

3 PSYCHOLOGISCHE GRUNDLAGEN DER RHYTHMUSERFAHRUNG

3.1 Grundkonzepte der psychologischen Rhythmusforschung

Rhythmus - eine Begriffsklärung aus musikpsychologischer Perspektive

In der Psychologie wird bekanntlich das Verhalten des Menschen untersucht – sein äußeres, direkt beobachtbares Verhalten ebenso wie die inneren, psychischen Vorgänge, u.a. Wahrnehmen, Denken, Erkennen, Fühlen, Erinnern, Vorstellen, Lernen und Motivation. Befassen sich Psychologen mit Rhythmusphänomenen, so stehen die inneren und äußeren Verhaltensweisen im Blickpunkt, mit denen Menschen zeitlich strukturierte Ereignisfolgen erzeugen oder auf zeitliche Strukturen in ihrer Umwelt reagieren. Rhythmen begegnen dem Menschen in vielen Lebensbereichen, so bei Abläufen und Bewegungen in der natürlichen Umwelt und im sozialen Zusammenleben, bei Körperbewegungen, beim Sprechen – und eben in der Musik. Dennoch gibt es in der Psychologie bislang keine griffige, allgemein anerkannte Rhythmusdefinition (vgl. Spitznagel 2000). Vielmehr hängt das jeweilige Rhythmusverständnis vom untersuchten Phänomenbereich sowie von den theoretischen und methodologischen Ansätzen ab, die bei der Erforschung dieser Phänomene angewendet werden.

In den vergangenen Jahrzehnten ist Rhythmusphänomenen innerhalb der musikpsychologischen Forschung große Aufmerksamkeit zuteil geworden (vgl. Gabrielsson 1986, 1993, Handel 1989, S. 382-459, Clarke 1999, Bruhn 2000, Schulze 2005). Innerhalb der musikpsychologischen Forschung dient „musikalischer Rhythmus“ vielfach als Überbegriff für eine Reihe von Phänomenen der Wahrnehmung und Gestaltung von musikalischen Strukturen in unmittelbar überschaubaren Zeitdimensionen. Der Musikpsychologe Alf Gabrielsson schlägt folgende Definition vor: „[...] the experience of rhythm includes some kind of grouping, accents, and

regularity within the limits of the psychological present“ (Gabrielsson 1993, S. 97). Die musikalische Rhythmus Erfahrung stützt sich demnach auf die Gruppierung von Klangereignissen¹ innerhalb enger zeitlicher Grenzen, auf die unterschiedlichen Gewichtungen oder Akzentuierungen der einzelnen Klänge sowie auf eine gewisse Regelmäßigkeit in der zeitlichen Struktur der Klänge.² Wohlgermerkt spricht Gabrielsson nicht von „Rhythmus“, sondern von „Rhythmus Erfahrung“ („rhythm experience“). Dies kommt nicht von ungefähr. In der Musikpsychologie stehen weder die physikalischen Strukturen des Schalls noch die extern kodierte Informationen (Notenschrift, mechanische oder elektronische Speichermedien) im Mittelpunkt des Interesses, sondern vielmehr die inneren und äußeren Verhaltensweisen der Menschen. Rhythmus wird daher als Phänomen des Gestaltens und Wahrnehmens betrachtet. Die Angemessenheit der psychologischen Perspektive erweist sich unter anderem daran, dass sich in manchen Fällen die wahrgenommenen rhythmischen Strukturen anhand der akustischen Stimuli nicht eindeutig bestimmen oder vorhersagen lassen. Bedenkt man weiter, dass in Musik normalerweise mehrere Klangfolgen, deren rhythmische Strukturen womöglich voneinander abweichen, simultan erklingen, so ergeben sich mehrere bzw. mehrdeutige Möglichkeiten der musikalischen Rhythmuswahrnehmung. Es genügt somit nicht, rhythmische Strukturen allein anhand von akustischen Daten oder Notentexten zu beschreiben. Vielmehr gilt es, die psychologischen Gesetzmäßigkeiten zu untersuchen, auf deren Grundlage Rhythmus wahrgenommen und erfahren wird.

Wie bereits im Einführungskapitel angedeutet, kann zwischen zwei Verwendungsweisen des Begriffs „musikalischer Rhythmus“ unterschieden werden. In einer ersten Bedeutung ist jede Klangfolge zugleich ein Rhythmus bzw. eine Rhythmusgestalt. In diesem Sinne ist auch jede Melodie oder melodische Gestalt zugleich eine Rhythmusgestalt; umgekehrt sind jedoch nur jene Rhythmusgestalten, die Tonhöhenverläufe aufweisen, zugleich Melodien. Aufgrund der sich ergebenden terminologischen Unschär-

-
- 1 Mit „Klang“ und „Klangereignis“ sind im Folgenden alle elementaren Einheiten der auditiven Wahrnehmung gemeint; dies beinhaltet auch Geräusche oder Geräuscheignisse. „Klang“ wird somit nicht auf die enge physikalische Bedeutung eines Gemisches aus Sinustönen eingeschränkt, sondern wird eher im Sinne des engl. Ausdrucks „sound“ verwendet; ausgeschlossen bleibt allerdings kontinuierliches Rauschen („noise“). Ich ziehe „Klang“ und „Klangereignis“ dem Ausdruck „auditives Objekt“ vor, für den Herbert Bruhn (2000) plädiert; andere Autoren wählen die Bezeichnungen „auditory event“ oder „auditory image“ (vgl. Bruhn 2000, S. 44).
 - 2 Klangfolgen, die nicht als geordnet oder strukturiert erfahren werden, könnten dagegen als *unrhythmisch* oder *arhythmisch* - analog zum Begriff der Atonalität - bezeichnet werden (vgl. Seidel 1976, S. 3). Gabrielsson schreibt hierzu: „We would thus experience non-rhythm if these conditions are not met, that is, if there were no perceived grouping, accent or regularity, or exceeding the temporal limits for what we can perceive as belonging together at once“ (Gabrielsson 1993, S. 97).

fe wird der Rhythmusbegriff im Folgenden jedoch vorwiegend in einer zweiten Bedeutung verwendet, nach der jede Klangfolge (so auch jede Melodie) einen Rhythmus bzw. mehrere Rhythmen oder Rhythmuskomponenten besitzt – im Sinne einer oder mehrerer aus der Klangfolge abstrahierten zeitlichen Struktur(en). Dabei fließen in die wahrgenommene Rhythmusstrukturen immer auch nicht-zeitliche Informationen über die Klangereignisse (z.B. Unterschiede der Lautstärke, Klangfarbe oder Tonhöhe) mit ein.

In der Psychologie werden Folgen von Klängen, deren Reihenfolge sich beschreiben lässt, in einem allgemeinen Sinne als auditive Sequenzen, sequenzielle Muster oder Patterns bezeichnet (vgl. Michon 1978, S. 90). Bezieht sich die Beschreibung nicht nur auf die Reihenfolge der Klänge, sondern auf die genaue zeitliche Struktur von Klangdauern und Pausen, so handelt es sich um *zeitliche* Sequenzen oder Patterns (vgl. Michon 1978, S. 90f). Die zeitlichen Eigenschaften der Klangfolgen – Reihenfolge, Dauern und Pausen – werden unter dem Begriff „Timing“³ zusammengefasst (vgl. Handel 1989, S. 384). Zentral für das Timing sind die Abstände zwischen den Anfangspunkten (Onsets) zweier aufeinander folgender Ereignisse, die als Interonsetintervalle (IOI) bezeichnet werden. Eine Reihe von musikpsychologischen Untersuchungen (z.B. Povel/Essens 1985, Parncutt 1994, Handel 1998) befassen sich aus Gründen der experimentellen Kontrollierbarkeit ausschließlich mit den Interonsetintervallen in Sequenzen klanglich identischer Ereignisse und lassen andere Faktoren, z.B. die zeitliche Strukturen des Tonhöhenverlaufs oder der Klangfarben, außer acht. Möglichen Rückschlüssen auf die musikalische Rhythmuserfahrung sind bei diesen Untersuchungen daher Grenzen gesetzt.

Die zeitliche Struktur der Aufmerksamkeit

Musikhören ist eine Tätigkeit, die Aufmerksamkeit erfordert. Das Wahrnehmen von Musik ist kein passives Entgegennehmen von Reizen, sondern ein aktiver Prozess der Informationsverarbeitung. Der aktive, konstruktive Charakter der menschlichen Wahrnehmung spiegelt sich im psychologischen Konzept der Aufmerksamkeit wider. Damit werden in der Psychologie alle Prozesse der Auswahl und aktiven Zuwendung beim Aufnehmen von Informationen aus der Umwelt beschrieben (Goldstein 2002, S. 131f). Aufmerksamkeit ist ein zentraler Aspekt der Informationsverarbeitung, der nicht nur dem Wahrnehmen, sondern ebenso dem Erinnern und Denken sowie zahlreichen praktischen Tätigkeiten zugrunde liegt. Aufmerksam-

3 Abweichend hiervon verstehen Clarke (1999) und Bruhn (2000) unter Timing die mikrorhythmischen Abweichungen von einer isochronen Pulsfolge; hierfür verwende ich den Ausdruck „Mikrotiming“.

keitsprozesse sind in der Regel bewusst. Allerdings gibt es auch viele automatisierte und unbewusst gewordene Abläufe. Dadurch wird möglich, dass gleichzeitig verschiedenen kognitiven Prozessen Aufmerksamkeit zuteil werden kann. Die Aufmerksamkeitsressourcen, über die der Mensch verfügt, sind jedoch beschränkt. Die Kapazität der Aufmerksamkeitsressourcen wurde in der Psychologie in zahlreichen Studien vorwiegend anhand der visuellen und auditiven Wahrnehmung untersucht (vgl. Anderson 2001, S. 75-106).

Während bei der visuellen Wahrnehmung hauptsächlich Auswirkungen der räumlichen Anordnung bestimmter Reize erforscht wurden, ist für die auditive Wahrnehmung deren zeitliche Struktur besonders wichtig (vgl. Jones 1986, Jones/Yee 1993). Das Modell der dynamischen Aufmerksamkeit („dynamic attending“), das die amerikanische Psychologin Mari Riess Jones entwickelt hat, setzt die zeitliche Struktur von Aufmerksamkeitsprozessen mit der zeitlichen Struktur der Umweltereignisse in Verbindung. In ihrem programmatischen Aufsatz *Time, Our Lost Dimension. Towards a New Theory of Perception, Attention, and Memory* (Jones 1976) beklagt Jones, dass in der psychologischen Forschung zu Wahrnehmung, Aufmerksamkeit und Gedächtnis die zeitliche Struktur der verwendeten Stimuli nur ungenügend berücksichtigt wird. Zwar wird die zeitliche Abfolge von Ereignissen, z.B. die äußerst komplexen Abfolgen von Phonemen beim Sprechen, ansatzweise reflektiert (z.B. bei Martin 1972). Die Wahrnehmung und mentale Repräsentation der Ereignisdauern und der zeitlichen Abstände zwischen Ereignissen wird dagegen kaum problematisiert. Eine zeitliche Struktur ist jedoch fast allen Abläufen im Verhalten wie auch in der Umwelt des Menschen eingeschrieben und ist somit „one of the defining properties of our world and so of ourselves“ (Jones 1976, S. 352). Die Bestimmung der zeitlichen Ordnungen, in denen Menschen leben, wird allerdings in der Regel der Physik und deren Zeitverständnis überlassen. Der physikalische Zeitbegriff ist jedoch, so Jones, dem menschlichen Erleben zeitlicher Strukturen nicht angemessen und vermag nicht, die Eigenheiten der menschlichen Zeitwahrnehmung zu erklären. Isaac Newton hat die Zeitvorstellung der modernen Naturwissenschaften in seinen *Mathematischen Prinzipien der Naturphilosophie* paradigmatisch formuliert:

„Die absolute, wahre mathematische Zeit fließt gleichmäßig an sich und ihrer Natur nach, ohne Beziehung auf irgend etwas Äußerliches. Sie wird mit einem anderen Ausdruck als ‚Dauer‘ bezeichnet: relativ, augenscheinlich und gewöhnlich ist ihr beliebiges, sinnliches und durch Bewegung gegebenes äußeres Maß (sei es nun exakt oder ungleichmäßig), das man gewöhnlich anstelle der wahren Zeit verwendet: wie z.B. die Stunde, der Tag, der Monat, das Jahr“ (zit. in: Cramer 1996, S. 37).

