.

Siebtens: Die Logik analysiert die Zusammenhänge und die auf ihnen beruhenden Operationen, welche Propositionen und andere Einheiten der logischen Form generell miteinander verbinden. Sie macht mithin die Bezüge deutlich, die sprachliche Ausdrücke auf Grund ihrer Bedeutung zueinander und zu den Sachverhalten haben, über die sie etwas sagen. Sucht man in der Musik nach einem Analogon zu einem solchen System von Bezügen und Operationen, so wäre dies eine Charakterisierung der Zusammenhänge zwischen musikalischen Lautstrukturen auf Grund der durch sie gezeigten geistlichen Struktur. Wir haben uns soeben deutlich gemacht, daß in der Musik Form und Bedeutung der Zeichen strukturell isomorph sind. Das heißt aber, daß die Strukturzusammenhänge in der Form musikalischer Zeichen die gleichen sind wie die in ihrer Bedeutung. Die „Logik der Musik“ ist damit nicht nur eine ohne Begriff und Urteil, sondern sie ist zugleich die Logik der musikalischen Formen selbst. Die Konsequenzen dieser Überlegung sind keineswegs trivial und können an dieser Stelle nicht einmal ernsthaft angedeutet werden.

14. Syntax

Ich habe bereits mehrfach auf die Unterscheidung zwischen Grundzeichen und zusammengesetzten oder komplexen Zeichen Bezug genommen. Die Rolle, die die Bildung komplexer Zeichen für die Zuordnung von Zeichenformen zu ihrer Bedeutung und damit für die unterschiedlichen Codierungs-

arten von Sprache und Musik spielt, ist nun genauer ins Auge zu fassen. Das führt uns zur Syntax als der letzten Strukturebene, die es im Rahmen des Vergleichs sprachlicher und musikalischer Zeichensysteme zu betrachten gilt. Ich werde wiederum zunächst die generellen Züge dieser Strukturbildung umschreiben und dann die Besonderheiten von Sprache und Musik kennzeichnen.

So wie Zeitmuster und Segmentstrukturen ihre psycho-physischen Grundlagen in der Regulation organismischer Zeitstrukturen und in den Bedingungen der Wahrnehmung und Erzeugung akustischer Signale haben und wie die Bedeutungsstrukturen auf den Systemen der kognitiven bzw. emotionalen Verarbeitung der Umwelt beruhen, so ist die generelle Grundlage der syntaktischen Struktur zunächst die organismische Fähigkeit, elementare Einheiten zu strukturierten Komplexen zusammenzufassen bzw. komplexe Gebilde in Komponenten zu gliedern. In diesem sehr weitgreifenden Sinn gehört die Fähigkeit zur Kombinatorik freilich zu fast allen organismischen Leistungen, und wir haben sie deshalb auch auf allen bisher erörterten Strukturebenen bereits angetroffen als Verknüpfungsprinzip der jeweiligen (zeitlichen, lautlichen, begrifflichen, emotionalen) Grundeinheiten. Faßt man Syntax ohne weitere Einschränkung als Verknüpfungsform gegebener Elemente auf, dann hat jede der bisher diskutierten Strukturebenen ihre eigene Syntax (nämlich die in Abschnitt 8 allgemein benannten Verknüpfungsoperationen über den jeweiligen Grundelementen). Eine solche Redeweise ist nicht unüblich und unter bestimmten Gesichtspunkten durchaus sinnvoll. In einem engeren Sinn, auf den es jetzt ankommt, meint Syntax jedoch nur solche Verknüpfungsoperationen, die den Bedeutungszusammenhang in den resultierenden komplexen Einheiten determinieren. Um diesen Unterschied zu verdeutlichen, greife ich noch einmal auf den Morse-Code zurück.

Alle Morsezeichen sind Folgen von Punkten und Strichen (Kürzen und Längen). Im weiteren Sinn beruht der Morse-Code demnach auf einer Syntax, die Längen und Kürzen aneinanderreihet. Dies geschieht jedoch in zwei grundsätzlich verschiedenen Schritten. Aus der Verknüpfung z. B. einer Kürze mit einer Länge läßt sich nicht ableiten, daß die resultierende Folge „kurz lang“ den Buchstaben *a* bezeichnet. Diese Bedeutung muß vielmehr aus der Code-Liste entnommen werden. Die Verknüpfung allein determiniert die Bedeutung der Folge nicht. Dagegen erzeugt die ebenfalls lineare Verbindung der Folge „kurz lang“ (mit der Bedeutung *a*) mit dem Element „lang“ (das die Bedeutung *t* hat) ein komplexes Zeichen mit der Bedeutung *at*. In diesem Fall organisiert die Verknüpfung die Bedeutung des neuen, komplexen Zeichens, nämlich als Aneinanderreihung der Bedeutungen der Einzelzeichen. Die Verknüpfung der Buchstabenzeichen bildet mithin die eigentliche Syntax des Morse-Codes, während die Verknüpfung der Elemente innerhalb der Buchstabenzeichen nicht zur Syntax im engeren Sinn gehört. (Die beiden Verknüpfungsarten werden auch formal gekennzeichnet durch verschieden große Abstände zwischen den Grundelementen.)

Im gleichen Sinn gehört in der Sprache die Verknüpfung der Merkmale bzw. Segmente innerhalb der Wortformen nicht zur Syntax, sondern nur die Verknüpfung der Wörter im Satz. Aus der unterschiedlichen Verbindung der Segmente *o*, *r* und *t* in den Wörtern *rot* und *tor* ergibt sich deren verschiedene Bedeutung sowenig wie aus der Verbindung Punkt Strich die Bedeutung *a* ableitbar ist. Wohl aber determiniert die unterschiedliche Verknüpfung der Wörter den Bedeutungsunterschied z. B. der beiden Sätze *Hans beobachtet die Katze* und *Die Katze beobachtet Hans*.

Diese Beispiele sind in vieler Hinsicht *simplifiziert*,⁴⁸ Sie machen jedoch den Gesichtspunkt deutlich, unter dem Syntax im engeren Sinn von Strukturbildung im allgemeinen zu unterscheiden ist. Ich will diesen Unterschied allgemeiner fassen und halte zu diesem Zweck zunächst in der weitgefaßten Bedingung (S 0) fest, was unter der Struktur *R* eines komplexen Gebildes *K* verstanden werden soll:

(S 0) *K* hat bezüglich der Einheiten $E(1), E(2), \dots, E(n)$ die Struktur *R* dann und nur dann, wenn die $E(i)$ (für $i = 1, 2, \dots, n$) Teile von *K* sind und *K* durch die $E(i)$ und *R* vollständig determiniert ist.

Damit ist nicht mehr gesagt, als daß man von einer Struktur nur sprechen kann in bezug auf die Elemente, aus denen ein Gebilde besteht, und daß die Struktur die Art ist, in der das komplexe Gebilde aus seinen Elementen hervorgeht. Die Verknüpfungen der Segmente in einer Wortform oder der Töne in einem Thema sind einfache Beispiele.

Was aber macht *R* zur syntaktischen Struktur von *K*? Der wesentliche Punkt in unseren Beispielen war der, daß *K* bestimmte zusätzliche Eigenschaften aufweist – im Fall sprachlicher Ausdrücke ist das die Bedeutung und alles, was von ihr abhängt –, die sich aus entsprechenden Eigenschaften der Elemente und der Struktur *R* ergeben. Ich will dies das *funktionale Prinzip der Syntax* nennen und in folgender Bedingung ausdrücken:

(S 1) *R* ist die syntaktische Struktur von *K* bezüglich *A*, wenn *A* die Gesamtheit der Eigenschaften von *K* ist, die durch *R* auf Grund entsprechender Eigenschaften $A(i)$ der Bestandteile $E(i)$ von *K* determiniert wird.

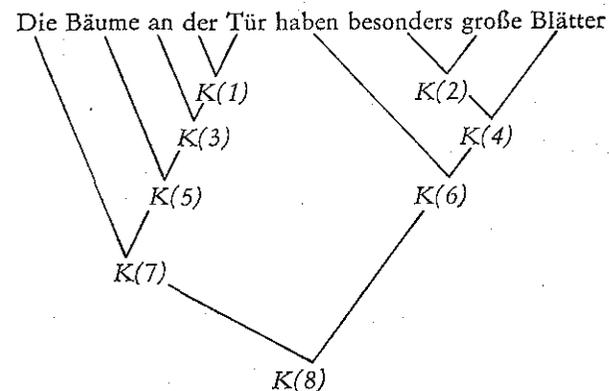
Die Formulierung dieses Prinzips verlangt eine Reihe von Zusätzen, auf die ich hier der Übersichtlichkeit wegen verzichte.⁴⁹ Wie es auf die oben betrachteten Beispiele anzuwenden ist, dürfte hinreichend deutlich sein: Phonetische Merkmale und Segmente haben keine Bedeutung, ihre Verknüpfung kann mithin auch nicht die Bedeutung von Wörtern determinieren. Wortbedeutungen aber legen zusammen mit der Art der Wortverknüpfung die Bedeutung von Sätzen und Texten fest. Wir werden sehen, daß sich aus (S 1) für die Musik Konsequenzen ergeben, die sich erheblich von denen für die Sprache unterscheiden.

Zunächst aber ist der Charakter der Verknüpfungsstruktur *R* noch etwas näher zu bestimmen. Ich beschränke mich dabei auf die Aspekte, die für den Vergleich von Musik und Sprache wesentlich sind.

Die einfache Aneinanderreihung von Grundelementen ist nur ein begrenzter Sonderfall der Strukturbildung, der schon im Verhältnis Wort–Satz nur scheinbar gilt. Tatsächlich beruhen Musik und Sprache auf der schrittweisen Zusammenfassung von Elementen zu Komplexen und dieser zu höheren Komplexen, die wiederum höhere Komplexe bilden usw. Dieses Verfahren ist grundlegend für die Organisation allen menschlichen Verhaltens. Schon eine relativ einfache Handlung, sagen wir Kaffeekochen, besteht aus mehreren Teilhandlungen, etwa Kaffeemahlen, Wasser erhitzen und Filtern des Kaffees, die wiederum aus Teilhandlungen aufgebaut sind, etwa das Mahlen aus dem Einfüllen des Kaffees in die Mühle, dem Mahlvorgang und dem Ausleeren. In der Tat haben Miller, Galanter und Pribram (1960) diese Tatsache zur Grundlage einer generellen Analyse komplexer Verhaltensstrukturen gemacht. Ich will die angedeutete Eigenschaft der Struktur *R* das *Hierarchieprinzip* nennen und durch die Bedingung (S 2) festhalten:

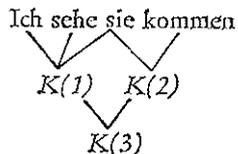
(S 2) Ein komplexes Gebilde *K* ist hierarchisch strukturiert, wenn mindestens eins seiner Teile ein Komplex *K'* ist, so daß die Struktur *R'* von *K'* Teil der umfassenderen Struktur *R* von *K* ist.

Im einfachsten Fall ergeben sich auf Grund des Hierarchieprinzips Beziehungen, die durch einen Stammbaum dargestellt werden können:



Die Komplexe $K(1), K(2)$ usw. zeigen, wie der Gesamtsatz $K(8)$ schrittweise aus seinen Teilen aufgebaut ist und wie dabei zugleich die Integration der Bedeutung der Teile bestimmt wird. Das Hierarchieprinzip läßt aber auch wesentlich verwickeltere Verhältnisse zu, die keinen einfachen Stammbaum ergeben. So kann etwa ein Arzt die Tür seines Ordinationszimmers öffnen, um einen Patienten zu entlassen und den nächsten zu empfangen. Eine Teilhandlung – Türöffnen – nimmt dann in zwei verschiedenen Komplexen einen unterschiedlichen Stellenwert ein. Ähnlich hat in dem Satz *Ich*

sehe sie kommen das Pronomen *sie* gleichzeitig zwei verschiedene Funktionen, es gehört zu *ich sehe sie* und zu *sie kommen*. *R* ist in solchen Fällen kein Baum, sondern ein Netz:



Von ähnlicher Art sind die häufigen Fälle, in denen ein Ton oder Akkord zugleich eine musikalische Passage abschließt und eine neue Sequenz eröffnet. Ein reizvolles und noch komplizierteres Beispiel sind die Übergänge zwischen den *Trois Danses* in der *Histoire du Soldat* von Strawinsky, in denen sich *Tango* und *Walzer* bzw. *Walzer* und *Ragtime* über viele Takte sehr artistisch überlappen. Den Spielraum der mit dem Hierarchieprinzip umschriebenen Möglichkeiten genauer oder gar vollständig zu charakterisieren, ist hier nicht möglich; ich beschränke mich deshalb auf diese Andeutungen verschiedener Arten komplexer Strukturbildung.

Da Sprache und Musik zeitlich und damit linear organisierte Zeichensysteme sind, ergibt sich als drittes Prinzip die Notwendigkeit, die Einheiten und Teilkomplexe einer komplexen Struktur, gleichgültig durch welche Zusammenhänge hierarchischer Art sie verknüpft sind, auf die Zeitachse zu projizieren. Ich will diese Bedingung das *Linearitätsprinzip* nennen, das sich folgendermaßen formulieren läßt:

(S 3) Ein komplexes Gebilde *K* ist linear bezüglich *D*, wenn die durch die Struktur *R* verknüpften Teile von *K* in bezug auf *D* geordnet sind.⁵⁰

Durch das Linearitätsprinzip kommt in die Struktur *R* eine zusätzliche Ordnungsbildung hinein, oder anders ausgedrückt: Es wird in *R* eine bestimmte Relation, nämlich die Anordnung in der Dimension *D*, ausgezeichnet. Für Musik und gesprochene Sprache ist *D* die Zeitachse, für geschriebene Sprache ist *D* die Zeilenanordnung von links nach rechts (bzw. von rechts nach links oder von oben nach unten).

Im trivialen Fall, der rein linearen Verknüpfung, fällt *R* mit der Anordnung in *D* zusammen. Das gilt z. B. für die Strukturbildung des Morse-Codes. Liegt Hierarchiebildung vor, dann ist der einfachste Fall der, daß alle Teile eines Teils auch linear in ihm enthalten sind. Die Anordnungsrelation läßt sich dann einfach als links-rechts-Orientierung im entsprechenden Strukturschema repräsentieren, wie ich das oben bereits getan habe. Man sagt dann, die Hierarchie ist „projektiv“ bezüglich der Anordnung. Sowohl in der Sprache wie in der Musik treten aber aus unterschiedlichen Gründen wesentlich kompliziertere Verhältnisse auf. In einem Satz wie *Er kündigte*

ein Lied an, das alle kannten bilden kündigte an und ein Lied, das alle kannten zusammengehörige Konstituenten, die aber linear aufgeteilt sind; der Satz ist nicht projektiv. Die Gründe dafür, daß natürliche Sprachen solche Diskrepanzen zwischen linearer und hierarchischer Struktur aufweisen, liegen in der Funktionsweise des menschlichen Kurzzeitgedächtnisses begründet. Grundsätzlich aber sind sprachliche Zeichen auch bei beliebiger Komplexität strikt linear, was heißen soll, daß ein Intervall von *D* immer nur von einem Elementarzeichen belegt sein kann. Für die Musik gilt diese Bedingung nicht: Jede Form von Mehrstimmigkeit beruht auf einer Linearisierung, in der mehrere strukturelle Bestandteile eines komplexen Zeichens den gleichen Platz in *D* einnehmen, und zwar so, daß das zur Struktur *R* des komplexen Zeichens gehört.

Die vier Bedingungen, die ich damit erläutert habe, sind von unterschiedlichem Gewicht und hängen auf verschiedenartige Weise miteinander zusammen: (S 0) ist Voraussetzung für alle weiteren Bedingungen – ohne Struktur keine Syntax –, während (S 1) bis (S 3) frei kombinierbar sind. Dabei ist (S 1) die Bedingung dafür, daß Syntax im engeren Sinn überhaupt zustandekommt, während (S 2) und (S 3) syntaktische Strukturen verschiedenen Charakters unterscheiden. Das folgende Schema illustriert durch je ein Beispiel solche Syntaxtypen. (Dabei besagt „+“, daß die Strukturen des Systems der entsprechenden Bedingung unterliegen, „–“ besagt, daß die Bedingung nicht zutrifft.)