Die Vorstellung einer absoluten Zeit, bei der ausdehnungslose Zeitpunkte auf eine Art Zeitstrahl aufgereiht sind, ermöglichte zwar die Entwicklung der Differentialrechnung, mit der Bewegungen physikalischer Objekte mathematisch exakt als eine Folge von Zuständen oder Zeitpunkten minimaler Dauer beschrieben werden können. Dabei wird jedoch die menschliche Erfahrung von Gegenwart ignoriert und die Erfahrung von Zukunft und Vergangenheit als gleichwertig beschrieben, da der Zeitstrahl prinzipiell umkehrbar ist und daher die Zeit in beide Richtungen verlaufen kann. Die Vorstellung eines reversiblen Zeitpfeils hat jedoch wenig mit der psychologischen Erfahrung von alltäglichen Ereignissen und Abläufen gemeinsam. Die von Newton angesprochene relative Zeit, die ja auch im Alltag zur Zeitmessung und zur Koordination sozialen Handelns verwendet wird, nutzt die Dauer bestimmter Bewegungen – z.B. der Pendelbewegung einer Turmuhr – als eher zufällige Maßeinheit, mit der die Dauer bestimmter Ereignisse und Abläufe losgelöst von deren realen zeitlichen Strukturen gemessen werden kann.

Zeitliche Kohärenz, Erwartung, Synchronisation

Der Konzeption der Zeit als einer Linie von Zeitpunkten sowie als Dauern mit zufälligen Anfangs- und Endpunkten stellt Jones die dynamische Ereigniszeit gegenüber, bei der Zeit ein konstitutiver Teil der Ereignisse und Abläufe der alltäglichen Umwelt ist:⁴

„Time periods play an intrinsic role in the relational structure of everyday events. Not only do we find characteristic markers that identify commonplace events, but these markers often outline characteristic time periods, ones of constrained lengths, that also go along with these events. That is, conversations, turntaking times, symphonies, and so on cannot take on any value; there are limits to the tempi found in action patterns, in speech, and in music and there are natural limits to whole event durations as well“ (Jones 1990, S. 208).

Ereignisse und Abläufe in der Erfahrungswelt des Menschen besitzen demnach charakteristische Zeitdauern; ihre Anfänge und Schlüsse sowie Einschnitte innerhalb der Abläufe sind in der Regel durch bestimmte Markierungen gekennzeichnet.

4 Jones (1990) prägt für die beiden entgegengesetzten Zeitvorstellungen der Physik und der Alltagserfahrung die Ausdrücke „high church time“ (physikalische Zeit) und „garden variety time“ (alltägliche Ereigniszeit). Diese Unterscheidung steht quer zur Subjektivierung der Zeit, wie sie von Philosophen wie Kant und Bergson vorgenommen wurde. Denn Zeit ist nach Jones keine bloße Wahrnehmungskategorie, die sich an die physikalische Zeit (Kant) oder die Phänomene (Bergson) anlehnt, sondern ist den Ereignissen und Abläufen der realen Welt eingeschrieben (vgl. Jones/Boltz 1989, S. 487).

Die zeitlichen Ereignisstrukturen können nun unterschiedliche Grade an zeitlicher Kohärenz aufweisen. Ein hoher Kohärenzgrad besteht bei all jenen zeitlichen Ereignisstrukturen, die hierarchisch organisiert sind – Jones spricht von „nested time levels“, von ineinander geschachtelten Zeitebenen. Die hierarchische Organisation bezieht sich dabei auf „[...] a time structure, in which the temporal distribution of markers reveals nested time levels that are consistently related to one another at a given level by ratio or additive time transformations“ (Jones/Boltz 1989, S. 465).

Nach Jones und Boltz lassen sich zwei Arten der hierarchischen Zeitstrukturierung unterscheiden. Bei der sog. Verhältniszeittransformation („ratio time transformation“, „vertikale“ Zeitverhältnisse) können die Zeitintervalle auf den verschiedenen hierarchischen Ebenen durch einfache Multiplikation oder Division mit einer kleinen ganzzahligen Konstanten ineinander überführt werden. Mathematisch lässt sich dieser Transformationstyp durch folgende Exponentialfunktion beschreiben: $\Delta T_n = \Delta T_0 \cdot C_t^n$ (mit $n = 1, 2, 3 \dots$). Die Zeiteinheiten (ΔT_n) der n-ten hierarchischen Ebene ergeben sich, indem die kleinste Zeiteinheit (ΔT_0) n Mal mit einer Konstanten (C_t) multipliziert wird (daher C_t^n). So führt z.B. die Konstante $C_t = 2$, ein bei musikalischen Strukturen weit verbreiteter Fall, zu einer jeweiligen Verdopplung der Zeiteinheit ΔT_n einer Ebene gegenüber der Einheit der vorhergehenden Ebene ΔT_{n-1} . Bei der additiven Zeittransformation („horizontale“ Zeitverhältnisse) werden Dauern der tieferen hierarchischen Ebene (kürzere Zeitintervalle) durch Addition von Konstanten in Dauern der höheren Ebene (längere Zeitintervalle) überführt: $\Delta T_n = \Delta T_0 + n \cdot C_t$ (mit $n = 1, 2, 3 \dots$). Allerdings sind nicht alle Abläufe in der menschlichen Erfahrungswelt streng hierarchisch organisiert. Nicht-hierarchische Ereignisstrukturen weisen jedoch einen weit geringeren Grad an zeitlicher Kohärenz auf und sind daher schwerer kognitiv zu verarbeiten.

Jones und Boltz (1989, S. 466f) nennen eine Reihe von alltäglichen Ereignissen und Abläufen, die hierarchisch organisiert sind und daher eine relativ hohe zeitliche Kohärenz besitzen, darunter verschiedene Formen der Körperbewegung (Gehen, Gesten, Tanz), Sprechen und Musik. Viele dieser Abläufe, z.B. die regelmäßigen Bewegungen der Gliedmaßen bei der Fortbewegung, lassen sich als Kombination der beiden Zeittransformationstypen beschreiben. Aus diesen Kombinationen entstehen, so Jones und Boltz, die typischen „rhythmischen Stile“ bestimmter Abläufe:

„[...] categories of rhythmic styles exist and can be formalized by particular combinations of ratio and additive time invariants. Both time transformations have meaning: The ratio base offers predictability for coordinative gestures, whereas additive time changes not only characterize an individual's style but they can also signal underlying velocity properties“ (Jones/Boltz 1989, S. 466).

Zur Veranschaulichung der beiden Zeittransformationstypen, die der zeitlichen Struktur von alltäglichen Bewegungsabläufen zugrunde liegen, verweisen Jones und Boltz auf das Beispiel einer streng metrisch organisierten Melodie, bei der sich die hierarchischen Ebenen aus der kleinsten Zeiteinheit (200 ms, Achtelnote) durch mehrmalige Multiplikation mit dem Faktor $C_t = 2$ ergeben, sowie auf die Fortbewegung einer Katze, die sich zwar auf den höheren Zeitebenen als Verhältniszeittransformation beschreiben lässt, bei der jedoch der Feinstruktur des Bodenkontaktes der vier Gliedmaßen auf der untersten hierarchischen Ebene ($n = 0$) eine additive Zeittransformation (mit einer Konstanten von 50 ms) zugrunde liegt (Abb. 2).

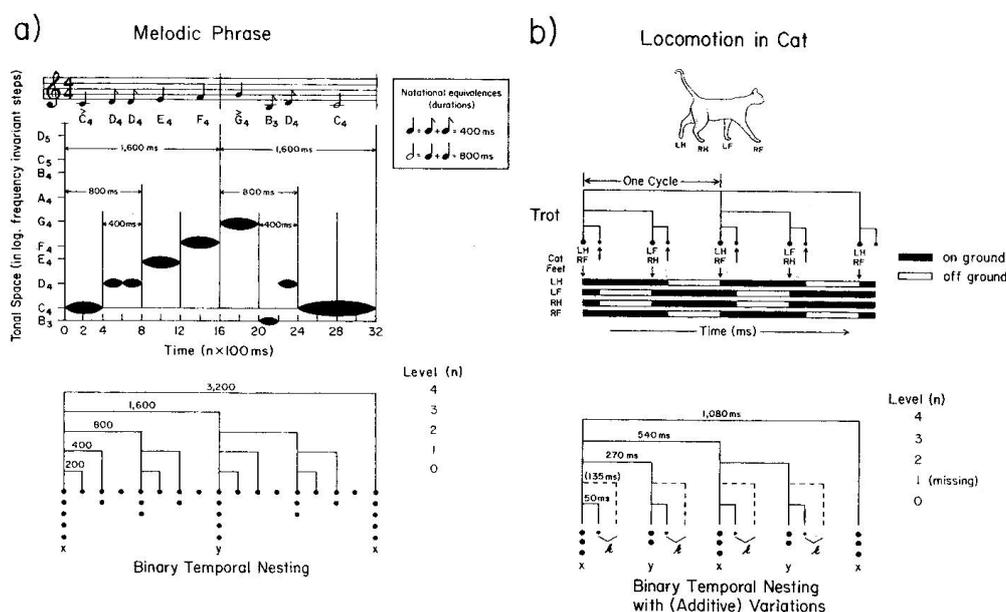


Abb. 2: Zwei Beispiele für hierarchische Zeitstrukturen mit fünf bzw. vier Ebenen; a) Verhältnistransformation, Beispiel einer Melodiephrase; b) Kombination von Verhältnistransformation (Ebenen 2 bis 4) und additiver Transformation (Ebene 0), Beispiel der Fortbewegung einer Katze (aus Jones/Boltz 1989, S. 463).

Der psychische Prozess der dynamischen Aufmerksamkeit vermag sich nun, so die Kernthese von Jones' Modell, mit der zeitlichen Struktur der wahrgenommenen Ereignisse zu synchronisieren. Bei der Synchronisation zwischen Aufmerksamkeit und Ereignisstruktur beginnt das Aufmerksamkeitssystem („the attender“) bestimmte Aktivierungsmuster mit der Zeitstruktur der Ereignisse zu teilen, wobei sich ein oder mehrere Aufmerksamkeitsrhythmen in ihrer Frequenz und Phase den entsprechenden Zeitspannen innerhalb des externen Ereignisablaufs anpassen. Der Aufmerk-

samkeitsfokus kann dabei auf verschiedene zeitliche Dimensionen mit unterschiedlich langen Zeitdauern gerichtet werden. Gerade die Flexibilität der zeitlichen Aufmerksamkeitsfokussierung ist ein zentrales Charakteristikum der menschlichen Wahrnehmung.

Jones und Boltz unterscheiden zwischen einem zukunftsorientierten und einem analytischen oder detailorientierten Aufmerksamkeitsmodus:

„[...] future-oriented attending is particularly functional when longer term preparatory attending and real time extrapolations of an unfolding event are involved. A future-oriented extrapolation is termed an *expectancy*. Analytic attending is more functional when relatively short-term activities are involved. Because relevant focal periods are usually less than that of the referent period, expectancies as such are not supported, although attending to local detail and grouping with respect to the referent time span“ (Jones/Boltz 1989, S. 471).

Die Verschiebungen des Aufmerksamkeitsfokus („attending shift“) können bewusst von bestimmten Zielen und Interessen geleitet werden. So kann z.B. ein Musikhörer seine Aufmerksamkeit auf größere formale Strukturen, auf den Bau melodisch-rhythmischer Phrasen oder auf Details der Klanggebung oder Artikulation richten. Bei zeitlich inkohärenten Ereignisstrukturen und bei überraschenden Ereignissen – etwa einem Hustenanfall des Sitznachbarn im Konzert – erfolgt in Ermangelung hierarchisch gegliederter Zeitstrukturen ein „Umschalten“ auf einen analytischen Aufmerksamkeitsmodus mit einer relativ kurzen Zeitdimension. Je nach Grad der zeitlichen Kohärenz der Ereignisse kommt es demnach zu unterschiedlichen Typen der Aufmerksamkeit. Hohe zeitliche Kohärenz aufgrund einer hierarchischen Ereignisstruktur ermöglicht eine zukunftsorientierte Aufmerksamkeit und den Aufbau von Erwartungen. Jones und Boltz setzen daher die Kohärenz zeitlicher Strukturen auch mit deren Vorhersehbarkeit gleich (vgl. Jones/Boltz 1989, S. 461). Dagegen dominiert bei nicht-hierarchischen und daher zeitlich inkohärenten Ereignissen ein analytischer Aufmerksamkeitsmodus, der sich auf Details in relativ kurzen Zeitspannen richtet.

Der hohe Stellenwert eines verlässlichen zeitlichen Erwartungsrahmens für die Musikwahrnehmung zeigt sich darin, dass die klanglichen Strukturen der meisten Musikrichtungen so gebaut sind, dass sich ein Grundschatz (s. unten) und eine regelmäßige Gliederung der Klangfolgen unmittelbar mitteilt. Natürlich kann die Synchronisation zwischen Aufmerksamkeit und Ereignisstruktur auch scheitern, insbesondere bei relativ inkohärenten Ereignisstrukturen. Oder aber eine zeitlich kohärente, hierarchische Struktur führt zunächst zu Erwartungen (zukunftsorientierter Aufmerksamkeitsmodus), die dann aber von dem tatsächlichen Ablauf nicht eingelöst werden.

Jones und Boltz sprechen von Erwartungsverletzungen und einem Überraschung erzeugenden zeitlichen Kontrast („temporal contrast“) (Jones/Boltz 1989, S. 473). In der Musik haben diese Überraschungen u.a. den Nebeneffekt, dass die bewusste Aufmerksamkeit erhalten bleibt und sich Musikhören nicht zu einem automatisierten Prozess der Informationsverarbeitung verselbständigt, dem keine bewusste Aufmerksamkeit mehr geschenkt werden muss. In der musikalischen Gestaltung bestehen zahlreiche Möglichkeiten, mit Erwartungen zu spielen, sie erst aufzubauen und dann mitunter zu enttäuschen, oder aber mehrere divergente, einander widersprechende Erwartungen zu initiieren.

Die Theorie der dynamischen Ereignisstruktur und der dynamischen Aufmerksamkeitsprozesse hat Implikationen für die Zeitpsychologie, so bei der Frage der Einschätzung von unterschiedlichen Zeitdauern. Zahlreiche Experimente zur Einschätzung von Zeitdauern haben bislang zu widersprüchlichen Ergebnissen geführt, die sich nur schwer durch eine einheitliche Theorie erklären lassen (vgl. Block 1990, S. 9-27, Jones/Boltz 1989, S. 459-461). Jones und Boltz weisen nun darauf hin, dass sowohl der Grad der zeitlichen Kohärenz der Ereignisse, deren Dauern von Versuchsteilnehmern geschätzt werden sollen, als auch der damit verbundene Aufmerksamkeitsmodus sowie die gewählte zeitliche Bezugsebene des Aufmerksamkeitsfokus (kurze oder längere Struktureinheiten) nachhaltigen Einfluss auf die Einschätzung von Zeitintervallen haben. Bei der analytischen Aufmerksamkeit stehen kürzere zeitliche Ebenen im Vordergrund; die entsprechenden Strategien bei der Dauernschätzung nennen Jones und Boltz mnemotechnische Strategien („mnemonic strategies“). Mit diesen Strategien versuchen Versuchsteilnehmer, aufeinander folgende Elemente zu Gruppen zusammenzufassen (Gruppieren) oder den Ereignisfolgen bestimmte zufällige Zeitintervalle und Zeitverhältnisse überzustülpen (Zählen, Uhren). Dabei kommt es zum bekannten Phänomen der als länger empfundenen Dauer eines Zeitintervalls, das mit einer größeren Zahl von Ereignissen gefüllt ist. Dagegen kann sich die Zeitdauernschätzung bei zeitlich kohärenten Ereignissen an strukturellen Markierungspunkten und einfachen Dauerverhältnissen orientieren. Dies führt zu weit genaueren Dauernschätzungen. Die Leichtigkeit und Genauigkeit der Zeitschätzung hängt somit unmittelbar vom Grad der zeitlichen Kohärenz der präsentierten Ereignisse ab. Da sich Menschen jedoch nur selten über die gerade fokussierte Zeitebene und den gerade verwendeten Aufmerksamkeitsmodus bewusst sind, kann es zu beträchtlichen Unterschieden in der Dauernschätzung kommen.

Dass die zukunftsorientierte Erwartung bei hierarchischen Zeitstrukturen tatsächlich weit verbreitet ist und die Dauernschätzung beeinflusst, weisen Jones und Boltz experimentell an der Längeneinschätzung von Variationen einer Melodie nach, deren symmetrischer Phrasenbau durch Stauchen

und Auseinanderziehen von zwei Phrasen in der Mitte und am Ende der Melodie – bei unveränderter Länge der gesamten Melodie – beeinträchtigt ist (vgl. Jones/Boltz 1989, S. 478-485). Tatsächlich traten die erwarteten Fehleinschätzungen beim Vergleich von ursprünglicher und variiertes Melodie in signifikantem Maße auf. Dieses experimentelle Ergebnis stimmt mit den Vorhersagen des sog. Kontrast-Modells von Jones und Boltz überein, das davon ausgeht, „[...] that time judgements are biased by the style of attunement failure (temporal contrast or mnemonic refinements) associated with a given attending mode (future-oriented or analytic)“ (Jones/Boltz 1989, S. 474).

Exkurs: Zeit in psychologischer und soziologischer Perspektive

Die Fähigkeit, aufgrund wahrgenommener zeitlicher Regelmäßigkeiten zukunftsorientierte Erwartungen über den regelmäßigen Fortgang von Ereignisfolgen aufzubauen, ist Grundlage jeglicher Synchronisierungsleistung. Für alle biologischen Systeme ist es lebensnotwendig, die Vielfalt ihrer Abläufe und Zyklen von Moment zu Moment untereinander abzustimmen und mit den zyklischen Veränderungen in ihrer natürlichen Umwelt zu synchronisieren: mit dem periodischen Wechsel von Tag und Nacht, den Gezeiten und Jahreszeiten. Vielfach greifen Organismen hierbei auf biochemische Prozesse, sog. biologischen Uhren, zurück (vgl. Birbaumer/Schmidt 2003, S. 542-550). Der biologische Tag des Menschen korrespondiert – genau wie auch der biologische Monat und das biologische Jahr – in etwa mit astronomischen Zyklen (Tag-Nacht und Gezeiten, Mondzyklus, Jahreszeiten).⁵ Im menschlichen Zeiterleben wird die unmittelbare Synchronisierung des biologischen Organismus mit den zeitlichen Strukturen der Umwelt durch ein Bewusstsein dieser Strukturen erweitert. Aufgrund dieser Überlegungen schlägt der Zeitpsychologe John A. Michon folgende Definition von Zeit vor: „Time is the conscious experiential product of the processes which allow the (human) organism to adaptively organize itself so that its behavior remains tuned to the sequential (order) relations in its environment“ (Michon 1985, S. 20).

Die bewusste Synchronisierung bezieht sich beim Menschen nicht mehr nur auf die vergleichsweise langen Zyklen in der natürlichen Umwelt, sondern ebenso auf die Koordination von Handlungen im sozialen Zusammenleben. In evolutionstheoretischer Sicht erlangt eine Spezies durch die

5 Allerdings wurde experimentell nachgewiesen, dass der Tag des Menschen (Wach-Schlafrythmus, Zyklus der Körpertemperatur) ca. 25 Stunden lang, also länger als der astronomische Tag ist; daher spricht man von einer *zirkadianen* Periodik (vgl. Birbaumer/Schmidt 2003, S. 542f).

Synchronisation der Handlungen ihrer Mitglieder eine Reihe von Entwicklungsvorteilen. Der Musikbiologe Björn Merker sieht in der Befähigung zur Synchronisation beim Singen und Trommeln einen Selektionsvorteil, der in der Entwicklungsgeschichte des Menschen eine zentrale Rolle gespielt haben könnte (vgl. Merker 2000). Obwohl einige Insekten ihre Rufe miteinander zu synchronisieren vermögen, fehlt diese Fähigkeit bei allen höheren Säugetieren, den Menschenaffen eingeschlossen. Gruppen von Männchen, die – wie beim Menschen – ihre Rufe untereinander synchronisieren und damit in der Lautstärke verstärken, so Merkers These, haben jedoch weit bessere Chancen, herumziehende Weibchen auf sich aufmerksam zu machen, und erlangen dadurch Fortpflanzungsvorteile. Weitere Beispiele für den funktionellen Einsatz von regelmäßigen Klangfolgen für die Handlungssynchronisation finden sich in der Menschheitsgeschichte zuhauf, so z.B. bei der Koordination von körperlicher Arbeit (Bücher 1909) und bei der gemeinsamen Fortbewegung (McNeill 1995).

Um Synchronisationsvorteile zu sichern, wurden im Laufe der Menschheitsgeschichte immer genauere Zeitmesseinrichtungen entwickelt, durch die soziales Handeln koordiniert werden kann (vgl. Wendorff 1985). Die Grundlage von Zeitmesseinrichtungen oder Uhren ist ein Vergleich der Dauern der zu messenden Vorgänge mit Zeitdauern von regelmäßigen zeitlichen Vorgängen (z.B. einer vollständigen Pendelbewegung), wobei die Anzahl dieser Zeiteinheiten durch einen Zähler notiert wird.⁶ Die verwendeten zeitlichen Bezugseinheiten stehen normalerweise in einem nur zufälligen Verhältnis zu den gemessenen Ereignisstrukturen. Die Kehrseite der immer präziseren Zeitmessung ist daher eine sinkende Natürlichkeit jener Handlungsstrukturen, die sich aufgrund sozialer Vereinbarungen den Zeitstrukturen der Messeinrichtungen anpassen müssen. Handlungszyklen lösen sich zunehmend von den biologisch vorgegebenen Zyklen wie etwa Schlafen und Wachen, Essensbeschaffung und -zubereitung, Säen und Ernten usw. Vielmehr wird in künstlichen Zeitplänen, z.B. Terminkalendern, festgelegt, wer wann was zu tun hat.

Nach Julius T. Fraser (1987, S. 190f) bestehen soziale Zeitpläne aus einer Folge von sozialen Gegenwart, deren Dauer durch die Möglichkeit gemeinsamen Handelns bestimmt wird. Je weniger Zeit benötigt wird, um Unkenntnis und Handlungsunfähigkeit zu überwinden, um so kürzer wird die Dauer der sozialen Gegenwart. Parallel zur Verfeinerung der Zeitmesseinrichtungen hat sich die Dauer der sozialen Gegenwart im Laufe der Neuzeit immer weiter verkürzt. So haben z.B. Verkehrsmittel und Nach-

6 Ein zweiter Uhrentyp (z.B. Feueruhren, Wasseruhren, Sanduhren, Räderuhren, Federuhren) basiert auf der Unterteilung eines geeichten Vorganges in gleiche Einheiten, die wiederum gezählt werden. Unterschiede zwischen verschiedenen Uhren bestehen in der Länge der zeitlichen Vorgänge, die zur Zeitmessung herangezogen werden, und in ihrer Exaktheit (vgl. Wendorff 1985).

richtentechnologien, die eine schnellere Informationsvermittlung ermöglichen und somit Phasen der Unkenntnis verkürzen, die soziale Gegenwart stark schrumpfen lassen.