	Struktur- bildung (S 0)	Funktio- nalität (S 1)	Hierarchie- bildung (S 2)	Linearität (S 3)
Verkehrszeichen	+	–	–	–
Freges „Begriffsschrift“	+	+	+	–
Zifferndarstellung der Zahlen	+	+	–	+
Strukturformel- system der Chemie	+	+	–	–
natürliche Sprache	+	+	+	+

Ich gehe auf die Eigenarten der angeführten Beispiele hier nicht im einzelnen ein.⁵¹ Die Tabelle soll lediglich die Kombinierbarkeit der Bedingungen (S 1) bis (S 3) andeuten.⁵² Bemerkenswert ist jedoch, daß die Musik, soweit ich sehe, das einzige Zeichensystem ist, dessen Syntax die Bedingung der Linearität erfüllt, ohne strikt linear im oben angedeuteten Sinn zu sein. Daß

sie diese Eigenart hat, liegt natürlich an der Art der Signalerzeugung, auf der die Musik beruht. Wie sie genutzt und interpretiert wird, wird noch zu erörtern sein.

Mit den Bedingungen (S₀) bis (S₃), für deren präzisere Fassung mathematische Mittel verschiedener Art zur Verfügung stehen, verfügen wir über einen zwar keineswegs kompletten, aber hinreichenden Rahmen, um die *Eigentümlichkeiten der Syntax von Musik und Sprache* zu vergleichen und zu charakterisieren. Sofern die Bedingung (S₁) den Kernpunkt der Syntax korrekt charakterisiert, daß nämlich syntaktische Strukturen Relationsgefüge in komplexen Zeichenformen sind, die deren Bedeutung aus der Bedeutung ihrer Bestandteile organisieren, ergeben sich dabei drei Hauptfragen:

- (a) Was sind die jeweiligen elementaren syntaktischen Einheiten?
- (b) Was sind die maximalen syntaktischen Komplexe?
- (c) Wie bewirken syntaktische Strukturen die Bedeutungsintegration?

Die Frage (a) ist zugleich die nach der „unteren“ Grenze der Syntax; für die Sprache haben wir sie *provisorisch* bereits beantwortet mit der Identifizierung des Wortes als des elementaren Zeichens. (b) fragt zugleich nach der „oberen“ Grenze der Syntax, und (c) muß die Konsequenzen der Bedingungen (S₁) bis (S₃) klären. Alle drei Fragen führen zu wesentlich verschiedenen Antworten für Musik und Sprache; die Differenzen führen in spezieller Weise den Unterschied zwischen Zeigen und Sagen und zwischen den Codierungsarten, von denen sie abhängen, weiter.

15. Zeichen und Text

Ich beginne die Erörterung mit der Frage nach der oberen Grenze der Syntax. Das Funktionalitätsprinzip der Syntax ist zunächst in gewissem Sinn eine spezielle Version der allgemeinen Sentenz, daß das Ganze mehr ist als die Summe seiner Teile. Der Satz *Die Männer unterdrücken die Frauen* ist mehr, genauer: er hat eine speziellere Bedeutung als eine Aufzählung von fünf Wörtern. Ein Lied ist mehr als eine Folge von Tönen, ein Sonatensatz ist mehr als eine Menge von Motiven und Figurationen. Musik und Sprache stellen nun beide die Möglichkeit zur Bildung immer weitergreifender komplexer Gebilde bereit: Wörter werden zu Sätzen verknüpft, aus Sätzen können Abschnitte, aus Abschnitten Kapitel, aus diesen Abhandlungen aufgebaut werden. Ähnlich baut sich eine Sinfonie oder eine Suite aus Sätzen auf, diese aus Teilen wie Exposition, Durchführung, Reprise, Coda, diese wiederum aus Themen, Motiven, Überleitungen und Figurationen. In allen diesen Strukturen lassen sich leicht die Prinzipien der Hierarchiebildung und der Linearisierung erkennen. Bezeichnen wir einmal alle diese Formen zusammenhängender musikalischer oder sprachlicher Gebilde ganz allgemein als Texte, dann ist folgende Frage zu stellen: Ist ein beliebiger, sinnvoll zusam-

menhängender Text, der ja zweifellos mehr ist als eine Ansammlung seiner Teile, ein komplexes, syntaktisch determiniertes Zeichen? Die Frage ist keineswegs unstrittig, und sie läßt sich für Musik und Sprache nicht auf die gleiche Weise beantworten.

Für die Sprache ist sie eindeutig zu verneinen, und das ist mit zwei Argumenten zu begründen, die zwei verschiedene Seiten des Problems betreffen, bildlich gesprochen, die Länge und die Tiefe der Zeichenstruktur. Zum ersten: Wer eine Sprache lernt, muß nicht nur den Wortschatz, sondern auch die grammatischen Regeln erwerben. Deren Geltungsbereich ist ziemlich genau begrenzt auf die Sätze der Sprache (was nicht wirklich eine Längenbegrenzung ist, da es streng genommen keinen längsten Satz einer Sprache gibt). Während es mithin verschiedene Regeln der Satzbildung im Englischen, Russischen, Deutschen gibt, gibt es keine verschiedenen Regeln der Textbildung. Die Komplexbildung über den Satz hinaus, heißt das, ist nicht mehr Sache der einzelnen Sprachen, sondern unterliegt generellen Prinzipien. Das könnte zu der Schlußfolgerung führen, daß hier Regeln ins Spiel kommen, die universell für alle Sprachen gelten, die also einen allgemeineren Charakter haben, aber deshalb nicht weniger die syntaktische Strukturbildung determinieren als die einzelsprachlichen Regeln, die innerhalb des Satzes gelten. Auch wenn das richtig wäre, hätten wir doch einen wichtigen Unterschied zwischen Satz und Text aufgespürt.

Hier aber kommt das zweite Argument ins Spiel. Die Bedingungen, nach denen die Zusammenhänge in einem Text, also zwischen den Sätzen, den Abschnitten, den Episoden einer Erzählung oder den Argumentationsschritten einer Darlegung etwa, gebildet und verstanden werden, hängen keineswegs nur von den in der Zeichenform manifesten Relationen ab. Es spielen hier vielmehr Kenntnisse über allgemeine Sachzusammenhänge, Vorstellungen über Situationsgegebenheiten, unausgedrückte Schlußfolgerungen, kurz: eine Vielzahl von Hintergrundfaktoren eine Rolle. Die Form eines Satzes legt seine Bedeutung fest, aber noch nicht seine volle, sinngemäße Einordnung in den Kontext. Was hier zum Tragen kommt, sind die in Abschnitt vier bereits erörterten Stufen beim Verstehen einer Mitteilung: Das Verstehen der Bedeutung der Sätze, die durch die sprachliche Struktur determiniert wird, ist Voraussetzung für das Erfassen des Sinnzusammenhangs in einem Text; aber in den Sinn des Textes gehen mehr Faktoren ein als die sprachlich bedingte Bedeutung. Diese verschiedenen Stufen des Verstehens, auf die die *Metapher der ‚Tiefe‘* zielte, spielen auch schon im einzelnen Satz eine Rolle. Der Punkt, auf den es ankommt, ist aber der, daß ein Satz zunächst eine Bedeutung auf Grund seiner Zeichenstruktur hat, die dann in weitere Sinnbezüge eingeordnet wird, während im Text die Zusammenhänge zwischen den Sätzen allein durch diesen zweiten Schritt konstituiert werden.

Ich fasse diese Argumentation zusammen: Die Sätze einer Sprache sind komplexe Zeichen, deren Bedeutung durch die Kenntnis der lexikalischen und syntaktischen Regeln der jeweiligen Sprache determiniert ist. Mit diesen

Zeichen werden Mitteilungen gebildet, die die Satzbedeutungen in allgemeine Kenntnisse, Kontext- und Situationszusammenhänge eingliedern.

Für die Musik entfällt im Hinblick auf das Verhältnis von Zeichen und Text zunächst das erste der beiden eben erörterten Argumente: Eine Unterscheidung, die der zwischen Satz und Text in der Sprache analog wäre, läßt sich nicht treffen, es gibt in der Musik keine Regeln, die den auf den Satz begrenzten der einzelsprachlichen Grammatik entsprächen. Ein musikalischer „Text“ – ob man dabei an ein Lied, einen Fugen-, einen Sonatensatz, eine Sinfonie, eine Oper oder an rituale Gebrauchsmusik denkt – besteht nicht wie ein sprachlicher Text aus Sätzen. Daß er gegliedert ist, aus Teilen besteht, die komplizierte Beziehungen untereinander haben können, ist unstrittig, doch sind diese Teile nicht Sätze in einem der Sprache vergleichbaren Sinn, weil sie nicht durch einen eigenen Typ von Regeln determiniert sind. Das einfachste Indiz dafür: Man lernt, um einen musikalischen Code zu verstehen, nicht Regeln der Satzbildung, die im sprachlichen Sinn von denen der Textbildung unterscheidbar sind.⁵³

Anders verhält es sich mit dem zweiten Argument: So wie ein sprachlicher Text mehr enthält als die Bedeutung der Sätze, ist auch ein musikalischer Text nicht erschöpft mit der Bedeutung der Zeichen. Diese Behauptung muß zunächst unsinnig oder zumindest unklar erscheinen, da ich soeben gezeigt habe, daß sich in der Musik so etwas wie größte komplexe Zeichen vom Text gar nicht unterscheiden lassen, die Textbedeutung und die Bedeutung des jeweils großen Zeichens zusammenfällt. Es genügt aber, daran zu erinnern, daß auch ein einzelner Satz bereits über seine Bedeutung hinaus auf eine zweite Verstehensstufe zu beziehen ist. (*Es regnet* kann eine erleichterte Feststellung, eine Warnung, eine Aufforderung, den Regenschirm mitzunehmen, und vieles andere sein.) Die Bedeutung musikalischer Zeichen welcher Größenordnung auch immer, die ich als ihre gestische Form charakterisiert habe, ist mithin nicht identisch mit dem Sinn eines musikalischen Textes. Sie ist eine Voraussetzung für sein Verständnis, so wie die logische Form sprachlicher Ausdrücke Bedingung für das Verständnis sprachlicher Texte ist. Die zweite Verständnisstufe indessen beruht in der Musik wie in der Sprache auf Zusammenhängen, die auf die Bedeutung der Zeichen Bezug nehmen, aber sich nicht in ihr erschöpfen. Sie sind naturgemäß für die Musik von ganz anderer Art als für die Sprache. Während dort die Bezüge zwischen den repräsentierten Sachverhalten den weiteren Verständnisrahmen determinieren, ist die Musik auf die Beziehungen im Ablauf der gezeigten gestischen Strukturen verwiesen. Der Aufbau solcher Beziehungen kann von sehr unterschiedlicher Komplexität sein. Faßt man ihre Identifizierung als kognitive Leistung auf, und das ist eine durchaus berechtigte Charakterisierung, dann erklärt das, warum Musik nicht nur emotionale, sondern auch geistige Prozesse auslöst. Es ist klar, daß dies Prozesse sind, die nicht Anleihen bei sprachlich vermittelten Erkenntnisleistungen machen müssen, sondern eine Grundlage *sui generis* haben, nämlich artikulierte gestische Strukturen, die

die Bedeutung musikalischer Zeichen konstituieren. Die Erfassung des Sinns, den die modifizierte Wiederkehr des durch das Hauptthema repräsentierten Gestus in der Durchführung des Sonatensatzes hat, ist ein spezielles Beispiel der kognitiven Leistungen, von denen hier die Rede ist. (In diesem Bereich ist das zu entwickeln, was oben mit dem Begriff der musikalischen Logik angedeutet wurde.)

Um zu resümieren: Syntaktische Strukturen, die die Bedeutung komplexer Zeichen organisieren, beziehen sich in der Sprache auf Sätze als die maximalen Einheiten, während es in der Musik eine analoge Grenzziehung nicht gibt. Wohl aber sind in Sprache und Musik gleichermaßen syntaktisch bedingte Bedeutungsstrukturen zu unterscheiden von Sinnzusammenhängen, die über die Bedeutung hinausgehen und nicht durch die Syntax der Zeichenstruktur allein determiniert sind. Um noch einmal die Metapher von Länge und Tiefe zu bemühen: In der Sprache ist die Geltung der Funktionalität der Syntax in Länge und Tiefe begrenzt, in der Musik nur in der Tiefe. Die hinter der Bedeutung liegenden Sinnzusammenhänge sind in der Sprache die kontextuellen und allgemeinen Sachbezüge, in die Propositionen eingeordnet werden, in der Musik sind es die Bezüge und Prozesse, in die die gestischen Strukturen integriert werden.

16. Motivierte und konventionelle syntaktische Strukturen

Die soeben getroffenen Feststellungen erhalten ihre eigentliche Begründung aus der verschiedenen Art, in der syntaktische Beziehungen die Bedeutungsstruktur organisieren. Dies führt zur Frage (c) aus Abschnitt 14. Für deren Erörterung muß ich auf die in Abschnitt 10 erläuterten Codierungsformen zurückgreifen, die ich zunächst ohne Unterscheidung elementarer und komplexer Zeichen diskutiert habe. Es ist nun zu verdeutlichen, wie der analog-motivierte oder der konventionelle Charakter eines Codierungssystems sich in seiner Syntax auswirkt, sofern das System komplexe Zeichen besitzt.

Der Schlüsselpunkt ist das Funktionalitätsprinzip (S 1), das die Verbindung formaler Relationen mit Bedeutungszusammenhängen garantiert. Daß diese Verbindung motivierter oder konventioneller Art sein kann, zeigt auf einfache Weise die Gegenüberstellung des Morse-Codes und der Zifferndarstellung der Zahlen. Beide haben eine formal gleiche, sehr einfache Syntax: die rein lineare Verknüpfung der Grundzeichen, nämlich der Punkt-Strich-Kombinationen im ersten Fall, der Ziffern im zweiten. Im Morse-Code ist diese Syntax auf direktem Weg analog interpretiert: Der linearen Verknüpfung zweier Zeichen entspricht die lineare Verknüpfung ihrer Bedeutung. Im Zahlencode dagegen ist die gleiche formale Operation mit einer konventionellen Interpretation verbunden: Jedem Verknüpfungsschritt entspricht die Anwendung einer komplizierten Rechenregel (vgl. Anm. 51).

Ein paralleles Verhältnis besteht auch zwischen den komplexeren, weil im Sinn von (S 2) hierarchischen Strukturen der Musik und der natürlichen Sprache. Mit anderen Worten: Die Syntax der Musik wird analog interpretiert, die Syntax der Sprache konventionell. Die Art, in der syntaktische Strukturen der Musik die Bedeutung der komplexen Zeichen organisieren, läßt sich in zwei Regeln (oder besser Prinzipien) formulieren:

- (R 1) Gliederungen in der Form musikalischer Zeichen entsprechen gleichen Gliederungen in der Bedeutung.
- (R 2) Die Zeitstruktur der Form musikalischer Zeichen entspricht der Zeitstruktur ihrer Bedeutung.

Mit (R 1) wird das Hierarchieprinzip (S 2) für die Musik interpretiert, mit (R 2) das Linearitätsprinzip (S 3). Beide entsprechen der mehrfach erörterten Isomorphie von Form und Bedeutung musikalischer Zeichen. (R 2) folgt überdies direkt aus der in Abschnitt 12 diskutierten Zeitstruktur der gestischen Form. (R 1) und (R 2) dürfen nicht verwechselt werden mit den (in bestimmten Grenzen konventionellen) Regeln für die musikalische Strukturbildung, die für verschiedene musikalische Codes erheblich differieren, und für die die Regeln der funktionellen Harmonik, des Kontrapunkts, der Zwölftontechnik oder auf andere Weise der indischen Rāgas als Beispiel dienen können. Dies sind in ihren Prinzipien und in ihrem speziellen Inhalt eigens zu charakterisierende Regelsysteme, die den Aufbau möglicher Zeichenformen determinieren. (R 1) und (R 2) legen nur fest, wie die so gebildeten komplexen Strukturen mit ihren Bedeutungen verbunden sind. Zwei einfache Beispiele mögen die Konsequenzen aus (R 1) und (R 2) andeuten. Die beiden in ihrem Charakter entgegengesetzten Motive des Hauptthemas im ersten Satz der *Jupiter-Sinfonie* von Mozart zeigen zwei aufeinanderfolgende Gesten, die in eben dieser Folge zu einem zweiteiligen Gesamtgestus verbunden werden. Zu einem „zweischichtigen“ Komplex kontaminiert wird etwa in Schuberts Vertonung des *Erlkönig* der durch den pochenden Grundrhythmus durchgängig festgehaltene Gestus atemloser Eile mit dem durch die Singstimme repräsentierten Gestus. (Das immer wiederkehrende Ostinato-Motiv macht diesen gestischen Komplex im Wechsel sogar dreischichtig.) Daß solche Feststellungen nur scheinbar trivial sind, wird deutlich, wenn wir sie der ganz andersartigen Syntax der Sprache gegenüberstellen.

Wer eine Sprache beherrscht, interpretiert deren Strukturen ebenso automatisch wie die der musikalischen Verknüpfung, ohne zu bemerken, daß einfache Analogieprinzipien wie (R 1) und (R 2) hier fehlen, daß syntaktische Relationen vielmehr nach verschiedenartigen konventionell entstandenen Regeln die Bedeutung determinieren. Ich belege zunächst mit einigen charakteristischen Beispielen, daß Analogieprinzipien vom Typ (R 1) und (R 2) in der Sprache nicht gelten.