Musikalisches Handeln ist als eine Form des sozialen Handelns in die Zusammenhänge der sozialen Organisation eingebunden. Zugleich vermag die Musikerfahrung jedoch, die alltägliche Zeitorientierung zumindest vorübergehend aufzuheben. Auf diese Weise entführt sie, so die Formulierung des Musikwissenschaftlers Hans Heinrich Eggebrecht, aus der Zeit – gemeint ist hier die soziale „Wirklichkeitszeit“ des alltäglichen Lebens:

„Diese Entführung wird bei der Musik noch dadurch verstärkt, dass das Hinhören auf das musikalische Spiel zugleich ein Hinhören ist auf ein Zeitspiel, eine gespielte Zeit: eine dem musikalischen Spiel eigene musikalische Zeit. [...] Die musikalische Zeit ist als Erlebniszeit weitgehend im Objekt, in der Musik selbst fixiert: Sie ist ‚komponiert‘ (‚zusammengesetzt‘), das heißt bis ins Detail durchorganisierte Zeit, ‚komponierte Erlebniszeit‘, die im Prozess der ästhetischen Identifikation die Wirklichkeitszeit zu sich hin auszulöschen vermag [...]“ (Eggebrecht 1996, S. 552).

Zeitliche Referenzebene, bevorzugter Tempobereich und Mikrotiming

Jones und Boltz gehen in ihrem Modell von einer zentralen Zeitebene aus („anchor or referent time level“, Jones/Boltz 1989, S. 470), mit der sich das Aufmerksamkeitssystem bevorzugt synchronisiert. Diese zentrale Zeitebene dient zugleich als Referenzebene, mit der kleinere und größere Zeitspannen verglichen werden. Die Annahme einer zeitlichen Referenzebene deckt sich mit experimentellen Befunden zum bevorzugten Tempobereich. Werden Versuchsteilnehmer aufgefordert, eine ihnen angenehme Schlagfolge zu klopfen, so ist diese Schlagfolge fast immer annähernd isochron und das gewählte „persönliche“ Tempo⁷ variiert zwischen Interonsetintervallen von ca. 330 und 880 ms, was 182 bzw. 68 Schlägen pro Minute (beats per minute, bpm)⁸ entspricht (vgl. Fraise 1982). Nur selten wird langsamer oder schneller geklopft (bis zu 1,4 s bzw. bis zu 200 ms pro Interonsetintervall). Ein repräsentativer Wert liegt nach Fraise bei 600 ms (100 bpm). Während die Tempounterschiede zwischen den Versuchsperso-

7 Fraise (1982, S. 154f) weist auf den Unterschied zwischen „persönlichem“ und „bevorzugtem“ Tempo hin; allerdings stimmen beide Tempobereiche weitgehend überein.

8 Eine Vergleichstabelle von Tempoangaben in den Maßeinheiten Schläge pro Minute (bpm), Dauer der Interonsetintervalle in Millisekunden (ms) und Häufigkeit pro Sekunde (Hz) sowie Formeln zum Umrechnen zwischen diesen Maßeinheiten finden sich im Anhang.

nen sehr groß sind, variieren sie bei wiederholten Messungen bei den einzelnen Versuchspersonen erstaunlich wenig (mit einer Reliabilität von $r = .75$ bis $r = .95$). In einem Versuch von Richard Parncutt, bei dem Versuchspersonen isochron zu vorgegebenen Klickmustern, die in verschiedenen Tempi vorgespielt wurden, mitklopfen sollten, war der bevorzugte Tempobereich etwas langsamer bei einem Mittelwert von 710 ms oder ca. 85 bpm; zwei Drittel der geklopfen Pulsfolgen lagen in einem Bereich von 420 bis 1190 ms (vgl. Parncutt 1994, S. 419).

Die Zeiteinheit, die mit dem musikalischen Tempo (von lat. tempus: Zeit) gemeint ist, wird seit dem 18. Jh. in Zeiteinheiten oder Schlägen pro Minuten (M.M. nach Mälzels Metronom oder beats per minute, bpm) ausgedrückt.⁹ Diese Zeiteinheit ist in der neuzeitlichen abendländischen Notenschrift zugleich der Bezugswert der Notensymbole und der Taktangaben (z.B. Viertelnote, Achtelnote oder halbe Note); alle anderen Notenwerte sind auf die Dauer dieses Zeitwertes bezogen. Der Bereich des musikalischen Tempos deckt sich weitgehend mit dem experimentell ermittelten bevorzugten Tempobereich. In einem Experiment von van Noorden und Moelants sollten Versuchsteilnehmer zu kurzen Musikausschnitten aus verschiedenen europäischen Musikepochen und aus populärer Musik mitklopfen. Es ergab sich ein durchschnittliches Klopftempo von knapp 500 ms (120 bpm) (van Noorden/Moelants 1999, S. 57ff). Köhlmann (1985) ermittelte in einem vergleichbaren Experiment (Mitklopfen zu Musikausschnitten verschiedener Stile einschließlich populärer Musik) ein Klopftempo, das in etwa doppelt so hoch war und bei ca. 250 ms (240 bpm) lag. Dies kann als Hinweis darauf gewertet werden, dass gerade beim Musikhören mehrere Referenzebenen möglich sind – etwa ein langsamerer Grundschlag und eine Grundschlagunterteilung im doppelten Tempo (ausgedrückt etwa in den Notenwerten einer Viertel- und einer Achtelnote).

Vermutlich hängt das bevorzugte Referenztempo von 500-700 ms oder 85-120 bpm mit den körperlichen Bewegungen des Menschen zusammen. Hierzu gehören vor allem die grundlegenden Arten der menschlichen Fortbewegung, das Gehen und Rennen, wiederholtes Nicken, aber auch Saugbewegungen, Kaubewegungen sowie der menschliche Herzschlag (vgl. Abschnitt 3.5). Aber auch die Grenzen des Kurzzeitgedächtnisses und die Dauer der psychologischen Gegenwart spielen in diesem Zusammenhang eine Rolle (vgl. 3.2). Nach Jones und Boltz (1989) wird das Tempo der Referenzebene sowohl vom Wahrnehmungskontext (Sinnesmodalität, Tempo) als auch von biologischen und psychologischen Faktoren (Alter, Grad der Erregung usw.) beeinflusst (Jones/Boltz 1989, S. 470).

9 Das subjektive Tempoempfinden hängt allerdings nicht nur mit dem metronomischen Tempo des Grundschlags, sondern zusätzlich mit der Ereignisdichte und der Art der musikalischen Bewegung zusammen.

Beim Wahrnehmen zeitlicher Strukturen wird freilich eine große Flexibilität und Dehnbarkeit der metronomischen oder isochronen Regelmäßigkeit akzeptiert. Rhythmische Strukturen – etwa in Aufführungen klassischer Musik oder beim Sprechen – werden auch dann als regelmäßig wahrgenommen, wenn das lokale Tempo starken Schwankungen unterliegt und wenn von einer strikten Regelmäßigkeit abgewichen wird. Jones (1986) versucht das Phänomen der Timing-Schwankungen in ein Zwei-Komponenten-Modell einzubauen. Dieses Modell berücksichtigt sowohl die oben beschriebenen Hierarchien der relativen Zeitverhältnisse als auch lokale Beschleunigungen und Verlangsamungen. Die hieraus resultierenden Geschwindigkeitsprofile („velocity or flow structure“) werden von Jones mit der Erfahrung von Bewegungen im Raum in Zusammenhang gebracht. Die Aufgabe des Musikers besteht dann darin, in der musikalischen Aufführung zwei Aspekte künstlerisch sinnvoll zusammen zu bringen: „[...] a performer ‚uses‘, in a generative fashion, both rhythmic and motional relative time information that are suggested by the composer. But because this information can raise conflicting timing goals, the successful artist is one who resolves the conflict in creative ways“ (Jones 1990, S. 229). Die Frage nach der Messbarkeit und möglichen Regelmäßigkeit sowie nach Funktionen und Wirkungen der kleinen Abweichungen von einer metronomisch exakten Periodizität, das sog. Mikrotiming, wird in Abschnitt 3.4 näher beleuchtet.

Zeitliche Regelmäßigkeiten und kognitive Schemata

Nach dem Modell der zukunftsorientierten Aufmerksamkeit führen zeitliche Regelmäßigkeiten von Ereignisfolgen zu Erwartungen bezüglich ihrer regelmäßigen Fortsetzung. Dies lässt sich als der besondere Fall eines kognitiven Skriptes oder Schemas beschreiben (Snyder 2000, S. 177ff). Ein kognitives Schema¹⁰ ist ganz allgemein eine mentale Struktur, durch welche die Wahrnehmung und kognitive Verarbeitung von Objekten, Situationen und Ereignissen auf gleichermaßen stabile wie flexible Weise ermöglicht wird. Nach Snyder gleicht ein Schema einem Bündel von Erwartungen, das es dem Menschen ermöglicht, sich durch eine Situation zu bewegen, ohne jedem Detail dieser Situation allzu große Aufmerksamkeit schenken zu müssen (vgl. Snyder 2000, S. 263). Aufgrund der schematischen Wissensbestände wird vielmehr die Aufmerksamkeit auf wenige besondere Eigenschaften von Ereignissen gelenkt, und ungenaue oder fehlende Beobachtungen werden durch die Standardwerte („default values“) ersetzt, welche das Schema vorgibt.

10 Zu Geschichte und Verwendung des Schemabegriffs in der Psychologie vgl. Anderson 2001, S. 153ff; zur Verwendung innerhalb der Musikpsychologie vgl. Snyder 2000, S. 95-105, und Louven 1998, S. 31-65.

Ein kognitives Schema lässt sich auch als ein Bündel von Leerstellen (sog. Slots) beschreiben, die durch feste Verknüpfungsvorschriften miteinander verbunden sind. Erklingen auditive Ereignisse nun mehrmals im gleichen Zeitabstand, so entsteht die Erwartung, dass zukünftig im gleichen Abstand weitere Klangereignisse auftreten werden. Die Leerstellen des Schemas, d.h. bestimmte Zeitpunkte, sind durch die Verknüpfungsvorschrift der Addition konstanter Zeitabstände miteinander verbunden. So entsteht ein schematisches Zeitraster, auf das sich der Hörer stützen kann, wenn er seine Aufmerksamkeit auf die kommenden Ereignisse richtet. Kleinere Abweichungen der Ereignisse von diesem Raster werden in einem gewissen Rahmen toleriert, größere Abweichungen führen jedoch zu einer Veränderung oder Anpassung des Schemas. Abschnitt 3.3 widmet sich dieser zentralen Fähigkeit der menschlichen Rhythmuswahrnehmung: der Bildung von Zeitrastern oder metrischen Erwartungsschemata aufgrund der wahrgenommenen Regelmäßigkeiten in der zeitlichen Struktur von Klangfolgen.

Die Erwartung von zeitlichen Regelmäßigkeiten ist nur ein Spezialfall von zahlreichen kognitiven Schemata bei der Musikwahrnehmung. Laut Thomas H. Stoffer besteht

„[...] die Funktion kognitiver Schemata [...] im Bereitstellen organisierter Gedächtnisinhalte, mit deren Hilfe die physikalischen Merkmalsstrukturen gehörter Musik bei ihrer Identifikation auf Korrespondenzen untersucht werden kann; zudem zielt die Verarbeitung auf eine systematische Einordnung neuer Inhalte in die bestehende Wissensstruktur, und bei Bedarf bewirkt sie deren Umorganisation (Assimilation/Akkomodation)“ (Stoffer 1998, Sp. 1859).

Das Konzept des kognitiven Schemas ist ein psychologisches Konstrukt, mit dem die Wahrnehmung und Klassifikation von Objekten und somit der gesamte Wissensbestand eines Individuums sowie dessen individuelle Konstruktion der Wirklichkeit beschrieben werden kann. Vergleichbare psychologische Konzepte sprechen von sog. Bezugssystemen und Prototypen (vgl. Stoffer 2005, S. 615 u. 632). Musikalische Bezugssysteme beziehen musikalische Merkmale auf einen aus dem vorangegangenen Kontext zu erschließenden musikalischen Vergleichsmaßstab. So werden etwa Tonhöhen in Bezug zu einem Grundton oder zur diatonischen Tonleiter wahrgenommen, oder eben die zeitliche Position von Klängen in Bezug auf ein periodisches Zeitraster. Ein musikalischer Prototyp ist eine idealtypische Repräsentation für vage definierte Kategorien von musikalischen Objekten oder Sachverhalten; Objekte lassen sich als unterschiedlich typische Vertreter einer Kategorie identifizieren. So lassen sich etwa bestimmte Schlagmuster als prototypisch für einen 4/4-Takt verstehen. Bei der Musikwahr-

nehmung spielen zahlreiche Bezugssysteme, Prototypen und Schemata eine wichtige Rolle: Tonale Bezugssysteme strukturieren die Wahrnehmung der einzelnen Tonhöhen, z.B. innerhalb der diatonischen Tonskala und des funktionsharmonischen Bezugssystems. Formschemata organisieren die Wahrnehmung größerer Einheiten und Abschnitte von Musikstücken. Klanggestalten werden zudem aufgrund von Ähnlichkeiten miteinander in Beziehung gesetzt – auch dann, wenn sie an voneinander entfernten Zeitpunkten eines Stückes auftauchen (vgl. Louven 1998, S. 58ff). Werden ähnliche rhythmische oder rhythmisch-melodische Gestalten in verschiedenen Stücken eines Stilbereichs verwendet, so können sich beim Hörer stil-spezifische Schemata herausbilden.

Dimensionen der Rhythmuserfahrung

Im Mittelpunkt der musikpsychologischen Forschung stehen vielfach die kognitiven Aspekte der Rhythmuserfahrung. Die Prozesse der kognitiven Verarbeitung von Klangfolgen stellen allerdings nur eine der Verhaltensdimensionen im Zusammenhang mit dem musikalischen Rhythmus dar. Nach Gabrielsson (1982, S. 159f) lassen sich drei Arten von Reaktionen auf musikalische Rhythmen („rhythm responses“) unterscheiden:

1. Physiologische Reaktionen, z.B. Veränderungen der Atem- oder Herzschlagfrequenz, der Muskelaktivität, des Hautwiderstandes oder bestimmter elektrischer Gehirnströme;
2. körperliches Verhalten, z.B. Klopfen des Fußes, Händeklatschen, Kopfbewegungen und Tanzen;
3. die musikalische Rhythmuserfahrung, die sowohl kognitive Aspekte als auch Bewegungsempfindungen und das emotionale Erleben umfasst. „Rhythm experience“, so Gabrielsson, „refers to such things as perceived grouping, accents, pulse and tempo, various motion characters (such as swinging, dancing, walking, etc.), feeling of excitement, tension, calmness, release etc.“ (Gabrielsson 1986, S. 140). Was umgangssprachlich als „rhythmisch“ bezeichnet wird, bezieht sich zwar in erster Linie auf das Vorhandensein einer regelmäßigen, leicht zu erfassenden zeitlichen Klangstruktur. Oftmals ist mit „rhythmisch“ jedoch zugleich die körperlich und emotional empfundene Intensität der ausgelösten Bewegungsempfindung gemeint.

Zu Beginn der 1970er Jahre versuchte Gabrielsson durch eine Reihe von explorativen Studien die Dimensionen der Rhythmuserfahrung näher zu bestimmen (Gabrielsson 1973a, b, c, d, vgl. auch Gabrielsson 1986, 1988). Er verwendete zwei Untersuchungsmethoden: Ähnlichkeitseinschätzungen zu zahlreichen paarweise dargebotenen Rhythmen (einstimmige und mehr-

stimmige Rhythmen, Tanzrhythmen einer synthetischen Rhythmusmaschine und Musikausschnitte) wertete er mit statistischen Methoden der multidimensionalen Skalierung aus; Adjektiv-Einschätzungen der Rhythmen unterzog er einer Faktorenanalyse. Zugleich spielten verbale Beschreibungen der Rhythmen durch die Versuchspersonen bei der Interpretation der Ergebnisse eine wichtige Rolle.

Die Anzahl und die Art der ermittelten Dimensionen oder Faktoren der Rhythmuserfahrung fiel in jedem Experiment unterschiedlich aus – abhängig von den verwendeten Rhythmen, aber auch in Abhängigkeit von der verwendeten Untersuchungsmethode, z.B. der Zusammensetzung der Adjektivliste. Darüber hinaus gab es große individuelle Unterschiede zwischen den Versuchsteilnehmern. Bei den Paarvergleichen ergaben sich Dimensionen, die sich vorwiegend auf strukturelle Eigenschaften der Rhythmen beziehen lassen. Bei den mehrstimmigen Rhythmen handelte es sich um die Dimensionen: Metrum, Schnelligkeit (metronomisches Tempo und Ereignisdichte), Einfachheit/Einheitlichkeit bzw. Komplexität/Variiertheit der Rhythmen, unterschiedliche „Basis-Pattern“ – ein Ausdruck der Versuchsteilnehmer, der sich auf Ähnlichkeiten in der Grundstruktur der Rhythmen bezieht. Bei den einstimmigen Rhythmen kamen weitere strukturelle Dimensionen ins Spiel: die genaue Timing-Struktur (kurze bzw. lange Tondauern), der Grad der Auftaktigkeit und der damit verbundenen Vorwärtsbewegung, der Grad der Synkopierung und der Akzentuierung des Taktanfanges.

Gabrielsson stellte allerdings fest, dass die Dimension des Metrums (3er- oder 4er-Periodizitäten) bei der Verwendung komplexerer Stimuli (z.B. begleiteter Melodien) in den Ähnlichkeitseinschätzungen seiner Versuchsteilnehmer bisweilen unwichtig wurde. Bei den polyrhythmischen Patterns aus der Rhythmusmaschine wurde die klangliche Gestaltung (helle Beckenklänge, Tonhöhen der Trommeln, dumpfer Klang der Basstrommel) zu einer eigenständigen Dimension. Bei den Musikausschnitten ergab sich eine weitere Dimension, die Gabrielsson als Bewegungscharakter interpretierte, da einige Versuchspersonen mit starker individueller Gewichtung in dieser Dimension davon sprachen, sie würden aufgrund vorgestellter Tanzbewegungen, die zu diesen Rhythmen passen, zu ihren Einschätzungen gelangen. Die Untersuchungen mit Adjektiv-Ratings erweiterte das Bild der Rhythmuserfahrung zusätzlich. Während einige der bereits genannten Dimensionen (Einfachheit/Einförmigkeit, Schnelligkeit sowie Instrumentierung – nun spezifiziert als klangliche Härte/Aggressivität gegenüber klanglicher Weichheit) wieder auftauchten, fehlten andere strukturelle Dimensionen völlig, so etwa die Metrums-Dimension. Zugleich wurden die Dimensionen des Bewegungscharakters differenzierter. Da die ermittelten Dimensionen sehr stark von der Art der verwendeten Adjektive abzuhängen schie-

nen, errechnete Gabrielsson Faktorenanalysen für drei unterschiedliche Gruppen von Adjektiven, die sich entweder auf kognitiv-strukturelle Aspekte, auf den Bewegungscharakter, oder aber auf emotionale Qualitäten beziehen. Für den kognitiv-strukturellen Aspekt ergaben sich nun nur noch zwei Dimensionen: Einförmigkeit/Einfachheit versus Variation/Komplexität und Grad der Klarheit oder Akzentuiertheit.¹¹ Für die Aspekte des Bewegungscharakters und der Emotionen ergaben sich dagegen jeweils mehrere Dimensionen. Bei den Adjektiven, die Bewegungscharaktere zum Inhalt haben, ermittelte Gabrielsson folgende Dimensionen: „[...] ‚uniform‘ versus ‚ununiform, limping, wavering‘; ‚solemn‘ vs. ‚swinging, pulsating‘; ‚rapid, agile, springy‘ vs. ‚crawling, stopping, hesitating‘; ‚floating, graceful, rocking, dancing‘ vs. ‚rugged, tense, stuttering, knocking, thumping‘; and ‚rocking, dancing‘ vs. ‚walking‘“ (Gabrielsson 1973c, S. 258). Ergänzt man die Dimension der Schnelligkeit und der Vorwärtsbewegung, so bestehen demnach in der musikalische Rhythmuserfahrung reichhaltige Möglichkeiten der Bewegungserfahrung. In den verbleibenden, emotional geprägten Adjektiven ergaben sich vier weitere Dimensionen, deren Pole sich durch die folgenden Adjektive beschreiben lassen: „[...] ‚playful, lively, vital‘ vs. ‚dull, heavy, restrained‘; ‚excited, violent, aggressive, tense, restless, hard‘ vs. ‚calm, soft, smoothed out, restrained‘; ‚free‘ vs. ‚mechanical, steady, monotonous‘; and ‚solemn‘ vs. ‚playful‘“ (Gabrielsson 1973c, S. 258).

Gabrielsson wendet sich zwar ausdrücklich gegen eine Verallgemeinerung seiner Ergebnisse. Andere Rhythmen und andere Adjektivvorgaben würden, so vermutet er, zu anderen Ergebnissen führen; außerdem sind die Unterschiede im Urteilsverhalten zwischen den verschiedenen Versuchspersonen teilweise beträchtlich. Dennoch belegen die Untersuchungsergebnisse eindeutig, dass die kognitiv-strukturellen Dimensionen durch motivationale und emotionale Dimensionen ergänzt werden müssen, will man der tatsächlich vorhandenen Vielschichtigkeit der Rhythmuserfahrung gerecht werden. In Abschnitt 3.5 sollen daher Überlegungen und empirische Untersuchungen zur Bewegungsdimension der Rhythmuserfahrung und deren Zusammenhang mit körperlichen und emotionalen Reaktionen diskutiert werden.

11 Gabrielsson mutmaßt, die geringe Zahl an kognitiven Dimensionen sei vielleicht auch darauf zurückzuführen, dass sich nur relativ wenige Adjektive direkt auf kognitiv-strukturelle Aspekte beziehen.

3.2 Grundlagen der auditiven Rhythmuswahrnehmung

Das menschliche Gehör ist ein äußerst differenziert arbeitendes Sinnesorgan, durch dessen Fähigkeiten eine Orientierung in der Umwelt erleichtert wird. In den akustischen Signalen, die in Form von schnellen Luftdruckschwankungen auf das Ohr treffen, überlagern sich fast immer eine Reihe von Klangereignisse unterschiedlicher Klangquellen. Es ist erstaunlich, mit welcher Genauigkeit das Gehör diese komplexen Schallsignale in auditive Einheiten, in verschiedene simultane, einander überlappende oder aufeinander folgende Klangereignisse und Klangströme aufzugliedern vermag. Die Gliederung der auditiven Szenerie in verschiedene Ereignisse und Ereignisströme ist Grundlage der räumlichen Orientierung in der akustischen Umwelt, der sprachlichen Kommunikation und der Musik. Die grundlegenden Mechanismen der auditiven Wahrnehmung sind auch bei der musikalischen Rhythmuswahrnehmung von zentraler Bedeutung, da hierdurch bestimmte Klangereignisse zu Segmenten, Gruppen oder Gestalten mit einem Anfang, einem Ende und einer inneren Struktur, zusammengefasst werden.

In der im vorherigen Abschnitt zitierten Definition bestimmt Gabrielson (1993, S. 97) Gruppierung, Akzentuierung, Regelmäßigkeit und die Grenzen der psychologischen Gegenwart als die zentralen Aspekte der musikalischen Rhythmuswahrnehmung. Unter Gruppierung versteht man die Gliederung von Klangfolgen in Klanggestalten, die als Ganzheiten wahrgenommen werden und gegenüber dem, was davor und danach erklingt, abgegrenzt sind. Voraussetzung hierfür ist, dass die auditive Szenerie aufgrund bestimmter Kriterien in verschiedene Klangströme unterteilt wird, denen die einzelnen Klangereignisse zugeordnet werden. Zur Struktur einer Klanggestalt trägt nicht nur die zeitliche Abfolge der sie konstituierenden Elemente bei. Vielmehr erhalten einzelne Klangelemente im Vergleich zu den sie umgebenden Elementen aufgrund verschiedener Faktoren ein größeres Gewicht, sie werden als akzentuiert wahrgenommen. Regelmäßigkeiten der zeitlichen Struktur von Rhythmen erleichtern die Wahrnehmung der Klanggestalten, indem sie den Aufbau von Erwartungen und eine Synchronisation der Aufmerksamkeit mit den Klangstrukturen ermöglichen (vgl. hierzu Abschnitt 3.3). Schließlich lassen sich musikalische Rhythmen von der musikalischen Form durch ihre kleinere, überschaubare zeitliche Dimension abgrenzen. Bei dieser Abgrenzung kann auf das Konzept der sog. psychologischen Gegenwart zurückgegriffen werden.