(1) *Der Veranstalter stellt das Programm wieder um.*
In der Bedeutung bilden *stellt um* eine Einheit, formal sind sie durch zwei intermittierende Einheiten *das Programm* und *wieder* getrennt. Umgekehrt bilden in

(2) *Diesen Fisch hat er versucht, an die Angel zu bekommen*
die ersten fünf Wörter eine relative Einheit – etwa wie in *Diesen Fisch hat er gefangen* – obwohl *diesen Fisch* in (2) natürlich zu *an die Angel bekommen* gehört. Noch offensichtlicher wird das Fehlen von Analogiebeziehungen, wenn wir die Zeichenform über die Bedeutung hinaus auf die Sachverhalte beziehen, über die etwas gesagt wird:

(3) *Die Hauptstadt von England liegt an der Themse.*
Die Einheit *Die Hauptstadt von England* enthält *England* als Untereinheit, London aber enthält nicht England als Teil. Die Beispiele (1) bis (3) zeigen zugleich, daß nicht nur Hierarchiebeziehungen, sondern auch Reihenfolgebeziehungen in der Zeichenform keine Analogie in linearen Beziehungen der Bedeutung haben.⁵⁴ In der Musik sind Erscheinungen der illustrierten Art durch die Bedingungen (R 1) und (R 2) ausgeschlossen.

Einen Schritt tiefer als die Beispiele (1) bis (3) führt die Feststellung, daß die syntaktischen Relationen in der Sprache, anders als die der Musik, unterschiedliche spezifische Funktionen haben, je nachdem, welche Einheiten sie verknüpfen. So sind zum Beispiel *dicke Bücher* und *viele Bücher* formal gleichartige Verbindungen zweier Wörter. Aber die Bedeutung von *viele* bezieht sich nicht nur in ganz anderer Weise auf die Bedeutung von *Bücher* als die Bedeutung von *dicke*. Es entstehen auch zwei Einheiten, die auf ganz verschiedene Weise in weitere Strukturen eingegliedert werden. Dieser zunächst ganz verdeckte Unterschied kommt in folgendem Vergleich zum Vorschein:

(4) (a) *Niemand liest dicke Bücher.*

(b) *Dicke Bücher werden von niemandem gelesen.*

Obwohl formal ganz unterschiedlich strukturiert, haben diese beiden Sätze ziemlich genau die gleiche Bedeutung. Nicht aber die gleichen Sätze mit *viel* anstelle von *dicke*:

(5) (a) *Niemand liest viele Bücher.*

(b) *Viele Bücher werden von niemandem gelesen.*

Wie die Verschiedenheiten in den Beispielen (4) und (5) zustandekommen, läßt sich hier nicht im einzelnen erklären. Sie sind jedoch charakteristische Belege der Tatsache, daß syntaktische Strukturen in der Sprache die Bedeutung in sehr viel differenzierterer Weise determinieren als dies die Prinzipien (R 1) und (R 2) in der Musik tun. Die Determination hängt in der

Sprache, um das wenigstens anzudeuten, von einem speziellen Kategoriensystem ab, durch das die elementaren und komplexen sprachlichen Zeichen klassifiziert werden. *Dick, dünn, rot* z. B. gehören in diesem Sinn zu einer anderen Kategorie als *viel, wenig, manch*. Mit verschiedenen Kategorien sind dann, um in Analogie zur Syntax des Zahlencodes zu sprechen, jeweils verschiedene „Rechenregeln“ verbunden. Dabei gilt hier wie in der Musik, daß die Regeln, die den formalen Aufbau komplexer Zeichen bestimmen, nicht identisch sind mit den Prinzipien, nach denen dieser formale Aufbau die Bedeutung determiniert. Gleiche formale Mittel können in verschiedenen Sprachen sogar ganz unterschiedliche Bedeutungszusammenhänge bestimmen und umgekehrt. Und das heißt zugleich, daß die Funktion, die der formale Aufbau komplexer Zeichen für deren Bedeutung hat, nicht durch Analogieprinzipien motiviert, sondern konventionsbedingt ist.⁵⁵ Der Spielraum der konventionellen Variation ist dabei wiederum abgesteckt durch universelle Bedingungen möglicher Sprachstrukturen.

Nach dieser erläuternden Gegenüberstellung läßt sich der unterschiedliche Rahmen, in dem sich die bedeutungsorganisierende Funktion der Syntax in Musik und Sprache bewegt, folgendermaßen abstecken:

Erstens: Jede Grammatik und jeder musikalische Code enthält ein System von Regeln zur Bildung komplexer Zeichenformen. Diese Regeln sind Verknüpfungsoperationen, die linear geordnete, hierarchisch gegliederte Strukturen erzeugen, oder Transformationen, die solche Strukturen modifizieren. Die Klasse von Operationen, die für solche Regeln genutzt werden, ist für Musik und Sprache sehr unterschiedlich. (Für die Sprache gibt es Ansätze zu ihrer systematischen Bestimmung [vgl. Chomsky, 1965], für die Musik steht ein solcher Versuch noch aus.) Zwei charakteristische Unterschiede der resultierenden Strukturen habe ich bereits genannt: In der Musik können mehrere Einheiten linear gleichzeitig auftreten, in der Sprache nicht; in der Sprache können hierarchisch zusammengehörige Einheiten linear getrennt sein, in der Musik nicht. Andere wesentliche Unterschiede kommen hinzu, so z. B. der, daß Transformationen in der Musik rein linearen Charakter haben können – die Bildung der Krebsform eines Themas ist die strengste Form einer solchen Operation –, während in der Sprache alle Operationen an hierarchische Strukturen gebunden sind.

Zweitens: Die so erzeugten strukturellen Beziehungen determinieren die Bedeutung der resultierenden komplexen Zeichen. In der Musik geschieht das durch die beiden Prinzipien (R 1) und (R 2), in der Sprache durch die „Rechenregeln“, die an die hierarchischen und linearen Beziehungen zwischen den verknüpften Einheiten und an deren Kategorienzugehörigkeit gebunden sind. Für die Musik ergibt sich aus (R 1) und (R 2) die Isomorphie der formalen und der Bedeutungsstruktur, die Sprache dagegen ist durch eine charakteristische Strukturverschiedenheit von Form und Bedeutung gekennzeichnet.

Drittens: Aus Abschnitt 15 wissen wir, daß die Regeln für Aufbau und

Interpretation komplexer Zeichen auf den Satz begrenzt sind, während es für die Musik eine solche Grenze nicht gibt. Mit anderen Worten, das durch (R 1) und (R 2) umschriebene Verfahren der Bedeutungsdetermination erstreckt sich in der Musik allemal auf den gesamten zusammenhängenden „Text“, die Bedeutungsdetermination der Sprache dagegen nur auf die Sätze. Deren lineare Abfolge hat keinen syntaktischen Charakter im Sinn der Bedingung (S 1) mehr. Zusammenhänge zwischen Sätzen werden nicht als Bedeutungs-, sondern als Sinnbezüge, also im Hinblick auf den zweiten Verstehensschritt, konstituiert. Dabei können mehrere aufeinanderfolgende Sätze entweder aufeinanderfolgende Ereignisse darstellen (etwa in einer Erzählung), oder verschiedene Teilsachverhalte einer einzigen Situation (etwa in einer Beschreibung), oder auch Schritte in einem Argumentationsgang, oder Kombinationen dieser Zusammenhangstypen. Eine einheitliche Interpretation der Satzfolge im Sinn von (R 2) gibt es nicht.

17. Elementarzeichen

Um das Bild der syntaktischen Funktionsprinzipien zu vervollständigen oder überhaupt erst zu einem Bild zu machen, müssen wir uns nun dem Problem der Elementarzeichen, also der Frage (a) aus Abschnitt (14), zuwenden. Dabei wird eine Reihe von früheren Feststellungen wieder aufzugreifen und weiterzuführen sein.

Grund- oder Elementarzeichen sind die kleinsten Zeichen, in die die komplexen Zeichen eines gegebenen Codes zerlegt werden können. Die Syntax baut auf ihnen auf und gibt ihnen damit zugleich ihren speziellen Charakter. Um die hier konzentrierten Unterschiede zwischen Musik und Sprache deutlich machen zu können, fasse ich die bisherigen Überlegungen zunächst in folgender Kennzeichnung der Bedingungen für ein Elementarzeichen zusammen:

- (Z) Z ist ein Elementarzeichen eines Codes C genau dann, wenn es folgende Bedingungen erfüllt:
- $Z = (F, B)$ ist ein geordnetes Paar aus der Zeichenform F und der Bedeutung B ;
 - F ist eine akustische Merkmalsstruktur, die festlegt, was Z entsprechend den Regeln von C zum Aufbau der Form komplexer Zeichen beiträgt, in denen es vorkommt;
 - B legt fest, was Z entsprechend den Regeln von C zur Bedeutung komplexer Zeichen beiträgt, in denen es vorkommt;
 - F gehört einer syntaktischen Kategorie k an, die die Verknüpfungsmöglichkeiten von Z im Code C determiniert;
 - B gehört einer semantischen Kategorie t an, die die semantischen Verbindungsmöglichkeiten von Z im Code C determiniert;
 - Z ist nicht nach Regeln von C aus anderen Zeichen aufgebaut.

Die Bedingungen (Z a) bis (Z e) rekapitulieren, was wir im Laufe unserer Überlegungen für (akustische) Zeichen allgemein konstatiert haben.⁵⁶ Die spezielle Bedingung für Elementarzeichen ist (Z f). Was sie – zusammen mit (a) bis (e) – besagt, ist für die Sprache zwar mit einigen Schwierigkeiten verbunden, jedoch im Prinzip ziemlich klar, für die Musik dagegen zunächst ganz offen. Sehen wir näher zu.

In der Sprache identifiziert (Z f), wie wir bereits wissen, im großen und ganzen die Wörter – genauer gesagt, die lexikalischen Einheiten – als die Grundzeichen des Codes. In der Differenz zwischen Wort und lexikalischer Einheit liegen die erwähnten Schwierigkeiten. Sie lassen sich stellvertretend an Beispielen wie *aufhören*, *umbringen* oder auch Redewendungen wie *schwedische Gardinen* demonstrieren. In einem Satz wie *Sie bringen ihn um* spaltet sich die lexikalische Einheit *umbringen* in zwei Wörter, ohne daß dem eine entsprechende Aufteilung der Bedeutung entspräche. Und die Bedeutung von *schwedische Gardinen*, nämlich „Gefängnisgitter“, ergibt sich nicht aus den beiden Wörtern mit ihrer jeweils eigenen Bedeutung. Probleme dieser Art beruhen darauf, daß die Struktur der Form und der Bedeutung sprachlicher Zeichen nicht isomorph ist. Trägt man dieser Tatsache mit geeigneten Mitteln Rechnung – und die linguistische Theorie hält diese Mittel bereit –, dann läßt sich sowohl für die syntaktische wie die semantische Ebene sagen: ein komplexes Zeichen läßt sich eindeutig in die Grundzeichen zerlegen, aus denen es aufgebaut ist.

Eben diese Feststellung ist in ganz anderer Weise problematisch für die Musik, trotz oder sogar wegen der *prinzipiellen Isomorphie von Form und Bedeutung*. Jeder, der ein nicht völlig triviales musikalisches Gebilde unter diesem Gesichtspunkt zu analysieren versucht, stößt auf dieses Problem. Aus wieviel Grundzeichen besteht z. B. das Hauptthema der *Fünften Sinfonie* von Beethoven? Falls eine solche Frage überhaupt richtig gestellt ist, sind zu ihrer Beantwortung jedenfalls Kriterien nötig, die in (Z f) nicht enthalten sind. Ich will für den Moment einmal annehmen, daß solche Kriterien zu finden sind und mit ihrer Hilfe eine Zerlegung komplexer Gebilde eines musikalischen Codes in Grundzeichen prinzipiell möglich wäre.⁵⁷ Wir müssen diese Annahme im folgenden modifizieren, können aber auf ihrer Grundlage dennoch einige fundamentale Feststellungen treffen.

Erstens: Jeder sprachliche Code enthält ein festes Repertoire von Elementarzeichen. Traditionell wird es als Lexikon der Grammatik gegenübergestellt. Ein „Lexikon“ musikalischer Grundzeichen gibt es in diesem Sinn nicht. Anders ausgedrückt: Wer sich einer Sprache bedient, kann jederzeit neue Sätze, aber nicht neue Wörter bilden. (Die in Anmerkung 35 erwähnte Möglichkeit, neue Wörter durch Definition einzuführen, hebt dieses Prinzip nicht auf, eben weil ein neues Wort definiert werden muß und nicht einfach gebraucht werden kann wie ein neues Thema.) Wer sich eines musikalischen Codes bedient, kann nicht nur neue komplexe Zeichen aus elementaren erzeugen, sondern stets auch neue Grundzeichen. (Daß die Möglichkeit be-

steht, bereits bekannte Themen bei der Bildung neuer komplexer Gebilde zu verwenden – Beispiel: die *Mozartvariationen* von Reger –, hebt wiederum das Prinzip nicht auf.) Kurz: Die Liste der Wörter einer Sprache ist abgeschlossen, die Liste der Grundzeichen eines musikalischen Codes ist offen.⁵⁸

Zweitens: Daraus ergibt sich, daß man die Wörter eines neuen Satzes schon kennen muß, um den Satz zu verstehen, nicht aber die Grundzeichen, um ein komplexes musikalisches Gebilde zu verstehen. In gewissem Sinn kann Musikverstehen deshalb einen Lernvorgang einschließen, der auf das Sprachverstehen so nicht zutrifft: Mit einem neuen Musikstück lernt man dessen Grundzeichen kennen. Es ist möglich, daß dabei Kategorien von Grundzeichen zu unterscheiden sind hinsichtlich des Lernresultates. Ein Grundzeichen kann auf seine Wiedererkennung hin angelegt sein und auf Grund dieser Identifizierbarkeit im musikalischen Zusammenhang funktionieren; es kann aber auch unterhalb dieser Schwelle, die im wesentlichen durch Mechanismen der Wahrnehmungsstrukturierung und der Gedächtnisfixierung bedingt wird, verbleiben und sozusagen anonym in den Zusammenhang komplexer Strukturen eingehen. Funktional handelt es sich dabei um so etwas wie Haupt- und Begleitmotive⁵⁹, strukturell sind diese Kategorien eher als Extremwerte eines Kontinuums aufzufassen. Jedenfalls aber bezieht sich der Lernvorgang nur auf die Form der Elementarzeichen, nicht die Zuordnung ihrer Bedeutung. Denn:

Drittens: Während sprachliche Grundzeichen arbiträr und konventionell sind, ist die Bedeutung musikalischer Grundzeichen durch ihre Form auf Grund des allgemeinen Analogieprinzips motiviert. Die formale Ähnlichkeit zweier Wörter wie *beginnen* und *besinnen* sagt nichts über ihre semantischen Beziehungen, nicht einmal über ihre Kategorienzugehörigkeit, wie *Tür* und *für* zeigen, ebenso wie formale Unterschiedlichkeit nicht Bedeutungsverschiedenheit anzeigt, wie man an *beginnen* und *anfangen* sieht. Die Form musikalischer Zeichen dagegen zeigt die Ähnlichkeit und Unterschiedlichkeit ihrer Bedeutung. Und dies wiederum heißt: Die Bedeutung musikalischer Grundzeichen muß aus der Kombination der Merkmale und Segmente hervorgehen, aus denen ihre Form aufgebaut ist.

Wir sind damit offensichtlich bei einem Paradox angelangt. Die Syntax der Musik hat nicht nur, wie wir in Abschnitt 15 gesehen haben, keine Grenze nach oben, sondern auch keine Grenze nach unten. Denn wenn sich die Bedeutung dessen, was wir bisher als Grundzeichen betrachtet haben, aus ihrer Struktur ergibt, dann muß das Funktionalitätsprinzip (S 1) auch innerhalb der musikalischen Grundzeichen wirksam sein, und konsequenterweise müßten dann die Merkmale der Laut- und Zeitstruktur als eigentliche Elementarzeichen gelten. Das aber ist widersinnig. Ein Tonhöhenmerkmal hat an sich keine Bedeutung, es zeigt auch nicht ein Fragment eines Gestus. Es ist wichtig, sich diesen Gedankengang klarzumachen, um den Sinn der folgenden These zu verstehen:

(T) Jede musikalische Struktur muß minimale Merkmalskonfigurationen enthalten, die als solche eine gestische Bedeutung haben.