Im Folgenden werden die wichtigsten Ergebnisse aus der musikpsychologischen und psychoakustischen Forschungsliteratur zur zeitlichen Wahrnehmung und Gruppierung von Klängen, zur psychologischen Gegenwart, zur Akzentwahrnehmung sowie zu Wechselwirkungen zwischen Akzen-

tuierung und Gruppierung zusammengefasst. Obwohl die einfachen Klangfolgen, die vielen dieser musikpsychologischen Untersuchungen zugrunde liegen, keinen Musikstücken entnommen wurden, sind die hierbei gewonnenen Erkenntnisse auch für die Wahrnehmung von Rhythmen in musikalischen Kontexten relevant, da sie grundlegende und allgemein gültige Wahrnehmungsmechanismen beschreiben.

Die zeitliche Wahrnehmung von Klangereignissen

Obwohl der Strom der im Gehör eintreffenden Schallinformationen kontinuierlich ist, werden innerhalb dieser auditiven Szenerie („auditory scene“, Bregman 1990) zumeist klar gegeneinander abgrenzbare Ereignisse wahrgenommen. Die Identifikation und Abgrenzung der Klangereignisse erfolgt aufgrund von Pausen zwischen den Ereignissen, aber auch aufgrund von Lautstärkeunterschieden und Eigenheiten der spektralen Energieverteilung. Nach psychoakustischen Untersuchungen, die bei Zwicker und Fastl (1999, S. 274f) zusammengefasst sind, muss das Lautstärkemaximum eines Ereignisses mindestens 120 ms hinter dem Maximum des vorherigen Ereignisses liegen, um als neues Ereignis wahrgenommen zu werden. Näher beieinander liegende Ereignisse verschmelzen dagegen zu einem gemeinsamen Klangereignis. Außerdem muss die Lautstärke der einzelnen Ereignisse mindestens 43 % der innerhalb des klanglichen Kontextes, z.B. einer Phrase, vorherrschenden maximalen Lautstärke betragen; die Lautstärkedifferenz, die zur Unterscheidung zweier Klangereignisse erforderlich ist, muss mindestens 12 % des Lautstärkemaximums betragen (Abb. 3).

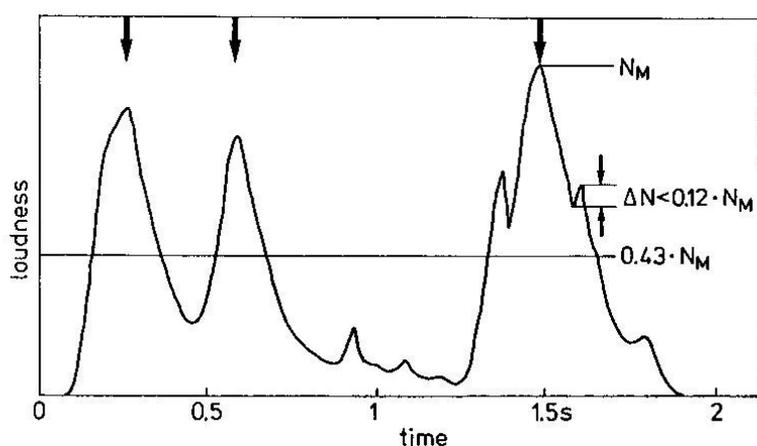


Abb. 3: Der Lautstärkeverlauf in einer akustischen Szenerie; N_M ist das lokale Lautstärkemaximum. Nur bei den drei mit Pfeilen gekennzeichneten Lautstärkemaxima werden separate Klangereignisse wahrgenommen (aus Zwicker/Fastl 1999, S. 275).

Justin London (2002, S. 535) führt weitere Belege aus der psychoakustischen Forschungsliteratur an, die alle dafür sprechen, dass die untere zeitliche Grenze für die Wahrnehmung und Produktion von rhythmisch relevanten Klangdauern bzw. Interonsetintervallen ca. 100-120 ms beträgt.¹²

Ein weiterer Kennwert für die zeitliche Auflösungsgenauigkeit des Gehörs ist die kleinste noch wahrnehmbare zeitliche Deplatziierung eines Klangereignisses innerhalb einer isochronen Klangfolge (vgl. im Überblick Friberg/Sundberg 1995). Friberg und Sundberg ermittelten diese Größe in Abhängigkeit vom Tempo der Klangfolge (60 bis 600 bpm). Dabei sollte in verschiedenen Tempi jeweils das vierte Element einer Folge von sechs gitarrenartigen Klängen (C⁴ mit einer extrem kurzen Einschwingzeit von 0,5 ms) zeitlich möglichst exakt justiert werden. Die durchschnittliche Genauigkeit dieser Einstellungen betrug bei Tempi unter 250 bpm jeweils ca. 2,5 % der Interonsetintervalle zwischen den Ereignissen; bei schnelleren Tempi jedoch konstant ca. sechs Millisekunden.

Der kleinste wahrnehmbare Abstand zwischen zwei auditiven Ereignissen hängt allerdings stark von deren akustischen Eigenschaften ab. Sehr kurze klickartige Impulse werden bereits bei einem Abstand von nur 2 ms als nicht mehr gleichzeitig erlebt, sondern voneinander unterschieden.¹³ Die Reihenfolge zwischen zwei Klicks kann allerdings erst dann verlässlich bestimmt werden, wenn zwischen ihnen ein Abstand von mindestens 20 ms besteht (Hirsh 1959, Warren 1993).¹⁴ Werden in Hörexperimenten statt kurzer, klickartiger Impulse etwas längere Klangereignisse (sog. „tone

12 Nach van Norden und Moelants (1999, S. 54) muss das Interonsetintervall zwischen zwei Klangereignissen allerdings mindestens 200 ms betragen, damit beide Klänge „being perceived as fully independent, without influencing each others loudness“. Neben Belegen aus psychoakustischen Experimenten führen sie Beobachtungen an, nach denen Versuchspersonen unterhalb dieser Temposchwelle Schwierigkeiten haben, kontrolliert isochrone Ereignisfolgen zu klopfen (vgl. hierzu auch London 2002, S. 536).

13 Diese sog. Gleichzeitigkeitsschwelle ergibt sich aufgrund von Vorgängen bei der Reiztransduktion im Innenohr sowie aufgrund der Übertragungseigenschaften der Hörnerven. Die Gleichzeitigkeitsschwelle ist bei den verschiedenen Sinnesmodalitäten unterschiedlich groß (vgl. Grüsser 1989, S. 114ff), bei auditiven Reizen jedoch mit Abstand am kleinsten.

14 Die Schwelle der Reihenfolge ist vermutlich von Übertragungseigenschaften der Nervenzellen nicht der Hörbahn, sondern des Gehirns abhängig, da festgestellt wurde, dass sich diese Schwelle bei bestimmten Gehirnverletzungen erhöht (Snyder 2000, S. 26). Folgen isochrone Klangimpulse mit einer Frequenz größer als 20 Hz (also maximal 50 ms Abstand zwischen den Impulsen) aufeinander, so lösen sie Tonhöhenempfindungen aus. Bei einer allmählichen Beschleunigung des Tempos werden zunächst Folgen von Klicks und kurzen Pausen gehört. In einem kritischen Übergangsbereich ist ein raues Geräusch hörbar, das ab einer Frequenz von 20 Hz (also 50 ms Abstand zwischen den einzelnen Impulsen) in einen tiefen Ton übergeht. Diese Tonfusionsschwelle beruht laut Grüsser (1989) auf Eigenheiten der Hörnerven. Beim Sehen spricht man von einer Flimmerfusionsschwelle, die bei 50 Hz oder 20 ms Abstand zwischen Licht und Dunkelheit liegt (vgl. Grüsser 1989, S. 96).

bursts“) verwendet, so zeigt sich, dass der empfundene Anfangspunkt nicht immer mit dem physikalischen Beginn des Schallereignisses übereinstimmt. Vielmehr wird der Beginn zumeist wenige Millisekunden später wahrgenommen. Bei Justierungsversuchen, bei denen verschiedenartige Klänge innerhalb einer isochronen Klickfolge so platziert werden sollten, dass die Abstände zwischen den Ereignissen als genau gleich lang oder isochron empfunden werden, ergaben sich die in Abbildung 4 dargestellten Resultate.

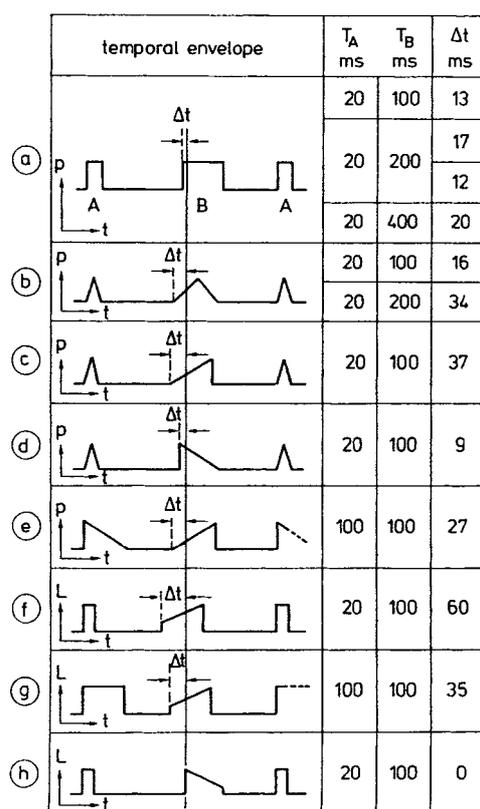


Abb. 4: Unterschiede Δt zwischen physikalischen und wahrgenommenen Anfängen von Klangereignissen mit unterschiedlichen Lautstärkehüllkurven; T_A und T_B bezieht sich auf die Dauern der Klangereignisse (aus Zwicker und Fastl 1999, S. 272).

Joos Vos und Rudolf Rasch (1981, 1982) kamen bei psychoakustischen Experimenten zu dem Ergebnis, dass der wahrgenommene Anfangspunkt eines Tones in einem moderaten Lautstärkebereich generell etwa 6 bis 15 dB unterhalb des größten Amplitudenausfalls liegt: „(a) the perceptual onsets of successively presented tones can be defined as the times at which the envelopes pass a relative threshold level; (b) within a range of 20 to 70

dB above masked or absolute threshold, the relative threshold for the perceptual onset lies about 6 to 15 dB below the maximum level of the tones“ (Vos/Rasch 1982, S. 316).¹⁵

Bei allen Schallereignissen, deren Lautstärkemaxima nicht schlagartig einsetzen, sondern allmählich anwachsen, liegt der wahrgenommene Anfangspunkt demnach zwischen dem Beginn des Einschwingvorgangs und dem Amplitudenmaximum. Einschwingvorgänge von Musikinstrumenten erstrecken sich abhängig von der Art der Klangerzeugung über 50 Millisekunden und mehr. Wie Rasch (1979, 1988) betont, ist bei Musikinstrumenten aufgrund der langen Einschwingzeiten ein weit größerer Abstand zwischen den Einsätzen der einzelnen Stimmen als bei kurzen Klicks möglich, bevor eine Asynchronizität der Einsätze empfunden wird. Wie Messungen an Aufnahmen klassischer Musik zeigten, sind hier Asynchronizitäten von 30-50 ms (bei langsamen Sätzen noch darüber hinaus) zwischen den Einsätzen mehrerer Stimmen durchaus üblich und werden nicht als störend oder ungleichzeitig empfunden (Rasch 1979, S. 129). Die Asynchronizität der Toneinsätze erleichtert in vielen Fällen die Wahrnehmung der verschiedenen Stimmen, da auf diese Weise Verdeckungen der Instrumentalklänge untereinander vermieden werden.