Ich will diese Konfigurationen, in freier Adaption dieses Terminus, „Motive“ nennen. Daß die These (T) gelten muß, ist logisch begründet, ihre Ausfüllung jedoch, die Bestimmung der Kriterien für die Identifizierung der Grundkonfigurationen, ist ein verwickeltes empirisches Problem. Sie kann sich keinesfalls an der Ausdehnung oder Länge bemessen: Einerseits muß z. B. der Hammerschlag in der *Sechsten Sinfonie* von Mahler zweifellos als Motiv im Sinn von (T) gelten, andererseits kann ein Motiv etwa ein über beliebig lange Sequenzen ausgedehntes rhythmisches Grundschema sein. Das bereits erwähnte durchgehende Pochen im *Erköning* ist eins von vielen Beispielen. Ein Motiv darf auch nicht einfach als Segmentfolge bestimmt werden. Es kann vielmehr auch ein rhythmisches Muster sein, das verschiedenen mit Segmentmerkmalen „gefüllt“ wird, oder eine Intervallfolge, die in verschiedenen Segmentkonstellationen erscheint. So baut sich zum Beispiel das Thema des Variationssatzes der Klaviersonate Op. 111 von Beethoven aus zwei Motiven auf, einem Intervallsprung und einem punktierten Rhythmus, die dann auch je für sich in andere Konstellationen eingehen. (Das Spätwerk Beethovens ist wegen der freien und komplizierten Kombinatorik minimaler Motive insgesamt ein interessantes Feld für die empirische Bestimmung der zur Präzisierung von (T) nötigen Kriterien.)

Entscheidend ist ein zweiter Punkt für das Verständnis von (T): Eine musikalische Struktur muß sich nicht lückenlos in Motive zerlegen lassen (wie ein Satz in Wörter), sie muß lediglich Motive enthalten. Motive können dann aber zusammen mit nichtmotivischen Merkmalen und Segmenten größere Zeichen bilden. Anders ausgedrückt: Ausgehend von den motivischen Konfigurationen, die als solche eine gestische Bedeutung haben, nehmen „angelagerte“ Merkmale und Segmente eine gestische Bedeutung an, und zwar nach dem generellen Analogieprinzip der Musik. Damit löst sich das genannte Paradox auf: Am „unteren“, elementaren Ende ist die musikalische Syntax in Motiven verankert; ihre Funktionsweise besteht darin, motivische mit nichtmotivischen Merkmalskonfigurationen zu (dann schon komplexen) Zeichen zu verknüpfen. Sie ist also in gewissem Sinn in der Tat nach unten offen, weil nur partiell in Motiven verankert, aber sie löst sich nicht in einzelne akustische Merkmale auf. Noch einmal verdeutlicht ein vergleichender Blick auf das System der Sprache die zuletzt formulierte Konsequenz. Dazu sind einige Hinweise auf die sogenannte morphologische Struktur der Sprache, die ich bisher ausgeklammert habe, nachzutragen. Wortformen bestehen nämlich nicht allgemein einfach aus bedeutungsindifferenten Segmenten und Silben (vgl. Anmerkung 48), sondern sind vielfach aus Teilen zusammengesetzt, die, vorsichtig ausgedrückt, bedeutungsrelevant sind. Diese Bestandteile heißen Morpheme. Sie können, verliert man die Unterschiede nicht aus den Augen, als sprachliches Analogon zu den Motiven aufgefaßt werden. In Fällen wie *bier*, *auf*, *Spiel* und vielen anderen ist ein Wort zu-

gleich ein einziges Morphem, das Morphem also zugleich ein Minimalzeichen im Sinn von (Z). (Ähnlich kann ein Thema in der Musik mit einem Motiv identisch sein.) Die Fälle, in denen das nicht gilt, die also morphologischen Kombinationsprinzipien unterliegen, will ich grob vereinfachend in zwei Klassen teilen, die in verschiedenen Sprachen sehr unterschiedlich vertreten sind. Die erste Klasse bildet den Bereich der sogenannten Wortbildungsmorphologie, die zweite den der Flexionsmorphologie.

Ich erläutere die erste an Fällen wie *versteinern*, *verdunkeln*. Dies ist ihre morphologische Zerlegung: *ver-stein-er-n*, *ver-dunkel-n*. *Stein* und *dunkel* sind Morpheme, die zugleich echte Grundzeichen sind (sogenannte „Stamm-morpheme“), *-n* ist ein Flexionsmorphem mit einer festen Funktion, auf die ich gleich zurückkomme, *ver-* ist ein Wortbildungsmorphem mit sehr „schwankender“ Bedeutung, wie der Vergleich mit *verglasen*, *vergeben*, *vertreiben* und vielen anderen Wörtern zeigt. Worauf es in unserem Zusammenhang ankommt, ist dies: Eine Wortform ist eindeutig aus einem oder mehreren Morphemen aufgebaut, die auf verschiedene Weise zu ihrer Bedeutung beitragen. Zum kleineren Teil geschieht das nach dem Funktionalitätsprinzip der Syntax auf Grund der Bedeutung der Morpheme, zum größeren Teil durch konventionelle, nicht regelbedingte Festlegung der Bedeutung der Verbindung, die damit ein Elementarzeichen im Sinne von (Z) wird und als lexikalische Einheit gelernt werden muß.

Bei der Flexionsmorphologie handelt es sich demgegenüber um einen strikt regelbedingten Sonderbereich der Syntax. Flexionsmorpheme, die im Deutschen vorwiegend als Endungen erscheinen, hängen direkt mit der Einordnung elementarer Zeichen in komplexe zusammen und unterliegen vollständig dem Funktionalitätsprinzip. Die Verbflexion, belegt etwa durch *leb-en*, (*ich*) *leb-e*, (*du*) *leb-st*, (*er*) *leb-t* usw. macht das deutlich.

Morpheme müssen nicht immer die Form eigenständiger Segmente haben, sie können (wie musikalische Motive) gewissermaßen Veränderungen an anderen Morphemen sein: was in (*er*) *leb-te* durch *-te* repräsentiert wird, kommt in (*er*) *gab* durch Veränderung des Vokals von *geb-en* zum Ausdruck. Ähnlich verhält sich in der Wortbildung *Leit-ung* zu *leit-en* wie *Trieb* zu *treib-en*.

Welche Form ein Morphem hat und was seine Konsequenzen für die Bedeutung sind, ist immer konventionell bedingt, für einzelne Einheiten oder durch Regeln für ganze Klassen. Die Anfügung von Morphemen, etwa *breit*, *breit-er*, *breit-er-er*, *breit-er-er-e* (etwa in *eine breitere Öffentlichkeit*) funktioniert deshalb ganz anders als die „Fortspinnung“ eines musikalischen Motivs, ebenso die Abwandlung eines Motivs im Vergleich zu einer Abwandlung wie *gib* gegenüber *gab* oder *geben*.

Der Vergleich von Motiven und Morphemen führt damit zu den beiden folgenden Konklusionen:

Erstens: Sprachliche Grundzeichen können aus mehr als einem Morphem bestehen, sie enthalten aber keine „nicht-morphemischen“ Merkmale und

Segmente; ihre Bedeutung ist konventionell festgelegt (in einem Rahmen, der durch Wortbildungsregeln strukturiert ist). Motive dagegen können mit nicht-motivischen Merkmalen und Segmenten verknüpft werden. Die resultierenden Strukturen sind nicht mehr Grundzeichen im strengen Sinn von (Z f). Ihre Bedeutung wird durch die Analogieprinzipien (R 1) und (R 2) determiniert.

Zweitens: Die Syntax der Sprache ist nach unten und oben begrenzt: nach oben durch den Satz⁶⁰, nach unten durch die Elementarzeichen, zu denen wir jetzt außer den lexikalischen Einheiten in gewissem Sinn die Flexionsmorpheme hinzunehmen haben. Die Syntax der Musik ist offen nach oben und nach unten. Nach oben, weil sie sich immer auf den ganzen „Text“ ausdehnt; nach unten, weil sie zwar auf den Motiven als den eigentlichen Elementarzeichen beruht, aber nicht vollständig; die Motive können ja mit nicht-motivischen Elementen, die also an sich keine Elementarzeichen sind, verknüpft werden. Drei Überlegungen sollen das verwickelte Problem musikalischer Elementarzeichen noch verdeutlichen.

Erstens: Auf welche Weise sprachliche Zeichen ihre Bedeutung erhalten, wird, sinngemäß vereinfacht, durch Lexikon und Grammatik determiniert. Musikalische Zeichen dagegen erzeugen ihre Bedeutung auf Grund des Analogieprinzips. Wenn man will, läßt sich auch das in zwei Komponenten aufgliedern, die freilich keine Parallele zu Lexikon und Grammatik bilden. Erstens haben Motive als Grundgestalten eine (zunächst meist relativ offene) gestische Bedeutung. Die Wirkung der Gestaltqualitäten zu bestimmen und in geeigneter Form zu charakterisieren, ist eine empirische Aufgabe, keine Sache der Spekulation. Zweitens ergibt die Verknüpfung von Motiven bzw. von Motiven mit nichtmotivischem Material komplexere Strukturen, deren Zusammenhänge entsprechend reichere, also definitere gestische Strukturen repräsentieren. An gestisch interpretierte Motive angelagert, erhalten damit auch nicht-motivische Elemente eine gestische Interpretation. Sowohl die Interpretation der motivischen Grundgestalten wie der Verknüpfung unterliegt dem Analogieprinzip. Was die Bedeutungszuordnung bei musikalischen Zeichen von der bei sprachlichen unterscheidet, ist deshalb nicht ihre unterschiedliche „Konstantisierung“, wie Knepler (1977, S. 134 ff.) annimmt, sondern der unterschiedliche Charakter der Bedeutung und die entsprechend verschiedener Art der Zuordnung. Im Hinblick auf den jeweiligen Bedeutungsbereich ist die Bedeutung etwa des Hauptthemas aus dem ersten Satz von Bartóks *Zweiten Violinkonzert* nicht weniger konstant und präzise als die des Wortes *Lust* oder *Höhe* oder des Satzes *Und die einen sind im Dunkel und die andern sind im Licht*. Löst man sich von der Fiktion, daß die Musik auf verschleierte Weise Aussagen macht (Knepler 1977, S. 134), dann verschwindet der Grund, von geringerer Konstanz zu sprechen. Die Bedeutungskonstanz der Musik wird durch ihre Struktur und die analoge Codierungsweise garantiert, so wie die der Sprache durch deren Struktur und die konventionelle Codierung.

Zweitens: Es läßt sich nun in einem strengeren Sinn begründen, warum „außermusikalische“ Bedeutungen konventionalisierter Signale nicht in die musikalische Bedeutung eingehen. Zum einen besagt ihr konventionalisierter Charakter gerade, daß sie nicht Gesten zeigen, daß es sich also um eine Bedeutung von anderer Art handelt, die – wie oben bereits gesagt – aus einem anderen Codierungszusammenhang stammen muß. Zum anderen kann deshalb gerade dieser Bedeutungsaspekt nicht von der musikalischen Syntax erfaßt werden, wenn solche Signale in einen musikalischen Zusammenhang eingehen. Das heißt aber: musikalisch funktionieren sie zunächst auf Grund ihrer „vorkonventionellen“ Bedeutung. Das schließt freilich keineswegs aus, daß außermusikalische Bedeutungen, die auf die eine oder andere Weise mit musikalischen Zeichen verbunden sind, in den Sinnzusammenhang eingehen, der durch die Bedeutung in jenem zweiten Verstehensschritt erzeugt wird. Und dies gilt nun nicht nur für echte, konventionalisierte Signale, sondern auch für offenere Fälle. Militärsignale in Mahlers Sinfonien und Liedern zum Beispiel, die ja keine wirklichen Signale, sondern wie Signale sind, unterliegen zunächst den Prinzipien der musikalischen Syntax und ihrer semantischen Interpretation, und erst durch diese hindurch ordnen sie das Signalhafte in den Sinnzusammenhang ein. Und Wagners Leitmotive erhalten ihren außermusikalischen Sinn aus dem komplexen Gesamtzusammenhang von szenisch-dramaturgischen, sprachlichen und musikalischen Strukturen, und im Sinngefüge dieses Ganzen funktionieren sie auch erst. (Die Rede vom Gesamtkunstwerk hat hier auch einen streng strukturellen Sinn.)

Drittens: Musikalische Grundzeichen, insbesondere solche Motive, die zur Kategorie der auf Wiedererkennung disponierten Gestalten gehören, zeigen einen Gestus auf deutlich abgehobene Weise, sie individualisieren ihn gewissermaßen. Die Bedeutung des umfassenderen Gebildes, in das sie eingehen, macht von dieser Identität Gebrauch. Alle Motive, die nicht unterhalb dieser Identifizierungsschwelle bleiben, haben damit etwas von der speziellen sprachlichen Kategorie der Eigennamen an sich. Das Gemeinsame von Motiven und Namen ist, daß beide immer auf das gleiche verweisen. Das muß erläutert werden. Alle Wörter und Wortverbindungen außer den Eigennamen können immer wechselnde Dinge und Eigenschaften der jeweils einschlägigen Klasse benennen. Der Ausdruck *ein Haus* zum Beispiel bezeichnet nicht immer dasselbe Haus, wenn er in einem Text wiederholt vorkommt. Nur Namen haben immer dasselbe Denotat, sooft sie auch wiederholt werden.⁶¹ Ebenso zeigt ein Motiv immer den gleichen individuellen Gestus. In diesem Sinn haben musikalische Zeichen mithin eher eine größere als eine geringere Konstanz als sprachliche. (Und in diesem Sinn wird in manchen Kompositionen ein Thema zum Individuum, zum „Helden“, dessen Schicksal, einer keineswegs unbegründeten hermeneutischen Redeweise zufolge, in der Komposition gezeigt wird.) Und dies ist der fundamentale Unterschied: Ein Name benennt ein Individuum, aber ein Motiv oder Thema zeigt einen Gestus, es benennt ihn nicht. Das Individuum, das ein Name benennt,

ist nicht die Bedeutung des Namens, sondern das, worauf die Bedeutung, die logische Form des Wortes, verweist. Der gezeigte Gestus aber ist die Bedeutung des Motivs, nicht das, worauf seine Bedeutung verweist. Und ganz entsprechend unterschiedlich funktionieren Namen und Motive im syntaktischen und semantischen Zusammenhang. Über ein benanntes Individuum wird Verschiedenes gesagt – *Hans wachte auf, Hans war deprimiert* usw. –, mit oder auch an einem Motiv (etwa bei seiner Verkürzung, Dehnung, Modulation, Umkehrung usw.) wird Verschiedenes gezeigt. Um den Vergleich simplifizierend auf die Spitze zu treiben: Der Ausdruck *Alexander* kann nicht sagen, daß Alexander traurig geworden ist, wohl aber kann die Moll-Variante eines Dur-Motivs eine solche Veränderung eines Gestus zeigen.

Zusammengefaßt: Die Art, in der sprachliche und musikalische Grundzeichen ihre Bedeutung erhalten und sie zur Bedeutung komplexer Zeichen zusammenfügen, ist der kardinale Unterschied in der Funktionsweise von Sprache und Musik. Auf ihm gründet sich der Unterschied zwischen Sagen und Zeigen.

18. Coda

Man kann die Überlegungen, die ich angestellt habe, eine Art vergleichender funktionaler Anatomie von Sprache und Musik nennen. Dabei ging es eher um die Kennzeichnung von Grenzen und Konturen als darum, den Inhalt der Konturen zu füllen. Das wiederum ergibt sich daraus, daß die Überlegungen in zweierlei Hinsicht abstrakt waren. Da das genaue Verständnis der dabei verfolgten Gesichtspunkte wichtig ist für den Sinn und die Konsequenzen der Überlegungen, will ich sie im Rückblick auf das Dargelegte noch einmal markieren.

Zum ersten Gesichtspunkt. Ich habe Musik und Sprache als Funktionsprinzipien zu erfassen versucht, nicht einzelne, vielfältig determinierte Phänomene beschrieben. Das ist nicht eine willkürliche Herauslösung eines beliebigen, wenn auch interessanten Moments. Es ist vielmehr die Heraushebung jeweils eines Determinationsgefüges, das in sich ebenso gerechtfertigt werden muß wie etwa die Gravitationsgesetze, die zusammen mit anderen Gesetzmäßigkeiten die Bahn fallender Blätter und Steine determinieren. Anders ausgedrückt: Auch wenn es überhaupt keine Erscheinung gibt, die allein durch die hier diskutierten Gesetzmäßigkeiten von Sprache bzw. Musik bestimmt wird, können diese Gesetzmäßigkeiten dennoch einen eigenständigen, empirisch zu ermittelnden Charakter haben.