Gliederung in Klangströme

Komplexe Schallinformationen, wie sie in der menschlichen Umwelt, beim Sprechen und in der Musik auftreten, werden vom Gehör als einzelne, gegeneinander abgrenzbare auditive Ereignisse wahrgenommen. Zugleich werden die Ereignisse in Klangströme aufgegliedert, die parallel erklingen und zumeist unterschiedlichen Schallquellen zugeschrieben werden. Hierdurch lassen sich verschiedene Umweltgeräusche gegeneinander abgrenzen, es wird möglich, in lauten Umgebungen einem Sprecher zuzuhören oder die Linien verschiedener Musikinstrumente voneinander zu unterscheiden. Seit den 1970er Jahren wurden die grundlegenden Prozesse der Aufgliederung oder Aufspaltung einer auditiven Szenerie in Klangströme („auditory stream segregation“) und die klanglichen Faktoren, die diesen Gliederungsprozessen zugrunde liegen, Gegenstand ausgiebiger Forschungen. Im Folgenden sollen die zentralen Ergebnisse dieser Untersuchungen dargestellt werden (vgl. van Noorden 1975, Bregman 1990, 1993, Handel 1989, S. 189-217, Goldstein 2002, S. 453ff).

15 Ashley (2002, S. 315) schlägt unter Hinweis auf Messmethoden der Phonetik vor, den Beginn eines Klages mit dem Beginn von dessen stationären Teil gleichzusetzen. Dies hätte jedoch zur Folge, dass sich die Anfänge von perkussiven und geräuschhaften Klängen mangels eines stationären Teils nicht bestimmen ließen.

In der menschlichen Umwelt ist der wichtigste Faktor, der zur Gliederung der auditiven Szenerie in verschiedene Klangströme beiträgt, die räumliche Ortung aufgrund von Laufzeit- und Intensitätsunterschieden des Schalls, die zwischen den beiden Ohren auftreten. Schallereignisse, die aus einer bestimmten Richtung des Raumes und daher mit konstanten Laufzeitunterschieden auf beide Ohren treffen, stammen in der Regel von ein und derselben Schallquelle. Bei fehlenden oder nur schwach ausgeprägten Rauminformationen stützt sich die Gliederung in Ereignisströme zudem auf Ähnlichkeiten im Klangspektrum der Klänge. Zeitliche Faktoren spielen bei der Gliederung in Klangströme ebenfalls eine Rolle. So werden Klänge, die in rascher Folge nacheinander erklingen, derselben Klangquelle und demselben Klangstrom zugeordnet. Beginnen oder enden mehrere Klänge oder Klangkomponenten gleichzeitig, so ist dies ein Indiz dafür, dass sie von derselben Klangquelle stammen.

Eine besondere Fähigkeit des Gehörs besteht in der Rekonstruktion von Klängen und Klangströmen, die durch einen anderen, lautereren Klangstrom verdeckt werden. Ähnlich wie bei der Anordnung visueller Objekte im Raum, bei der die Objekte einander teilweise verdecken und dennoch in der visuellen Wahrnehmung zu vollständigen Gestalten ergänzt werden können, werden Klangströme auch dann als kontinuierlich erlebt und in der Wahrnehmung ergänzt, wenn sie zeitweise durch lautere Klänge oder Geräusche verdeckt werden. Bereits kleine Verzögerungen zwischen dem Einsatz des maskierten und des maskierenden Klanges reichen aus, um auch den maskierten Klang durchgängig wahrnehmbar werden zu lassen.

Bei Klangfolgen mit einheitlicher Klangfarbe, Lautstärke, jedoch unterschiedlichen Tonhöhen kommt es in rascherem Tempo zur spontanen Bildung zweier (oder mehrerer¹⁶) Klangströme. Dabei richtet sich die Aufmerksamkeit auf einen der Klangströme, der sich als Figur vor dem Hintergrund der restlichen Klänge abhebt. Allerdings ist bei der künstlichen Mehrstimmigkeit – wie bei aller Mehrstimmigkeit – eine willentliche Verlagerung der Aufmerksamkeit auf andere Klangströme und somit eine Umkehrung des Figur-Grund-Verhältnisses möglich. Außerdem kann die Aufmerksamkeit in bestimmten Grenzen zwischen den Klangströmen geteilt werden, sodass beiden Klangströmen simultan Aufmerksamkeit geschenkt wird.

Das Phänomen der Aufspaltung einer Klangfolge in zwei Klangströme aufgrund von Tonhöhen sprüngen wird in verschiedenen Musikkulturen zur

16 Auch wenn eine Aufspaltung in mehr als drei Klangströme unwahrscheinlich ist und bislang experimentell nicht untersucht wurde, hält sie Bregman (1990, S. 464f) durchaus für möglich.

Schaffung einer sog. künstlichen Mehrstimmigkeit genutzt¹⁷, so etwa in Johann Sebastian Bachs Partita für Violine Solo (Abb. 5).



Abb. 5: Ausschnitt aus dem Präludium der Partita für Violine Solo, No. 3 in E-Dur, BWV 1006, von Johann Sebastian Bach (T. 1-12). In den Takten 3, 5 und 9ff wird die Linie aufgrund der Intervallsprünge als zweistimmig wahrgenommen.

Bei der künstlichen Mehrstimmigkeit ist die Wahrnehmung zweier Klangströme sowohl vom Tempo der Tonfolge als auch von der Größe der Intervalle zwischen den einzelnen Tönen abhängig. Van Noorden (1975, S. 2ff) stellte experimentell fest, dass bei Klangfolgen mit zwei alternierenden Tonhöhen erst ab einem Tempo von ca. 300 bpm (oder fünf Tönen in der Sekunde) und bei Intervallen, die größer als eine Oktave sind, die Spaltung in zwei Klangströme automatisch erfolgt; bei schnelleren Tempi (> 600 bpm) sinkt die hierzu erforderliche Intervallgröße auf ca. eine kleine Terz. In einem Zwischenbereich mit einem Tempo der Tonfolge von weniger als ca. 300 bpm kann jedoch der Hörer in Abhängigkeit von der Intervallgröße seine Aufmerksamkeit bewusst so lenken, dass er entweder einen kohärenten Klangstrom oder aber zwei Klangströme wahrnimmt (vgl. Abb. 6).

Eine Aufgliederung in zwei Klangströme kann auch aufgrund anderer klanglicher Eigenschaften der alternierenden Töne erfolgen, so etwa aufgrund unterschiedlicher Klangfarben und unterschiedlicher Lautstärken – jeweils in Abhängigkeit von der Geschwindigkeit der Klangfolge (Bregman 1990, S. 478ff). Die Mechanismen der Gliederung in Klangströme wirken auch dann, wenn Klangsequenzen aus mehr als zwei alternierenden Klangtypen bestehen. Wiederum werden aber zumeist nur zwei Klangströme –

17 Van Noorden (1975, S. 2ff) spricht von „rhythmic fission“; vgl. auch Bregman 1990, S. 461ff, und Wegner 1993.

Figur und Hintergrund – wahrgenommen, in denen ähnliche Klänge zusammengefasst werden.

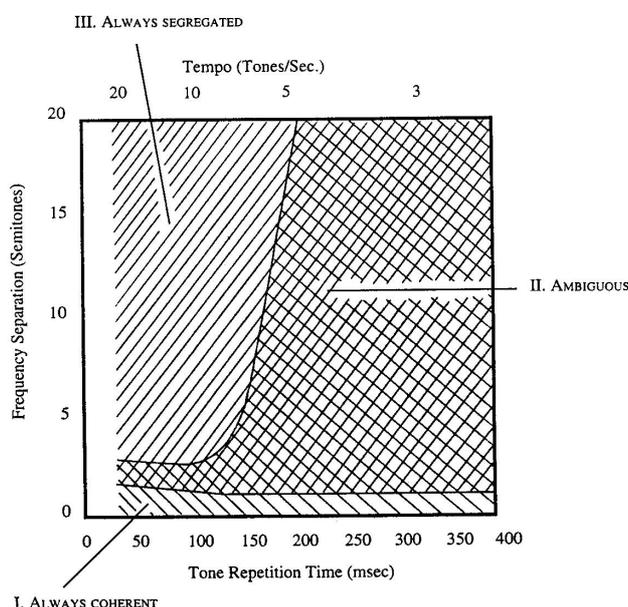


Abb. 6: Wahrnehmung zweier Klangströme bei einer Folge zweier alternierender Töne in Abhängigkeit vom Tempo (horizontale Achse) und der Intervallgröße (vertikale Achse). Im Bereich I wird keine Aufspaltung, im Bereich III eine automatische Aufspaltung in zwei Klangströme vorgenommen; im Bereich II ist abhängig von der Hörstrategie sowohl eine Aufgliederung als auch eine kohärente Wahrnehmung möglich (nach van Noorden 1975, aus Wegner 1993, S. 211).

Gruppierung von Klangfolgen

Die menschliche Wahrnehmung tendiert dazu, die auditiven Ereignisfolgen innerhalb eines Klangstromes in Einheiten (Segmente, Gruppierungen oder Gestalten) zu unterteilen, die eine Art organisches Ganzes darstellen und als gegenwärtig wahrgenommen werden. Grundlegend hierfür sind vor allem zwei Gestaltbildungsprinzipien: Das Prinzip der zeitlichen Nähe und das Prinzip der Ähnlichkeit. Klänge, die nahe beieinander liegen, werden zu Gestalten oder Gruppierungen zusammengefasst; bei vergleichsweise großen Abständen zwischen zwei aufeinander folgenden Klangereignissen entsteht dagegen eine Gruppierungsgrenze. Das Prinzip der Ähnlichkeit kann sich auf mehrere Merkmale beziehen: auf Ähnlichkeiten in der Tonhöhe, der Klangfarbe, der Lautstärke, der Tondauer, der Artikulation oder der räumlichen Position der Schallquelle. Aufgrund dieser Ähnlichkeiten

werden aufeinander folgende Klangereignisse zu Gestalten zusammengefasst; aufgrund von Veränderungen in einem oder mehreren dieser Parameter entstehen Gruppierungsgrenzen. Allerdings ist das Prinzip der zeitlichen Nähe für die Gestaltwahrnehmung wichtiger als Ähnlichkeitskriterien der Tonhöhe oder Klangfarbe. Während bei der regelmäßigen Tonfolge in Abb. 7a die Gruppierungsgrenzen aufgrund der Tonhöhe wahrgenommen werden, ist die zeitliche Nähe ausschlaggebend, sobald zwischen den Tönen Pausen eingefügt werden (Abb. 7b).



Abb. 7: Gruppierung einer Tonfolge aufgrund a) der Ähnlichkeit der Tonhöhe und b) der zeitlichen Nähe.

Erstaunlicherweise erfolgt eine Gruppierung selbst dann, wenn die Klangfolgen aus völlig identischen Ereignissen bestehen, die in einem identischen zeitlichen Abstand, also isochron erklingen (Abb. 8).

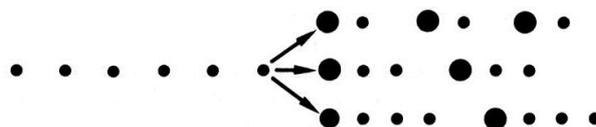


Abb. 8: „Subjektive Rhythmisierung“ einer isochronen Klangfolge (aus Handel 1989, S. 387).