In der Tat ist die Hervorbringung und Rezeption sprachlicher und musikalischer Äußerungen stets in mehr oder weniger vielfältige Zusammenhänge eingeordnet. Sie teilen dies mit anderen durch Strukturen und Regeln bedingten Verhaltensformen. Selbst eine ganz autonom geregelte Tätigkeit wie Schachspielen kann ganz verschiedene Funktionen haben: Vergnügen, Aus-

tragung von Rivalität, geistiges Training, Höflichkeitsbezeugung oder Zeitvertreib, und all dies auf der Grundlage der Geltung und Kenntnis der Schachspielregeln. Da Musik und Sprache auf besonders stark automatisierten, weitgehend unwillkürlich wirksamen Kenntnissen beruhen, versteht es sich um so mehr, daß sie unter entsprechenden Bedingungen untereinander und mit anderen Formen des inneren und äußeren Verhaltens in komplexe Verbindungen eingehen. Wie aber Vergnügen am Schachspiel nur durch dessen Regeln hindurch zustandekommt, so auch die verschiedenartigen Wirkungen von Musik und Sprache nur durch deren formale Struktur und die ihr zugeordnete Bedeutung hindurch. Das gilt für die Oper nicht weniger als für Tanz, Ritual oder andere Gebrauchsfunktionen der Musik, und natürlich für die praktisch unübersehbare Fülle verschiedener Funktionen der Sprache. Es wäre aber ebenso sinnlos, diese weitergehenden Funktionen in die Musik und die Sprache und ihre Bedeutungsstrukturen hineinzu projizieren, wie es sinnlos wäre, Vergnügen oder Geltungsdrang in die Regeln einer Schachpartie zu verlagern oder die vom Wind gekrümmte Bahn eines fallenden Blattes in das Fallgesetz.

Das eben Gesagte betrifft nicht erst die Einordnung sprachlicher und musikalischer Signale in die verschiedenartigen Handlungs- und Situationszusammenhänge, es läßt sich oft schon am akustischen Signalereignis selbst dingfest machen, ja in speziellen Fällen sogar in den sprachlichen und musikalischen Strukturen. Die Poesie ist insgesamt ein Beispiel für diesen Fall. In einer Gedichtstrophe wie

*Über allen Gipfeln ist Ruh,
in allen Wipfeln spürest du
kaum einen Hauch*

überlagern sich zwei Strukturen: die sprachliche, die den Lautformen eine syntaktische Verknüpfung und durch diese als Ganzes eine Satzbedeutung zuordnet, und ein metrisches Muster mit Versgliederung und Reimbindung, das einen (wenn auch an sich sehr wenig definiten) Gestus fixiert. Beide Strukturen bestimmen sich gegenseitig, nicht nur in ihrer formalen, auch in ihrer Bedeutungsseite. In der Vertonung entsteht als weitere Anreicherung ein neues, komplexeres Signal, das zusätzlich die musikalische Struktur und deren wesentlich definitere gestische Bedeutung trägt. Damit kommen zusätzliche Determinationen ins Spiel.

Schließlich ist bei allen sprachlichen und musikalischen Gebilden, insbesondere bei denen, die als Kunst verstanden werden, aber nicht nur bei ihnen, die Einordnung der Bedeutung in weitgreifende Sinnzusammenhänge der Hintergrund ihrer konkreten Wirkung. Ich habe dies den zweiten Schritt des Verstehens genannt (was keine zeitliche Reihenfolge der Schritte unterstellt). Er variiert je nach dem Kenntnishintergrund des einzelnen Senders und Empfängers, er kann Lernvorgänge einschließen oder erforderlich

machen. In diesem Schritt ist die Tatsache begründet, daß Musik kognitive Prozesse vermittelt, und zwar Musik gerade auch als gewissermaßen reines Funktionsprinzip genommen, also im Spezialfall der sogenannten absoluten Musik, und keineswegs erst im Zusammenhang mit manifestem oder unterstelltem Text oder gegenständlichem Programm. Propositionen formen, Aussagen machen ist eine spezielle Form kognitiver Leistung, für den Menschen eine zentrale, aber mitnichten die einzige. Die Musik kann, ohne Aussagen zu bilden, auf dem Weg über die Bezüge in ihrer gestischen Form, geistige Prozesse von eminenter Komplexität auslösen, so wie die Sprache auf dem Weg über begriffliche Strukturen emotionale Prozesse von großer Intensität mobilisieren kann. Das Verständnis der Zeichensysteme in ihrer jeweils primären Funktionsweise ist Voraussetzung für das Erfassen gerade auch dieser Möglichkeiten.

Nun zum zweiten Gesichtspunkt. Ich habe die so herausgelösten Funktionsprinzipien von Sprache und Musik zwar an gelegentlichen Beispielen illustriert, im Grundzug jedoch generalisierend und allgemein erörtert. Dabei waren verschiedene Stufen von Allgemeinheit im Spiel, die ich nicht immer als solche abgehoben habe. Ich will sie zusammenfassend ordnen und die Art kenntlich machen, in der sie jeweils für sich systematisch aufzubauen wären.

Die weitestgehende Allgemeinheit ist eine rein begriffliche Analyse der Konzepte Kommunikation und Zeichensystem. Sie hat rein logischen Charakter und grenzt nur die Bedingungen der Möglichkeit empirischer Unterscheidungen und Bezüge ab.⁶² Der so zu bestimmende, keineswegs simple Rahmen, von dem vor allem die Abschnitte vier und fünf, aber auch viele andere Passagen handeln, bildet den gemeinsamen Hintergrund für Sprache und Musik.

Eine zweite Stufe bildet innerhalb dieses Rahmens die wiederum rein begriffliche Bestimmung verschiedener Typen von Zeichensystemen. Dies kann im Hinblick auf verschiedene Bedingungen geschehen, etwa das Medium der Zeichenformen und die damit verbundenen Konsequenzen, den Bereich der Bedeutungen, die Formen der Codierung betreffend. Insbesondere ist hier ein allgemeiner Begriff von Sprache zu fixieren, der zunächst nicht auf natürliche Sprachen eingegrenzt ist, sondern auch künstliche und formale Sprachen, ja sogar Computersprachen einbezieht. Ich habe diese Stufe nicht gesondert thematisiert, doch gehört hierher die Unterscheidung der Codierungsformen und der Funktionen des Sagens und Zeigens: Sprachen sind Zeichensysteme, die etwas sagen, Musik gehört mit bildnerischen Darstellungen zusammen zu den Zeichensystemen, die etwas zeigen.

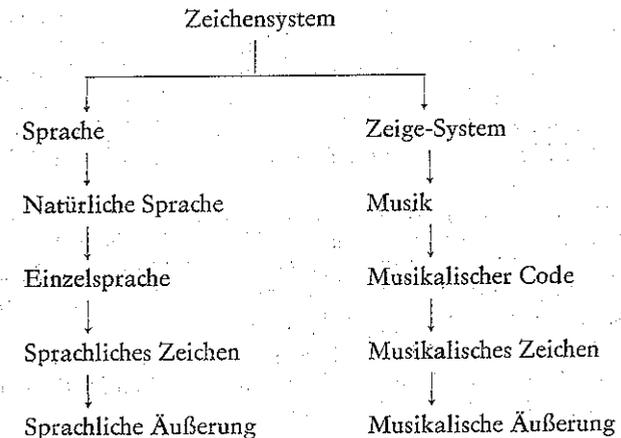
Die dritte Stufe stellt die universellen, aber empirisch bedingten Gegebenheiten der natürlichen Sprache und der Musik in Rechnung. Welchen Rahmen die Merkmalauswahl zur Verfügung hat, welche Formen syntaktische Strukturen und Operationen haben können: das ist nur durch die Analyse empirischer Gegebenheiten zu begründen. Im wesentlichen auf dieser Stufe

haben sich die – durchaus unvollständigen – Vergleiche des systematischen Aufbaus von Musik und Sprache in den Abschnitten 8 bis 17 bewegt, auch wenn sie durch diese Stufe hindurch oder zu ihrer Fundierung begriffliche Analysen einschlossen.

Die nächste Stufe, die ich in keinem Fall ernsthaft betreten habe, bestünde in der Charakterisierung der Elemente und Regeln einer einzelnen natürlichen Sprache, eines speziellen musikalischen Codes. Auch dies sind noch immer abstrakte Gebilde, sie treten nur vermittelt in Erscheinung durch die Strukturen, die sie determinieren. Im Unterschied zur dritten Stufe bewegt sich diese jedoch im Bereich gesellschaftlich-kulturell bedingter, also historischer Erscheinungen: Sprachen und musikalische Codes ändern sich mit den Gemeinschaften, die sie benutzen, und entsprechend den Bedürfnissen und Zielen, die sie verfolgen.

Erst die fünfte Stufe schließlich betrifft die Struktur der einzelnen sprachlichen und musikalischen Gebilde, mit denen Philologen und Musikhistoriker primär befaßt sind. Ich habe diese Stufe nur anekdotisch und zum Zweck illustrierender Beispiele berührt – und auch das ausdrücklich im Hinblick auf deren abstrakte, nämlich die Struktur und die Funktionsprinzipien betreffende Seite.

Ich fasse die Stufen noch einmal in einem Schema zusammen, das die Abhängigkeit zwischen den konzeptuellen und empirischen Bedingungsgefügen zeigt:



Zu jeder dieser Stufen gehört der jeweilige empirische Gegenstand und dessen Spiegelung in der theoretischen Analyse.

Mit dieser Kennzeichnung der Abstraktionsstufen ist nun auch der Abstand und der Weg deutlich, der zwischen den Klärungsversuchen auf den oberen drei Stufen und der Analyse konkreter Äußerungen liegt. Für die

Linguistik ist dieser Weg ein klar vorgezeichnetes und in aufschlußreichen Fragmenten auch ausgeführtes Programm. Ein entsprechendes Programm für die Musikwissenschaft ins Auge zu fassen, bedeutet nicht nur, viele verschiedenartige Erkenntnisse und wohlbegründete Befunde in einen systematischen Zusammenhang zu bringen. Es heißt vor allem, eine entsprechende Blickrichtung zu gewinnen. Damit ist gemeint, daß die Analyse einzelner Strukturen nicht lediglich auf genaue Beschreibung, sondern auf die generalisierende Ermittlung des zugrunde liegenden Codes, und diese auf die Charakterisierung musikalischer Universalien hinauslaufen muß. Das wiederum verlangt einerseits die Bereitstellung geeigneter theoretischer Instrumentarien, vor allem mathematisch struktureller Systeme, zum anderen die Entwicklung empirischer Unterscheidungskriterien und von Verfahren zu ihrer Prüfung. Die Musikwissenschaft würde dadurch, wie vor ihr die Linguistik, den Weg zur systematischen Erklärung ihres Gegenstandes einschlagen, wobei „Erklären“ ohne Verlust der Spezifik der Disziplin den gleichen Sinn annimmt wie in der Naturwissenschaft: Herleitung der Einzelercheinungen aus allgemeinen Gesetzmäßigkeiten.

Diese Erwägung schließlich verweist auf den Sinn, den die oft abgelegenen und in ihrem Mangel an konkreter Substanz gelegentlich unbefriedigenden Überlegungen in ihrer Gesamtheit haben könnten: Um konkrete, an den interessanten Einzelfall gerichtete Fragen stellen zu können, bedarf es eines Orientierungsrahmens. Er ersetzt weder die konkreten Fragen noch gar die Antworten. Ihn bewußt zu machen, ihm nicht nur implizit zu folgen, erzeugt jedoch Fragen und Zusammenhänge, die sonst nicht entstehen. Traditionelle, angestammte Fragen und Resultate werden dabei nicht selten umzuordnen oder anders zu formulieren, aber nicht aufzugeben, sondern zu integrieren sein. Zumindest die folgenden Fragenkomplexe, für die linguistische Entwicklungen nur zum Teil abwandeltbare Modelle liefern, sind in diesem Sinn formulierbar:

(a) Voraussetzung für jede systematische Analyse ist ein Repertoire empirisch-diagnostischer Konzepte, das dem Zeichencharakter des Gegenstands angemessen ist. Akzeptabilität, syntaktische Wohlgeformtheit, strukturelle und semantische Mehrdeutigkeit sind Beispiele für solche diagnostischen Konzepte, die sich in der Linguistik als unentbehrlich erwiesen haben. Sie dienen auf oft diffizile Weise zur Klassifizierung der konkreten Erscheinungen und damit zur empirischen Rechtfertigung erklärender theoretischer Annahmen. Ein Beispiel zur Andeutung:

(1) *Der Termin, den Hans hoffte einhalten zu können, ist verstrichen.*

(2) *Der Termin, den Hans hoffte, daß er einhalten könnte, ist verstrichen.*

(1) ist ein korrekter Satz der deutschen Sprache, (2) ist defekt, er ist „ungrammatisch“. Das ist eine empirische Feststellung, die aber ein Konzept von „Ungrammatikalität“ voraussetzt. Diese Feststellung besagt an sich noch

nicht viel, sie wird aber interessanter und erlaubt wichtige Schlüsse für die Regeln der deutschen Syntax und für die Prinzipien der Syntax generell, wenn man sie in systematische Zusammenhänge stellt, etwa die Beobachtung, daß sowohl (3) wie (4) grammatisch sind:

(3) *Hans hoffte, den Termin einhalten zu können.*

(4) *Hans hoffte, daß er den Termin einhalten könnte.*

Das läuft auf systematisches Experimentieren hinaus. Die Frage ist hier: Was passiert, wenn man einen daß-Satz unter bestimmten Bedingungen mit einer Infinitiv-Konstruktion vertauscht. Solche Fragestellungen verwandeln eine rein deskriptive Vorgehensweise in prädiktive Theoriebildung. Auch für die Musik können (und müssen) solche Fragestellungen verfolgt werden. Denn nur so läßt sich auf nicht-spekulative Weise klären, was z. B. die Mindestmerkmale eines Motivs sind, welchen Verknüpfungsoperationen es unterliegen kann, welchen Charakter welche Veränderung hat.

(b) Außer solchen Beobachtungskonzepten und -techniken müssen geeignete Systeme theoretischer Begriffe und Strukturen entwickelt werden. Der formale Aufbau der Strukturebenen, die ich diskutiert habe, ihre Einheiten und Relationen müssen systematisch bestimmt werden. Um empirisch prüfbare Kriterien im Sinn von (a) zur Unterscheidung von Motiven, zur Identifizierung gestischer Gliederungen ausdenken und anwenden zu können, muß man über einen hinreichend präzisierten Begriff von Motiv und Gestus verfügen. Das ist nicht nur eine Frage der inhaltlichen Überlegung, sondern auch der formalen Kennzeichnung. Mit diesen Mitteln können dann Analysen einzelner Gebilde und die Beschreibung musikalischer Codes aufgebaut werden. Hier wäre auch rational zu präzisieren, was die vieldeutige Rede vom „musikalischen Material“ besagt.

(c) Zugleich mit der Beschreibung der Regeln und Einheiten konkreter, historisch bedingter Codes, genauer: durch sie hindurch, als Grundlage und Korrektiv, ist die Charakterisierung musikalischer Universalien in Angriff zu nehmen. Was sind mögliche Verknüpfungs- und Transformationsoperationen, welchen (psychologisch bedingten) Beschränkungen unterliegen sie, welche Merkmale und Beziehungen können überhaupt unterschieden und genutzt werden? Das sind Fragen, die in diesen Zusammenhang gehören.

Die Verfolgung der Probleme unter (a) bis (c) würde nicht nur die hier angestellten Überlegungen präzisieren oder korrigieren, sie würde auch schrittweise die Lücke zwischen den „höheren“ Stufen und der empirischen Grundlage schließen, die dem Schema der Abstraktionsebenen abzulesen ist. Manches, was unter diese Fragen fällt, ist bereits Thema der systematischen Musikwissenschaft; vieles ist neu zu entwickeln. An Hand eines gesicherten Verständnisses der so umschriebenen strukturellen und funktionalen Aspekte der Musik sind auch die folgenden Probleme, die weit stärker das Interesse der Musikwissenschaft absorbiert haben, neu zu durchdenken.

(d) Die Verwendung von Sprache und Musik – ich habe das ausführlich begründet – involviert mehr als die Aktivierung der Codekenntnisse. Wie aber das weitergehende Verständnis, das ich nur vage umschrieben habe, strukturiert ist, wie es zustandekommt, läßt sich nur angemessen verstehen, wenn man den Anteil und die Funktionsweise der Zeichenkenntnisse dabei berücksichtigt. Die Zusammenhänge werden hier fraglos um vieles verwickelter und unübersichtlicher als im Rahmen der selbst schon sehr komplexen Zeichenstrukturen. Dennoch bilden sie kein amorphes Gebiet, in das man das exakt nicht Faßbare abschieben kann und über das nur einfühlsame Spekulationen möglich sind. Tatsächlich belegt eine Vielzahl konkreter Werkanalysen durchaus das Gegenteil. Zu gewinnen ist hier jedoch eine klarere Bestimmung des Zusammenwirkens struktureller Aspekte der Musik und der verschiedenen Formen von Hintergrundkenntnis, Lernprozeß, Situations-einbettung.

(e) Auch die Frage, die in vieler Hinsicht der Kernpunkt geläufiger Musik-auffassung ist, die nach der ästhetischen Funktion, kann erst auf der Grundlage angemessener Erkenntnisse über den strukturellen Aufbau musikalischer Zeichen präziser gestellt werden. Ich habe diese Frage überhaupt nicht thematisiert, und es bleibt zu zeigen, daß sie einen rationalen, nicht trivialen Charakter annimmt, wenn sie auf strukturell wohlverstandene Gebilde bezogen wird, statt vorab auf unanalytierte Komplexe.

(f) Schließlich müssen alle unter (a) bis (e) umschriebenen Faktoren zusammenwirken, um zumindest Elemente überprüfbarer Argumentation in das Gebiet hineinzutragen, das zuerst und zuletzt eine Kunst ist: die Wertung der Qualität von Musik, ihrer individuellen und gesellschaftlichen Bedeutsamkeit.

Was es heißt, daß der Mensch Sprache *und* Musik besitzt, sie im Rahmen seiner phylogenetisch erworbenen Anlagen historisch entwickelt und sich in beiden auf verschiedene Weise individuell verwirklicht, diese Frage etwas genauer zu formulieren: das war alles in allem die Absicht des vorliegenden Versuchs.

Anmerkungen

* Der vorliegende Beitrag ist die überarbeitete Fassung zweier Rundfunkvorträge, die ich im Mai 1977 im Programm von Radio DDR II gehalten habe. Sie wurden von Dieter Boeck in einem nicht nur formalen Sinn angeregt und betreut. Der Inhalt meiner Überlegungen ist wesentlich beeinflusst durch das, was ich in Gesprächen und Diskussionen mit zahlreichen Kollegen und Freunden gelernt habe. Neben vielen anderen habe ich dafür vor allem Professor Dr. Friedhart Klux, Professor Dr. Georg Knepler, Dr. Hans Geißler, Dr. Hans Grüß, Dr. Eberhardt Klemm, Dr. Ewald Lang und Dr. Doris Stockmann zu danken. Nicht nur im Sinn der üblichen Formel muß ich hinzufügen, daß Unzulänglichkeiten selbstverständlich von mir zu verantworten sind. Denn

was ich von meinen Gesprächspartnern gelernt habe, ist keineswegs immer mit deren Auffassung identisch. Dies gilt nicht für meinen Freund Klaus Baumgärtner, der mir in nur äußerlich unterbrochenen Gesprächen verständlich gemacht hat, worauf ich hinauswollte und was dafür nötig wäre. Er ist deshalb eigentlich an jedem Gedanken beteiligt, vor allem aber auch am Bewußtsein von dem, was fehlt. Anzumerken ist noch, daß ich den ursprünglichen Charakter der Vorträge nicht nur in der Darlegungsform und im weitgehenden Verzicht auf wissenschaftlichen Apparat beibehalten habe, sondern auch in dem Versuch, so weit wie möglich zu vereinfachen, wichtige Nebenaspekte beiseite zu lassen, um die Hauptlinien deutlich zu machen, auf die es mir jeweils ankommt. Wie weit das ohne Schaden möglich war, bleibt abzuwarten.

¹ Die damit unterstellte Zweiteilung in Sprache und andere Zeichensysteme ist natürlich eine Vergrößerung. Sie unterschlägt die mannigfaltigen Unterschiede und Verwandtschaften, nach denen Zeichensysteme im Hinblick auf Entstehung, Funktion, verwendete Mittel und Strukturmerkmale zu gruppieren sind. Die sogenannten formalen Sprachen der Logik und Mathematik etwa sind auf ganz andere Weise verwandt mit der natürlichen Sprache, als die Zeichensprachen der Taubstummen. Beide aber teilen wesentliche Charakteristika mit ihr, die auf Systeme der Verkehrszeichen oder der Nationalflaggen z. B. nicht zutreffen. Weitere Unterschiede kommen ins Spiel, wenn diese vom Menschen hervorgebrachten Zeichensysteme denen der Tierkommunikation gegenübergestellt werden. Da eine differenzierte Typologie der Zeichensysteme hier aber nicht vorauszusetzen ist – für einen neueren Versuch dazu vgl. etwa Eco (1976) –, halte ich mich weitgehend an die Festlegung, den Terminus Sprache für die verbalen Zeichensysteme des Menschen zu reservieren.

² Die Einschränkung auf bestimmte Grenzen der Genauigkeit ist notwendig, weil z. B. unser Unterscheidungsvermögen für Gesichter oder Stimmen größer ist als die Charakterisierungsschärfe sprachlicher Mittel. Ein Photo oder ein Phantombild identifiziert deshalb eine Person genauer als die beste Personenbeschreibung. Diese Grenze ist prinzipieller Natur, sie ist im begrifflich-abstraktiven Charakter begründet.

³ Für eine kurze Charakterisierung der systematischen Literaturwissenschaft vgl. Bierwisch (1965), für größere Überblicke Ihwe (1972), Marcus (1973) und Freeman (1970). Daß in diesem Gebiet sowohl methodologische Anregungen aus der Linguistik wie inhaltliche Bezüge zur Sprachstruktur eine wesentliche Rolle spielen, liegt auf der Hand. Nichtsdestoweniger handelt es sich der Sache nach um einen Aspekt der Literaturwissenschaft.

⁴ Aufschlußreich ist auch ein wissenschaftsgeschichtlicher Blick auf dieses Schema: In allen drei Bereichen hat im 19. Jahrhundert die historische über die systematische Betrachtungsweise dominiert. Unter Anknüpfung an ältere Traditionen haben sich in der ersten Hälfte des 20. Jahrhunderts die systematischen Disziplinen schrittweise emanzipiert und ihre eigene Themenstellung entfaltet. Es ist übrigens nicht schwer, wenn auch nicht trivial, das Schema auf den visuellen Bereich zu erweitern und die bildenden Künste in ihrem systematischen und historischen Aspekt einzubeziehen.

⁵ Um das einzusehen, kann man sich die (in dieser Form fiktive) Situation vorstellen, daß eine unbekannte Sprache aus Dokumenten rekonstruiert wird, die in einer bekannten Sprache abgefaßt sind und interlinear Schriftzeichen enthalten, deren Lautwert unbekannt ist, die aber als Übersetzung der bekannten Sprache betrachtet werden müssen. Von dieser fiktiven Sprache wäre dann nur ihre geschriebene Form bekannt. Dennoch könnte man neue, einsprachige Texte in dieser Sprache abfassen und verstehen.

⁶ Ich deute wenigstens drei Punkte einer solchen Analyse an. Erstens: Ihrer Entstehung nach sind alle Schriften, soweit man sehen kann, zunächst Bilderschriften, d. h.

schematisierte Darstellungen begrifflicher Inhalte. Sie entwickeln sich erst schrittweise zu Wort-, Silben- und schließlich Buchstabenschriften, die strikt die Lautgestalt zum Bezug haben. Die musikalische Notation beginnt dagegen stets mit abkürzenden Hinweisen auf Lautmerkmale, meist in Form von diakritischen Zeichen, die den zu singenden Texten hinzugefügt werden. Die Notation beginnt also in dem Bereich, in dem sich auch ihre ganze systematische Ausgestaltung vollzieht: im Bereich der Lautmerkmale. Sie geht nicht von einer Bezugsebene in eine andere über. Zweitens: Funktional sind Schrift und Notation Zeichensysteme, deren Bedeutung auf bestimmte Aspekte der primären Systeme verweist, die durch sie sekundär repräsentiert werden. In der Musik ist dieser Aspekt fast ausschließlich die Lautstruktur. (Ich sage fast ausschließlich, weil der Status von Angaben wie *con spirito*, *cantabile* oder Mahlers *wie ein schwerer Kondukt* nicht eindeutig ist.) In der Sprache hingegen sind, wie wir gesehen haben, Bezüge zu verschiedenen Aspekten möglich, die Lautstruktur setzt sich erst allmählich als dominierender, weil optimaler Bezug, durch. Und drittens: Strukturell ist die musikalische Notation von Anfang an weitgehend durch Analogien zwischen Schriftzeichen und Lautmerkmalen motiviert, sie ist, um einen später zu erklärenden Begriff zu verwenden, ein vorwiegend ikonisches Zeichensystem. Die Schrift hingegen ist in ihren frühen Phasen durch Ähnlichkeiten mit dem Inhalt der sprachlichen Zeichen motiviert, später ist sie ein überwiegend unmotiviertes, arbiträres Zeichensystem: Die Buchstaben E, I und F z. B. zeigen in ihrer Form nichts von den lautlichen Unterschieden und Ähnlichkeiten der phonetischen Segmente, die sie repräsentieren. Dagegen wird durch die Platzierung der Noten die relative Tonhöhe auf streng systematische Weise repräsentiert. Diese Andeutungen, die durch einen Blick in einschlägige Standardwerke – etwa Jensen (1958) und Apel (1962) – beliebig angereichert werden können, müssen genügen, um die Schlussfolgerung zu belegen: Die Dominanz der Lautstruktur schlägt sich in der Musik im systematischen Aufbau der Notationssysteme nieder, während die in bestimmten Grenzen arbiträre Lautstruktur der Sprache sich in dem zu ihr relativ arbiträren Charakter der Schriftsysteme ausdrückt. Kurz, für die schriftliche Fixierung steht in der Musik die akustische Realisierung im Vordergrund, in der Sprache dagegen die Diskrimination von Bedeutungen.

Der einzelne Spieler ist einerseits Teil eines kollektiven Senders, andererseits zugleich vollständiger Empfänger der kollektiv produzierten Mitteilung. (Und dies unabhängig davon, ob an der Produktion nicht beteiligte Hörer dabei sind oder nicht.) Es ist klar, daß dabei zahlreiche interessante Zusatzprozesse auftreten, deren Betrachtung hier aber beiseite gelassen werden muß. Man sieht jedoch, inwiefern sich das Gruppenmusizieren vom früher betrachteten reinen Monologisieren unterscheidet.

⁸ Vollständigkeitshalber erwähne ich noch eine dritte Komplizierung in der Bestimmung der Kommunikationspartner. Sie ergibt sich aus der Möglichkeit zeitlicher und räumlicher Trennung von Sender und Empfänger. Sie hat sich mit den technischen Medien und Aufzeichnungsmöglichkeiten verstärkt, reicht aber weiter zurück, mindestens bis zur Entstehung schriftlicher Fixierung. Die Zusatzbedingungen, die mit dem Lesen eines Briefes oder der Produktion und dem Abhören einer Schallplatte verbunden sind, können hier aber mit dem gleichen Argument ausgespart werden, das ich soeben in bezug auf den komplexen Erzeugungsprozeß der Mitteilung angeführt habe.

⁹ Ich habe die wesentlichen Merkmale eines gelungenen Kommunikationsakts beschrieben. Sie gelten auch, wenn der Sender das Signal in einem mehrstufigen, kollektiven Prozeß produziert, und sie erfassen auch Fälle, in denen der Sender nicht einen bestimmten, im voraus identifizierten Empfänger im Sinn hat, nur muß der Empfänger bemerken, daß der Sender die fragliche Mitteilung beabsichtigt hat. Man kann sich nun

klar machen, daß alle Arten von defekten Kommunikationsakten, also Mißverständnisse, Täuschungen (absichtliche und unabsichtliche) oder unvollständiges Erfassen der Mitteilung, nur auf dem Hintergrund dieser den gelungenen Akt bestimmenden Merkmale möglich sind. Das gilt auch für den Fall, daß jemand ein Signal aufnimmt, für das er nicht als (potentieller) Hörer bestimmt war, wo also z. B. jemand zufällig ein Gespräch anhört oder es absichtlich belauscht. Das im einzelnen nachzuprüfen überlasse ich dem Leser und gehe im folgenden vom normalen, gelungenen Kommunikationsakt als Grundlage aus.

¹⁰ Häufig, aber nicht notwendigerweise sind es physikalische Eigenschaften des Signals und deren natürliche Wirkung auf den Organismus, so z. B. der Lärm in Beispiel (E). Ähnlich, wenn jemand, die „Sprache der Blumen“ benutzend, Rosen schickt, die auch ohne die verschlüsselte Liebeserklärung willkommen sind. Ein Gegenbeispiel wäre ein Mann, der in einer Synagoge die Kopfbedeckung abnimmt in der vom Rabbiner erkannten Absicht, diesen zu brüskieren. Die „natürliche“ Wirkung $W(n)$ kommt hier nicht durch den unbedeckten Kopf, sondern durch die Vorschrift zustande, in der Synagoge eine Kopfbedeckung zu tragen.

¹¹ Man wird einwenden, daß die Mitteilung sehr unterschiedlich ausfällt, je nachdem, ob man den Satz *Hier ist gesperrt* brüllt oder ruhig ausspricht. Das ist nicht zu bezweifeln, aber dieser Unterschied kommt gerade nicht durch den sprachlichen Anteil der Mitteilung zustande. (Es heißt nicht zufällig: Der Ton macht die Musik.) Sprachliche Äußerungen realisieren eben meist nicht nur sprachliche Mittel. Wir werden das sogleich deutlicher sehen.

¹² In der oben erwähnten Arbeit von Grice (1957) wird der Versuch gemacht, die Bedeutung von Zeichen durch ihre Zurückführung auf die mit ihnen verbundene Mitteilung zu erklären. Ziff (1967) hat gezeigt, daß das grundsätzlich nicht möglich ist, daß Bedeutung ein unabhängig zu klärendes Konzept ist.

¹³ Das Verstehen der Mitteilung ist mehr als das Verstehen der Bedeutung der vorgegebenen Zeichen. Sein Ergebnis ist spezieller, und weil es von mehr Faktoren abhängt als das Verstehen der Bedeutung, können in ihm auch leichter Differenzen auftreten zwischen dem, was der Sender zu bewirken beabsichtigt und was im Hörer als Wirkung entsteht. (Streng genommen sind solche Differenzen Mißverständnisse, d. h. Fälle von nicht vollständig gelungener Kommunikation.) Nach der anderen Seite hin ist das Verstehen der über die Zeichenbedeutung hinausgehenden Mitteilung zu unterscheiden von sekundären Wirkungen, die zwar durch einen Kommunikationsakt ausgelöst werden, aber nicht Teil von $W(k)$ sind. Der Schreck, den die Frage *Wo bast du dein Portemonnaie?* verursacht bei einem, der daraufhin den Verlust bemerkt, gehört nicht zur Mitteilung, die der Frager macht. Zweifelloser ist die Grenze hier gelegentlich schwer zu bestimmen, was jedoch nichts gegen die notwendige Differenzierung besagt.

¹⁴ Die Verbindung der gleichen musikalischen Gestalt mit verschiedenen sprachlichen Äußerungen – klassisches Beispiel: der Choral „O Haupt voll Blut und Wunden“ in der Matthäuspassion und „Wie soll ich dich empfangen“ im Weihnachtsoratorium von Bach – gehört, wie letzten Endes überhaupt das Zusammenwirken von Musik und Sprache, auch in diesen Problembereich. Ich komme darauf noch genauer zurück.

¹⁵ Es ist weder möglich noch notwendig, in unserem Zusammenhang auf die kaum überschaubare Literatur zur Zeichenproblematik in auch nur annähernd repräsentativer Weise Bezug zu nehmen. Wegen der verschiedenartigen Zusammenhänge, in denen von Zeichen die Rede ist, ergäbe sich dabei ein nicht nur terminologisch sehr heterogenes Bild. Ich werde deshalb so wenig Voraussetzungen wie möglich machen und versuchen, die Festlegungen so zu begründen, daß sie für sich genommen einsichtig sind. Einen mit

Vorsicht heran zu ziehenden Hintergrund bilden Ansatzpunkte, die Saussure (1916) formuliert und Hjelmselv (1953) präzisiert hat.

¹⁶ Jenseits terminologischer Differenzen gibt es eine immer wieder auflebende Kontroverse darüber, ob Zeichen unilaterale oder bilaterale Gebilde sind. Dieser Streit reduziert sich auf die Entscheidung, ob man die Bedeutung als einen Teil des Zeichens oder als eine Eigenschaft (mathematisch gesprochen: als eine Funktion) seiner Form ansehen will. Im ersten Fall ist die Beziehung zwischen F und B eine Relation, ein Zeichen $Z = (F, B)$ ein geordnetes Paar, im zweiten Fall ist die Beziehung zwischen F und B eine Abbildung. Die Wahl der Redeweise ist beliebig, ich folge der ersten.

¹⁷ Wegen dieser Isomorphie der Struktur von Form und Bedeutung bei musikalischen Zeichen hat Hjelmselv (1953, S. 113) nahegelegt, die Musik als ein System zu betrachten, durch das die Formen direkt auf Mitteilungen bezogen werden, ohne die Vermittlung durch eine Bedeutungsstruktur, so daß sie, in seiner Terminologie, nicht als Zeichen-, sondern als Symbolsystem gilt. Ob wir von zwei isomorphen Strukturen mit einer identischen Abbildung zwischen ihnen oder von einer einzigen Struktur sprechen, ist formal gesehen ebenso beliebig wie die Entscheidung über die Uni- oder Bilateralität von Zeichen (vgl. Anmerkung 16).

¹⁸ Interindividuelle Unterschiede mögen für die musikalische Disposition größer sein als für die sprachliche, bleiben aber vermutlich in nicht pathologischen Fällen in einem vorgegebenen Rahmen. „Unmusikalität“ ist zweifellos weitgehend ein Produkt bestimmter soziokultureller Bedingungen, die die Entwicklung der natürlichen Disposition hemmen, nicht eine Folge des Fehlens dieser Disposition.

¹⁹ Einen Überblick über die Probleme der Entstehung von Sprache und Sprechen gibt der Sammelband von Harnad, Steklis und Lancaster (1976). Einen bemerkenswerten Versuch, die Entstehung musikalischer Verfahren phylogenetisch einzuordnen, unternimmt Knepler (1977, S. 70–177).

²⁰ Zur Besonderheit der perceptiven Verarbeitung von Sprachlauten vgl. Liberman et al. (1967). Metzke und Stejngart haben genetisch fixierte Reaktionsdispositionen für Sprachlaute bei Kleinstkindern bereits vor dem Beginn des Spracherwerbs konstatiert. Die Eigenständigkeit der sprachlichen Lauterzeugung wird am drastischsten deutlich durch ihre selektive Störung in Fällen von motorischer Aphasie. – Die damit belegte angeborene Disposition zur Lautsprache macht noch einmal deutlich, daß die in Abschnitt 3 erörterte Möglichkeit von nicht gesprochener Sprache nur als künstliches Epiphänomen vorgestellt werden darf.

²¹ Die Mechanismen kognitiver Strukturen und Prozesse als solcher können dabei im Prinzip vorausgesetzt werden. Einen Überblick über einschlägige Forschungsergebnisse und Modellvorstellungen gibt Klix (1976), den Zusammenhang konzeptueller Strukturen mit ihrer sprachlichen Codierung behandeln besonders ausführlich Miller und Johnson-Laird (1976).

²² Mit dieser Auffassung unterscheide ich mich deutlich von der z. B. von Knepler (1977) vertretenen These, daß Musik begriffliche Bedeutungen haben und im sprachlichen Sinn „semantisiert“ werden kann, und zwar nicht nur durch die eben erwähnte Bindung an sprachliche Texte. Mir scheint, daß dieser Gedanke jedoch auf Undeutlichkeiten beruht, die ich durch die klarere Unterscheidung von Bedeutung und Mitteilung zu beheben versucht habe. Was Knepler (und andere) an triftigen Sachbeobachtungen herausgearbeitet haben, läßt sich, wie ich deutlich machen möchte, auf andere Weise angemessener erfassen.

²³ Der Charakter solcher Erzeugungssysteme ist für die Sprache ausgiebig studiert worden, insbesondere im Gefolge der entscheidenden Arbeiten von Chomsky (1957,

1965). Die Übertragung der dabei ermittelten Grundprinzipien auf die Musik ist keineswegs trivial und wird uns noch beschäftigen. Daß Versuche in dieser Richtung möglich und sinnvoll sind, zeigen Ansätze wie die von Lindblom und Sundberg (1970), die Erzeugungsregeln für eine Klasse schwedischer Kinderlieder angeben.

²⁴ Einen instruktiven Überblick über den gesamten Bereich der theoretischen Linguistik gibt Lyons (1968), eine kürzere Orientierung ermöglicht Bierwisch (1966).

²⁵ Vgl. z. B. Halle (1970) für die Grundzüge einer Theorie der Metrik und ihrer Beziehung zur Sprachstruktur.

²⁶ Es ist möglich, die Überlegung in dieser Richtung noch einen Schritt weiterzutreiben und auch die Dauer als eine Merkmalsdimension der Segmente anzusehen (wie das in der Tat die Notation der Tondauer nahelegt). Die Zeitmuster ergeben sich dann einfach aus der Verknüpfung der Segmente und bilden keine eigene Ebene. Eine solche Auffassung ist formal möglich und wegen des abhängigen Charakters der Zeitmuster in der Sprache dort auch angemessen. In der Musik spielt die Zeitorganisation jedoch eine so fundamentale Rolle, daß ihre Zurückführung auf Merkmale von Segmenten kein zutreffendes Bild des Strukturprinzips gibt. Aus diesem Grund habe ich die Zeitmuster als eigene Ebene behandelt. Es ist aber nützlich, sich die Verwandtschaft von Strukturprinzipien vor Augen zu führen, die auf verschiedenen Ebenen wiederkehren.

²⁷ Bei kontinuierlicher Abstufung ist die Menge potentieller musikalischer Segmente theoretisch unendlich. De facto wird sie zunächst durch die Diskriminationsfähigkeit des Wahrnehmungsapparats begrenzt, dann durch diskrete Skalenbildung weiter eingeschränkt.

²⁸ Technisch erscheint dieser Aspekt als „Instrumentation“. Beispiele wie Weberns Fassung des Badischen *Ricercar* aus dem *Musikalischen Opfer* zeigen, daß dabei nicht nur Tonfolgen zusammengefaßt, sondern auch punktuelle Kontraste gebildet werden können. Die Organisation reiner Geräuschfolgen schließlich findet sich nicht nur in frühen Kulturen, sondern auch in komplizierten Kompositionen, für die Varèse ein sehr charakteristisches Beispiel ist.

²⁹ Es ergibt sich dabei die Frage, ob Akkorde als mehrere simultan verknüpfte Segmente oder aber als ein einziges, gewissermaßen komplexes Segment gelten müssen.

³⁰ Natürlich können mehrere Personen gleichzeitig sprechen oder auch singen. Sofern dabei intendierte Ordnungsstrukturen im Spiel sind, beruhen sie jedoch, wie man sich leicht klar macht, nicht auf sprachlichen Regeln, während Beziehungen zwischen zwei oder mehr Stimmen Teil der musikalischen Struktur sind.

³¹ Grundlegende Arbeiten sind Trubetzkoy (1939), Jakobson (1941). Wesentliche Gesetzmäßigkeiten vor allem der phonologischen Regelsysteme werden in Chomsky und Halle (1968) dargestellt.

³² Eins der Ziele des oben erwähnten Ansatzes von Grice (1957) besteht gerade darin, die Bedeutung nichtkommunikativer „Anzeichen“ (Grice spricht von der „natürlichen Bedeutung“ solcher Signale) vom intendierten Effekt kommunikativer Äußerungen zu unterscheiden.

³³ Letzteres gilt für die Beziehung zwischen Bedeutung und Form der Zeichen, es ist nicht zu verwechseln mit unscharfen Grenzen, die bei der Signalerzeugung entstehen können. Wenn z. B. bei flüchtiger Schrift die Ziffern 1 und 7 oder die Buchstaben π und μ einander graduell ähnlicher werden, so ändert das nichts am diskreten Charakter der Ziffern- und Buchstabenzeichen: Die Bedeutung des Zeichens 1 geht nicht allmählich in die des Zeichens 7 über, wenn die Realisierungen miteinander verschwimmen, es entstehen höchstens Mißverständnisse.

³⁴ Allerdings mit motivierten Einsprengseln. Bereits erwähnt wurden die Onomatopoeika. Darüber hinaus scheint es verdeckte und noch wenig erforschte Motivierungen zu geben, die die Konventionalisierung steuern. So wird zum Beispiel die Vokalqualität in der Reihenfolge *i - e - a - o - u* einer ganz abstrakten Skala vom Nahen zum Ferneren zugeordnet. *Bim-bam-bum, Sing-sang, ding-dong, Schnick-schnack* und ähnliche Bildungen (nicht nur im Deutschen) halten die Vokalabfolge fest. Auf diesem Hintergrund ist es kein reiner Zufall, daß in *ich* gegenüber *du, hier* gegenüber *dort* oder *bin, singe, gebe, schiebe* gegenüber *war, sang, gab, schob* jeweils die vorderen Vokale *i* und *e* dem „Nahbereich“, die hinteren Vokale *a, o, u* dem „Fernbereich“ zugeordnet sind. Cooper und Ross (1975) haben erste Gesichtspunkte und zahlreiche Beispiele für Erscheinungen dieser Art gesammelt. Insgesamt handelt es sich hier jedoch nur um regulative Tendenzen, die das Prinzip der Konventionalität nicht aufheben.

³⁵ Wiewohl Wörter im Prinzip als die nach Form und Bedeutung schon bekannten Grundzeichen für die normale Sprachverwendung immer schon vorausgesetzt, also als bereits gelernt genommen werden, können stets auch neue Wörter in einer Sprache eingeführt werden. Dafür gibt es in jeder Sprache spezielle Mittel. Am deutlichsten sind Definitionssätze, also Sätze, die Wörter bekannt machen, wie z. B. *Das Velum ist der hintere, weiche Teil des Gaumens*. Das heißt aber andersherum: Die Tatsache, daß es in der Musik kein Analogon zur sprachlichen Definition eines Wortes gibt, ist nur eine andere Seite der eben aufgestellten Behauptung.

Man kann sogar noch schärfer sagen: In der Musik sind Definitionen nicht nur nicht erforderlich, sondern sogar unmöglich. Es gibt keinen Weg, die Bedeutung eines Motivs musikalisch zu definieren. Wir werden noch sehen, was es mit der ganz anderen Möglichkeit auf sich hat, sie mit sprachlichen Mitteln zu beschreiben.

³⁶ Miller und Johnson-Laird (1976) haben in einem weitgreifenden Überblick diesen Zusammenhang von der Strukturierung des Wahrnehmungsfeldes bis zur Konstituierung der konzeptuellen Grundeinheiten und deren Kombination in komplexen logischen Formen dargestellt und für einen repräsentativen Ausschnitt aus dem System der englischen Wortbedeutungen im einzelnen analysiert.

³⁷ Freges „Begriffsschrift“ ist in diesem Sinn der Versuch, eine Idealsprache zu konstruieren, in der die Leibnizsche These voll eingelöst ist und die Verknüpfung der Grundzeichen eindeutig die Beziehungen ihres Anteils an der logischen Form darstellt. Wittgenstein (1921) analysiert die Bedingungen einer solchen Idealsprache, in der ein Satz ein Bild eines Sachverhalts ist, das in seiner Zusammensetzung aus Elementen von Zeichen das zeigt, was es mit der Struktur des Sachverhalts gemeinsam hat (und haben muß, damit es ein Bild ist), in der Form der Elementarzeichen dagegen das verschlüsselt, was das logische Bild eines Sachverhalts nicht zeigen kann. Die Grenzen einer solchen Idealsprache sind in der Folgezeit ausführlich erörtert worden, am schärfsten in Wittgensteins Spätphilosophie.

³⁸ Eine strengere, experimentallpsychologisch ausgerichtete Fassung erhält dieser Zusammenhang bei Simonov (1975). Emotion wird dort als Ergebnis der Wechselwirkung von Bedürfnis und der für die Zielerreichung nötigen Information bestimmt. Mangel an erreichbarer Information bedingt negative Emotion, Informationszunahme, die die Zielerreichung wahrscheinlicher macht, bedingt positive Emotion.

³⁹ Einen Überblick über verschiedene Ansätze zu einer solchen Analyse sowie einen ausführlichen Ordnungsversuch einschlägiger empirischer Befunde gibt Izard (1977). Ihrer Darstellung sind auch die folgenden Prinzipien entnommen.

⁴⁰ Dabei wird, anders herum angesehen, der mit der gesprochenen Sprache immer schon gegebene Sachverhalt genutzt, daß die sprachliche Codierung keineswegs alle

Merkmale des Signals bindet und somit Raum läßt für „gestische Information“, die im Sprechtempo, in der Lautstärke und Modulation in Erscheinung tritt. In der gesungenen Sprache werden diese Merkmalsbereiche systematisch codiert und damit ihre primäre Wirkungsweise verstärkt, differenziert oder aufgehoben.

⁴¹ Im Normalfall wird eine Proposition durch einen Satz ausgedrückt, doch gibt es zahlreiche Abweichungen vom Normalfall. *Fritz*, geäußert als Antwort auf die Frage *Wer war dort?* drückt die gleiche Proposition aus wie *Fritz war dort*, obgleich es an sich kein Satz ist. *Den Stuhl dorthin!*, *Ich leider auch*, *Den erst später* sind andere, situationsabhängige Beispiele für solche Abweichungen, oder besser Unvollständigkeiten.

⁴² Den hier nur unzulänglich erläuterten Zusammenhang zwischen der Struktur, die sich in den Bedeutungen von Zeichen zeigt, und dem, was durch sie gesagt werden kann, hat Wittgenstein (1921) am eindringlichsten analysiert. Allerdings ging es ihm dabei um eine Idealsprache, die die logische Form direkt in der Zeichengestalt erkennen läßt. (Vgl. Anmerkung 37). Die Beziehung

Sachverhalt – Proposition – Satz
wird dabei in der direkten Korrespondenz
Sachverhalt – Satz

zusammengefaßt, die Strukturentsprechung zwischen Sachverhalt und Proposition mithin auf die Zeichenform projiziert. Sagen und Zeigen werden damit zu zwei einander bedingenden Aspekten der Form der Zeichen. Hält man indessen Satz und Proposition, *Form und Bedeutung der Zeichen auseinander*, dann bleibt bei konventionellen Zeichen das Zeigen auf die strukturelle Korrespondenz zwischen dem Sachverhalt und seiner kognitiv-begrifflichen Repräsentation, eben der Proposition, beschränkt. Die Äußerung eines Satzes sagt dann etwas, weil die mit ihm verbundene Proposition die logische Struktur des Sachverhalts zeigt, aber gedanklich, nicht im geäußerten Signal.

⁴³ Die beiden Möglichkeiten haben ganz enge Parallelen bei anderen inneren und äußeren Zuständen und Handlungen: *Er denkt nach*, *Ich springe* illustrieren die erste Möglichkeit, die analysierende Beschreibung kognitiver Prozesse, motorischer Abläufe oder praktischer Handlungen belegen die zweite.

⁴⁴ Die Tatsache, daß sich das, was die Musik zeigt, auf andere Weise, nämlich durch abstraktive Annäherung, auch sagen läßt, ist die eigentliche Grundlage für einen großen Teil der musikalischen Hermeneutik. Die Möglichkeiten und Grenzen dieses Weges sind damit einsehbar: Was Musik bedeutet, läßt sich so gut und so wenig auf sprachlichem Wege ausdrücken, wie eine Bildbeschreibung ein Bild ersetzen kann. Für die Methodologie der deskriptiven Musikwissenschaft, die aller Theoriebildung zugrunde liegen muß, ergeben sich daraus bedenkenswerte Anhaltspunkte. Etwas vergrößernd lassen sich alle Beschreibungen als reine oder gemischte Versionen folgender drei Möglichkeiten bestimmen:

- (a) formal-strukturell: Beschreibung des Aufbaus der Lautmuster,
- (b) hermeneutisch: Beschreibung der gestischen Bedeutung,
- (c) „inhaltlich“: Beschreibung (einer Variante) des Sachbezugs, der mit dem codierten emotionalen Muster verbunden ist (oder sein kann).

Auf Grund der Überlegungen im Abschnitt 12 ist klar, wie (c) mit musikalischen Zeichen zusammenhängt, daß aber dabei nicht die musikalische Wirkungsweise musikalischer Signale erfaßt wird. Auch mehr oder weniger programmgebundene Musik wie Honeggers *Pacific 231* oder Vivaldis *Vier Jahreszeiten* hat ihre Wirkung nicht durch die „Inhalte“, sondern in dem, was aus ihrem Anlaß gezeigt wird. Die Beschreibungsform (a) ist insofern die präziseste, als für sie ein sprachlich-begriffliches Instrumentarium bereitsteht, das die strukturellen Eigenschaften musikalischer Lautmuster genau

zu erfassen gestattet. Sie erfaßt den Sinn, die Bedeutung musikalischer Gebilde, genau in dem Maß, in dem er sich in diesen Mustern zeigt, aber sie kann ihn nicht direkt benennen. Ein Thema als wichtig, aufrüttelnd, fröhlich, eine Variation als gedrückt oder überschwänglich zu charakterisieren, schließt deshalb den Übergang zur Beschreibungsförmigkeit (b) ein. In der Tat glaube ich, daß angemessene Beschreibungen musikalischer Gebilde die Verbindung von (a) und (b) verlangen; sie ist durch (c) zu ergänzen, wenn die Erfassung der Erfahrungsanlässe notwendig ist, da ja musikalische Wirkungen nicht generell mit der Bedeutung musikalischer Zeichen identisch sind, wie ich in Abschnitt 4 erläutert habe.

⁴⁶ Natürlich reicht „Scheinhaftigkeit“ nicht aus, um Kunst zu konstituieren. Insbesondere muß der Als-ob-Charakter genauer qualifiziert werden, wobei u. a. das überindividuell Verbindliche des im Als-ob-Modus Präsentierten, also seine gesellschaftliche Repräsentanz, eine wesentliche Rolle spielt. Eine neuere Erörterung dieser Problematik findet sich bei Kirsch (1976). Ich kann und brauche dieses Problem nicht zu verfolgen, da es hier nur auf die unzweifelhafte Als-ob-Modalität von Kunst ankommt.

⁴⁶ Der Unterschied zwischen ‚echter‘ und stellvertretend gezeigter Emotion ist in Brechts Konzeption der Verfremdung in genau diesem Sinn zum bewußt gehandhabten Faktor gemacht. Er kommt aber keineswegs erst durch die Verfremdung zustande, sondern ist immer schon da: Auch konsequenteste „Einfühlung“ bringt eine Darstellerin der Margarete nicht dazu, Selbstmord zu begehen.

⁴⁷ Eben darum sind Signale, musikähnliche Pfeifen- und Trommelsprachen und ähnliche Systeme primär keine Musik: Sie werden – wegen ihrer Konventionalität – ursprünglich im faktischen Modus gebraucht: Das Wecksignal fordert tatsächlich zum Aufstehen auf, es fingiert diese Situation nicht.

⁴⁸ Wie wir sogleich sehen werden, ist z. B. die Betrachtung der linearen Verknüpfung im allgemeinen keineswegs auszeichnend. Bereits der Aufbau der Wortformen beruht, wie ich in Abschnitt 9 erläutert habe, auf komplizierteren Operationen als der bloßen Aneinanderreihung der Segmente: Einerseits sind Segmente nicht-lineare Verbindungen von Merkmalen, andererseits werden Segmente nicht einfach zu Wörtern, sondern zunächst auf Grund rein lautlich bedingter Regeln zu Silben verknüpft: *Lau-te, lau-fen, Lau-ge, Wo-ge* deuten diese Gliederung an. Noch weit komplizierter wird das Bild, wenn man den sogenannten morphologischen Aspekt der Wortstruktur berücksichtigt. Unter diesem Gesichtspunkt stellt sich z. B. *laufen* nicht als Verknüpfung der Silben *lau* und *fen*, sondern des Stammes *lauf* und des Infinitivsuffixes *en* dar. Um die Grundzüge unseres Problems zu verdeutlichen, muß ich diesen Aspekt hier zunächst außer Betracht lassen. Ich nehme ihn in Abschnitt 17 wieder auf.

⁴⁹ So muß einerseits der Charakter der Eigenschaft *A* näher festgelegt werden. In der hier gegebenen Formulierung könnte *A* z. B. auch die Länge eines Wortes sein, die sich aus der Länge der Segmente und ihrer linearen Verknüpfung ergibt, und damit würde es Fälle geben, in denen die syntaktische Struktur mit der Struktur im allgemeinen Sinn entgegen der anvisierten Unterscheidung zusammenfällt. Da uns aber im folgenden ohnehin *A* vornehmlich im Sinn der Bedeutung von Zeichen interessiert, ist eine allgemeinere Kennzeichnung überflüssig. Zum anderen muß (S 1) durch eine Klausel abgesichert werden, die bestimmte, durch Zusatzfaktoren bedingte Fälle ausnimmt. Sonst würde z. B. ein Ausdruck wie *zur Schule bringen* eine bestimmte syntaktische Struktur aufweisen, aber nicht die Redewendung *zur Strecke bringen*, denn nur im ersteren, nicht aber im letzteren Fall ist die Bedeutung des ganzen Ausdrucks durch die Bedeutung der Wörter und ihre Verknüpfung determiniert. Dennoch müssen Redewendungen der angedeuteten Art als syntaktisch strukturiert gelten. Solche Erscheinungen

spielen zwar eine besonders zu berücksichtigende Rolle, heben aber das generelle Prinzip, auf das es hier ankommt, nicht auf.

⁵⁰ Es genügt hier, einen intuitiven Begriff der Anordnung zugrunde zu legen. Eine exaktere mathematische Bestimmung müßte zunächst die Teile von *K* als Intervalle bezüglich *D* festlegen und dann die Ordnung für die unteren und oberen Grenzen dieser Intervalle definieren.

⁵¹ Eine genauere Betrachtung zeigt, daß sie eng mit Funktion und Genese der verschiedenen Systeme zusammenhängt. Hier nur vier Hinweise. Erstens: Verkehrszeichen sind zwar strukturierte Gebilde, besitzen aber im eigentlichen Sinn keine Syntax, da sich aus der Kombination von Form- und Farbwerten nicht die Bedeutung der Zeichen ergibt. (Sie sind darin den Wörtern, nicht den Sätzen der Sprache vergleichbar.) Zweitens: Freges Begriffsschrift (vgl. Anmerkung 37) ist eine Kunstsprache, die in gewissem Sinn die Struktur der Gedanken in Zeichenformen zu fixieren versucht und folgerichtig zwar hierarchische Beziehungen benutzt, sie aber nicht der Linearisierung unterwirft. (Daß die Gedanken selbst schon eine „Syntax“ im eigentlichen Sinn haben, also die Bedingung (S 1) erfüllen, ist eine sinnvolle Charakterisierung unter der Voraussetzung, daß man das, was sie repräsentieren oder widerspiegeln, als ihre funktionale Eigenschaft ansieht, so wie die Bedeutung die funktionale Eigenschaft von Zeichenformen ist.) Drittens: Die Strukturformeln der Chemie sind eine Kunstsprache, die Verknüpfungsstrukturen von Atomen in die Fläche projiziert und dabei weder lineare noch hierarchische Beziehungen auszeichnet. Viertens: Die Zifferndarstellung der Zahlen, d. h. der Aufbau von Ziffernfolgen, nicht der einzelnen Ziffern, beruht auf einer rein linearen Verknüpfung, wobei eine generelle Rechenregel die Funktion der Verknüpfung festlegt. Die Rechenregel lautet: In einer Ziffernfolge gibt die Stellenzahl einer Ziffer, gezählt von rechts (bzw. vom Komma), die Zehnerpotenz an, mit der die durch die Ziffer codierte Zahl zu multiplizieren ist; die Summe der so gebildeten Werte ist die durch die Ziffernfolge repräsentierte Zahl.

⁵² Die volle Kombinatorik der drei Bedingungen (S 1) bis (S 3) ergibt 8 Möglichkeiten. Die fehlenden drei sind, wie die Verkehrszeichen, Systeme, bei denen die Strukturbildung keinen funktionalen Charakter hat, also nicht im engeren Sinn syntaktisch ist. Ich gehe deshalb hier nicht auf sie ein.

⁵³ Um Mißverständnisse abzufangen, die aus der Doppeldeutigkeit des Wortes „Satz“ entstehen können: Ein Satz als Teil einer Komposition (englisch „movement“) kann sehr spezifische Strukturen aufweisen, doch entsprechen sie eher denen eines Abschnitts oder Kapitels, mit Sicherheit aber nicht denen eines sprachlichen Satzes. Das wird im nächsten Abschnitt deutlich werden.

⁵⁴ Letzteres ist trivial, wenn wir die früher begründete Tatsache berücksichtigen, daß die logische Form, die die Bedeutung sprachlicher Ausdrücke bildet, gar keine zeitlich-lineare Struktur besitzt. Es gilt aber selbstverständlich auch für die Sachverhalte, auf die sich Sätze beziehen: *Der Wolkenbruch folgte einer langen Dürre* nennt die Ereignisse in umgekehrter Folge ihres Eintretens.

⁵⁵ Eine gewisse Form von Motiviertheit bringen allerdings auch die konventionsbedingten Prinzipien, denen die Interpretation der sprachlichen Syntax unterliegt, in die Codierungsform sprachlicher Zeichen. Es ist dies die Motiviertheit, die sich aus dem Funktionalitätsprinzip (S 1) für die Syntax grundsätzlich ergibt. Ohne diese Motiviertheit könnte man nicht beliebige neue Sätze verstehen, sondern müßte sie wie Wörter Satz für Satz neu lernen. Nur ist eben diese Motivierung selbst konventioneller Natur.

⁵⁶ Für die Sprache habe ich syntaktische Kategorien im Sinn von (Z d) in Abschnitt 16, semantische Kategorien im Sinn von (Z e) in Abschnitt 11 zumindest andeutungsweise

erläutert. Die Unterscheidung von *k* und *t* ist notwendig, da z. B. zwei verschiedene Zeichen zur gleichen semantischen, aber zu verschiedenen syntaktischen Kategorien gehören können. Für die Musik müssen syntaktische und semantische Kategorien eines Zeichens wegen der Isomorphie von Form und Bedeutung grundsätzlich gleich sein. Wie diese Kategorien aussehen (Exposition, Coda, Überleitung könnten Kategorien solcher Art sein), muß im Rahmen der jeweiligen Codes begründet werden. Ob Elementarzeichen verschiedenen Kategorien angehören, ist ein offenes Problem.

⁵⁷ Eine solche Zerlegung besteht natürlich nicht einfach in der Identifizierung etwa von Haupt- und Seitenthema eines Sonatensatzes. Sie muß auch Begleitfiguren, Überleitungen, Modulationen als Grundzeichen identifizieren oder sie in solche zerlegen. Die Bedeutung solcher „Nebenzeichen“ wäre dann umschreibbar als ein Hintergrundgestus oder ein gestischer Übergang im Gefüge der gestischen Gesamtform.

⁵⁸ Das ist im streng formalen Sinn unrichtig, falls ein musikalischer Code auf einem endlichen System von Merkmalsdimensionen und Merkmalswerten beruht und Grundzeichen nicht beliebig lang sein können. Dann gibt es nämlich eine endliche Kombinatorik möglicher Merkmalsstrukturen. Es ist jedoch klar, daß diese Strukturen nicht im gleichen Sinn Teil der Musikkenntnis sind wie die Wörter einer Sprache Teil der Sprachkenntnis sind.

⁵⁹ Mir scheint, daß dies die eigentliche Grundlage für die Unterscheidung ist, die Knepler (1977, S. 108 ff.) mit der Gegenüberstellung von logogenen und biogenen Elementen der Musik faßt und allerdings ganz anders motiviert. Eine genauere Begründung dieser Behauptung hängt mit der Auffassung über das Verhältnis der Funktionsprinzipien von Musik und Sprache zusammen, die sich bei Knepler wesentlich von der hier begründeten unterscheidet.

⁶⁰ Dazu gibt es eine Ausnahme: die Gruppe von Regeln, die die anaphorischen Beziehungen von Pronomina determinieren, etwa für *ihn* in *Hans war das nicht. Ich kenne ihn gut genug* wirken über die Satzgrenze hinaus. Der Sondercharakter dieser Regeln hebt das Prinzip nicht auf.

⁶¹ Daß viele Personen den Namen *Franz Müller* haben, ändert daran nichts. Im Konfliktfall wird es nämlich automatisch nötig, die Namen durch Zusätze oder Änderungen zu unterscheiden und die Eindeutigkeit wiederherzustellen.

⁶² Daß in diesem Rahmen dann faktische, empirische Zusammenhänge aufgedeckt und erklärt werden müssen, ist an die Untersuchung einzelner Zeichensysteme und Typen von Zeichensystemen und Kommunikationsformen gebunden. Sie geben dem logischen Rahmen seinen sachlichen Gehalt, sie bereichern das allgemeine Konzept mit Tatsachenerkenntnissen an. (Vgl. die Bemerkungen in Anmerkung 1.)

Literaturverzeichnis

- Apel, Willi: *Die Notation der polyphonen Musik*. Leipzig 1962.
- Bierwisch, Manfred: *Poetik und Linguistik*. In: Helmut Kreuzer und Rul Gunzenhäuser (ed.): *Mathematik und Dichtung*. München 1965.
- : *Strukturalismus. Geschichte, Probleme, Methoden*. In: Kutsbuch V. Frankfurt a. M. 1966.
- Chomsky, Noam: *Syntactic Structures*. The Hague 1957.
- : *Aspects of the Theory of Syntax*. Cambridge, Mass. 1965. Deutsch: *Aspekte der Syntax-Theorie*. Berlin 1970.
- Chomsky, Noam; Halle, Morris: *The Sound Pattern of English*. New York, Evanston, and London 1968.
- Cooper, William E.; Ross, John R.: *World Order*. In: Papers from the Parasession on Functionalism. Chicago Linguistic Society. Chicago 1975.
- Eco, Umberto: *A Theory of Semiotics*. Bloomington 1976.
- Freeman, Donald C. (ed.): *Linguistics and Literary Style*. New York 1970.
- Grice, Paul: *Meaning*. Philosophical Review LXVI (1957).
- Halle, Morris: *On Meter and Prosody*. In: Manfred Bierwisch and Karl Erich Heidolph (ed.): *Progress in Linguistics*. The Hague 1970.
- Harnad, Stephen R.; Stekliš, Horst D.; Lancaster, Jane (ed.): *Origins and Evolution of Language and Speech*. Annals of the New York Academy of Sciences Vol. 280. New York 1976.
- Hjelmslev, Louis: *Prolegomena to a Theory of Language*. Bloomington 1953.
- Ihwe, Jens: *Linguistik in der Literaturwissenschaft*. München 1972.
- Izard, Carol E.: *Human Emotions*. New York and London 1977.
- Jakobson, Roman: *Kindersprache, Aphasie und allgemeine Lautgesetze*. Uppsala 1941.
- Jensen, Hans: *Die Schrift in Vergangenheit und Gegenwart*. Berlin 1958.
- Katz, Jerrold J.: *Semantic Theory*. New York, Evanston, and London 1972.
- Kirsch, Rainer: *Das Wort und seine Strahlung*. Berlin 1976.
- Klix, Friedhart: *Über Grundstrukturen und Funktionsprinzipien kognitiver Prozesse*. In: Friedhart Klix (ed.): *Psychologische Beiträge zur Analyse kognitiver Prozesse*. Berlin 1976.
- Knepler, Georg: *Geschichte als Weg zum Musikverständnis*. Leipzig 1977.
- Lieberman, A. M.; Cooper, F. S.; Shankweiler, D. P.; Studdert-Kennedy, M.: *Perception of the Speech Code*. Psychological Review LXXIV (1967).
- Lindblom, Björn; Sundberg, Johan: *Towards a generative theory of melody*. Svensk tidskrift för musikforskning LII (1970).
- Lyons, John: *Introduction to Theoretical Linguistics*. Cambridge 1968.
- Marcus, Solomon: *Mathematische Poetik*. Frankfurt a. M. und Bukarest 1973.
- Metze, Erika; Stejngart, K. M.: *Wechselbeziehungen im Funktionssystem der Sprache*. In: Manfred Bierwisch (ed.): *Psychologische Effekte sprachlicher Strukturkomponenten*. Berlin 1978.
- Miller, George A.; Johnson-Laird, Philip N.: *Language and Perception*. Cambridge 1976.
- Miller, George A.; Galanter, Eugene; Pribram, Karl H.: *Plans and the Structure of Behavior*. New York 1960.
- Peirce, Charles S.: *Collected Papers*. Vol. II. Cambridge, Mass. 1932.

- Riemann, Hugo: *Große Kompositionslehre*. 3 Bände. Berlin und Stuttgart 1902, 1903, 1913.
- Sartre, Jean Paul: *Esquisse d'une théorie des émotions*. Paris 1939.
- Saussure, Ferdinand de: *Cours de linguistique générale*. Paris 1916.
- Simonov, P. V.: *Die höhere Nerventätigkeit des Menschen. Motivational-emotionale Aspekte* (russ.). Moskau 1975.
- Stockmann, Doris: *Die Erforschung vokaler und instrumentaler Praktiken im mittelalterlichen Rechtsleben*. In: Deutsches Jahrbuch der Musikwissenschaft für 1973 bis 1977, Leipzig 1978.
- Trubetzkoy, Nikolai Sergejewitsch: *Grundzüge der Phonologie*. Prag 1939.
- Wittgenstein, Ludwig: *Logisch-philosophische Abhandlung* (1921). In: Ludwig Wittgenstein: *Schriften*. Bd. 1. Frankfurt a. M. 1960.
- Ziff, Paul: *On H. P. Grice's account of meaning*. Analysis XXVII (1967).