Die experimentellen Untersuchungen zur Segmentierung isochroner Klangfolgen, der „subjektiven Rhythmisierung“, führten zu folgenden Ergebnissen:¹⁸ Es werden vorwiegend Gruppierungen von zwei oder drei, seltener von vier Elementen wahrgenommen. Tempo-Effekte sind bei den spontanen Gruppierungsprozessen relativ gering, bei höherem Tempo vergrößert

18 Bereits Wilhelm Wundt (1911) entdeckte einen spontanen Gruppierungsmechanismus einheitlicher, isochroner Klangfolgen in Gruppen von mehreren Klängen; weitere Untersuchungen hierzu wurden von Bolton (1894) und Woodrow (1909) durchgeführt. Die Kenntnis des Phänomens ist natürlich älter, vgl. z.B. Sulzer 1794, Bd. 4, S. 92-96. Die Bezeichnung „subjektive Rhythmisierung“ wird von Handel (1989, S. 386) mit dem Hinweis in Frage gestellt, dass wahrgenommene Rhythmen immer „subjektiv“ sind, da sie sich nicht eindeutig und zwingend aus der physikalischen Struktur der Schallereignisse ergeben.

sich die Anzahl der Elemente in den Gruppierungen nur geringfügig. Das erste Element jeder Gruppe erscheint als akzentuiert, z.B. als lauter. Die Abstände zwischen Elementen innerhalb der Gruppen erscheinen kürzer als zwischen verschiedenen Gruppen.

Bei der Musikwahrnehmung lässt sich eine spontane Gruppierung in kleinere Einheiten („primitive Gruppierung“) von einer Gruppierung in Phrasen unterscheiden, deren maximale Länge der Dauer der psychologischen Gegenwart entspricht (Snyder 2000, S. 31-39).¹⁹ Die elementaren oder „primitiven“ Gruppierungen werden innerhalb der Dauer der psychologischen Gegenwart zu längeren Einheiten, den musikalischen Phrasen, zusammengefasst. Die Grenzen von Phrasen und größeren musikalischen Einheiten entstehen durch simultane Veränderungen mehrerer klanglicher Parametern, z.B. Klangdauer bzw. Interonsetintervall, Tonhöhe und Lautstärke, aber auch durch gelernte Regeln der Phrasenbildung.

Psychologische Gegenwart und Arbeitsgedächtnis

Das Gehör fasst die Klangereignisse innerhalb eines Klangstromes zu Einheiten oder Gruppen zusammen, die als gegenwärtig erlebt werden. Wenn eine geschlossene rhythmische Gestalt immer weiter verlangsamt wird, so fällt sie in Einzelklänge auseinander, sobald der Abstand zwischen zwei aufeinander folgenden Klangereignissen größer als anderthalb bis zwei Sekunden wird. Dieses Phänomen steht in Zusammenhang mit der sog. psychologischen Gegenwart²⁰, das bereits 1890 von William James in seinen *Principles of Psychology* beschrieben wurde: „The practically cognized present is no knife-edge, but a saddle-back with a certain breadth of its own on which we sit perched, and from which we look into two directions in time“ (zit. nach Dowling/Harwood 1986, S. 179).

Innerhalb der Zeitspanne der psychologischen Gegenwart sind alle Bewusstseinsinhalte direkt verfügbar, sodass keinerlei Erinnerungsleistungen erforderlich sind (Michon 1978, S. 92). Das Zeitfenster der Gegenwart verschiebt sich nicht kontinuierlich auf einer abstrakten Zeitachse. Wir hören das Ticken einer Uhr nicht in einem Augenblick als „tick-tack“ und im nächsten als „tack-tick“, sondern als eine Wiederholung von identischen Einheiten („tick-tacks“ oder „tack-ticks“). Das Gegenwartsfenster verschiebt sich jedoch mit Überlappungen, sodass der Eindruck eines kontinu-

19 Ein Computermodell der Gruppierung von Klangfolgen aufgrund von Tonhöhen- und Lautstärkeunterschieden der Klangereignisse hat Neil P. Todd entworfen (Todd 1994, vgl. Clarke, 1999, S. 480f).

20 Für die psychologische Gegenwart finden sich in der Literatur auch die Bezeichnungen „psychische Gegenwart“, „bewusste Gegenwart“ (Michon 1978), „Wahrnehmungsgegenwart“ („perceptual present“, Fraise 1978) und „psychische Präsenzzeit“ (Stern 1897).

ierlich verstreichenden Zeitflusses, eines Bewusstseinstroms entsteht. Hierdurch wird sowohl eine Orientierung in die unmittelbare Vergangenheit als auch eine Orientierung in die Zukunft (Antizipieren) ermöglicht (Michon 1978, S. 93).

Die exakte Dauer der psychologischen Gegenwart lässt sich nicht eindeutig bestimmen; in der psychologischen Literatur finden sich Angaben zwischen anderthalb und drei Sekunden (vgl. Michon 1978, Block 1990, S. 5ff). Nach Richard A. Block (1990) besitzt die psychologische Gegenwart eine Obergrenze der zeitlichen Kapazität, die bei ca. fünf Sekunden liegt, wobei diese Grenze von der inneren Struktur der wahrgenommenen Ereignisse abhängt. Auch Michon betont, dass die zeitliche Ausdehnung der psychologischen Gegenwart nicht nur von psychischen Eigenschaften abhängt, sondern ebenso von der zeitlichen Struktur und Segmentierung der wahrgenommenen Ereignisfolgen (Michon 1978, S. 89). Zur Dauer der psychologischen Gegenwart bei der Musikwahrnehmung gibt es in der Literatur ebenfalls widersprüchliche Angaben (vgl. Berz 1995, S. 354). Manche Autoren beziehen die Obergrenze auf die Anzahl der erklingenden Töne und kommen hier auf Folgen von sieben bis fünfzehn Tönen.

Vor dem Hintergrund von Jones' Unterscheidung zwischen einem analytischen und einem zukunftsorientierten Aufmerksamkeitsmodus (vgl. 3.1) ist es nicht verwunderlich, dass experimentell gewonnene Angaben zur Dauer der psychologischen Gegenwart beträchtlich variieren – abhängig von der Bestimmungsmethode und den Bewusstseinsinhalten innerhalb dieser Zeitspanne. Eine zeitliche Kapazität von ca. anderthalb Sekunden ist sicherlich einem detailorientierten, analytischen Aufmerksamkeitsmodus angemessen. In musikalischen Kontexten mit einer kohärenten, stark hierarchischen Organisation der zeitlichen Struktur treten jedoch vermutlich auch längere Ausdehnungen der psychologischen Gegenwart auf.²¹ Dabei gilt es jedoch zu berücksichtigen, dass nach Cowan (1984) der unmittelbare auditorische Sinnesspeicher nur akustische Reize, die maximal 20 Sekunden zurückliegen, speichern kann. Block bringt die Dauer der Gegenwart des Wahrnehmens und Erlebens ebenfalls mit der Kapazität der menschlichen Informationsverarbeitung in Verbindung: „The perception and production of rhythm, as in a piece of music or in a series of coordinated movements, depends on structural and dynamic properties of information-

21 So schreibt etwa der Psychologe Edward G. Boring 1933 in *The Physical Dimensions of Consciousness*: „[...] conscious present can certainly include a rhythmic grouping that occupies a second or a second and a half, and that with somewhat less ‚immediacy‘ [...] may extend to include a rhythm of a quarter or perhaps even half a minute“ (zit. nach Block 1990, S. 5). Kauffman und Carlsen (nach Berz 1995) nennen sogar eine zeitliche Ausdehnung der psychologischen Gegenwart bei der Musikwahrnehmung von über drei Minuten.

processing systems underlying the psychological present [...]“ (Block 1990, S. 6).

Nach dem verbreiteten Gedächtnismodell von Atkinson und Shiffrin (1968) wird bei der Informationsverarbeitung zwischen Kurzzeit- und Langzeitgedächtnis unterschieden. Da es als erwiesen gilt, dass das Kurzzeitgedächtnis über Komponenten sowohl der Speicherung als auch der Verarbeitung von Informationen verfügt, sollte nach Berz (1995) die Bezeichnung Arbeitsgedächtnis vorgezogen werden.²² Berz nennt Hinweise darauf, dass für verschiedene Sinnesmodalitäten unterschiedliche Arbeitsgedächtnisse existieren. So wurde experimentell beobachtet, dass instrumentale Hintergrundmusik die Leistungen der sprachlichen Verarbeitung nicht beeinträchtigt, wohl aber Vokalmusik oder im Hintergrund abgespielte Sprachaufnahmen (Salame/Baddeley 1989). Berz postuliert aufgrund dieser Beobachtung die Existenz eines musikalischen Arbeitsgedächtnisses neben einem sprachlich-phonologischen Arbeitsgedächtnis.

Berz weist außerdem darauf hin, dass sich die Kapazität des musikalischen Arbeitsgedächtnisses durch Bekanntheit des musikalischen Materials und durch Verarbeitungsstrategien des Langzeitgedächtnisses vergrößern lässt. So wächst die momentane Erinnerungsleistung für Melodien durch die wiederholte Präsentation der Stimuli stark an (Pembroke 1986). Die Vermutung liegt nahe, dass sich im Langzeitgedächtnis sehr schnell interne Repräsentationen der Melodien aufbauen. Die bessere Erinnerungsleistung von musikalisch trainierten Hörern kann damit in Verbindung gebracht werden, dass diese von einem reichhaltigeren Erinnerungsschatz an melodischen Schemata profitieren können (Berz 1995, S. 354). Wissensstrukturen des Langzeitgedächtnisses, seien es nun konkrete musikalische Erinnerungen oder aber schematisch aufgebaute Erwartungen über einen regelmäßigen zeitlichen Verlauf, haben demnach direkten Einfluss auf die musikalische Wahrnehmung, auf die Kapazität des musikalischen Arbeitsgedächtnisses und damit auf die Dauer der psychologischen Gegenwart.

Akzentuierung von Klängen

Innerhalb von Klangfolgen erhalten manche Klangereignisse ein größeres Gewicht als andere, sie erscheinen im Vergleich zu den sie umgebenden Klängen als akzentuiert. Die Akzentuierung von Klängen hat u.a. zur Folge, dass sie besser wahrgenommen und leichter erinnert werden als weniger akzentuierte Ereignisse. Veränderungen der akzentuierten Ereignisse werden in stärkerem Maße bemerkt und haben einen größeren Einfluss auf die wahrgenommene Struktur einer Klanggestalt als Veränderungen weniger akzentuierter Klänge. Es gibt eine Reihe von Möglichkeiten, Klangereig-

22 Berz beruft sich hierbei auf den Gedächtnispsychologen Alan Baddeley (1990).

nisse zu akzentuieren. In zahlreichen Hörexperimenten wurde versucht, mögliche Akzentuierungsarten zu ermitteln (vgl. Monahan/Caterette 1985, Monahan 1993, Jones 1993, Müllensiefen 2004). Hier die wichtigsten Faktoren, die zur Akzentuierung eines Klangereignisses beitragen:

1. Ein lauterere Ereignis erscheint akzentuiert.
2. Ein längeres Ereignis erscheint akzentuiert; nach längeren Ereignissen entstehen vielfach Gruppierungsgrenzen.
3. Ein Ereignis, das in Klangfarbe oder Artikulationsweise von den umgebenden Ereignissen abweicht, erscheint akzentuiert.

Für Ereignisfolgen mit wahrnehmbarem Tonhöhenverlauf gilt außerdem:

4. Das Klangereignis nach einem Tonhöhenprung erscheint akzentuiert.
5. Die Klangereignisse an den Extrempunkten des Tonhöhenverlaufs (tiefster oder höchster Ton) erscheinen akzentuiert.

In Musik, die metrisch oder tonal organisiert ist, tragen zwei weitere Faktoren zur relativen Akzentuierung von Klangereignissen bei. Fällt ein Ereignis auf einen metrisch betonten Zeitpunkt oder besitzt es eine harmonisch wichtige Tonhöhe, so erscheint es als besonders gewichtet. Allerdings setzt dies voraus, dass beim Hörer aufgrund des vorherigen Klanggeschehens oder aufgrund parallel erklingender Begleitstimmen ein metrisches bzw. tonales und harmonisches Wahrnehmungsschema aktiviert ist. Musikalisch wenig trainierte Hörer oder Hörer aus anderen Musikkulturen werden diese Akzentuierungsweisen vermutlich nicht bzw. schwächer empfinden.

Da es sich bei den genannten Akzentuierungsmöglichkeiten jeweils um relative Akzentuierungsgrade handelt, muss zunächst offen bleiben, wie stark die jeweiligen Unterschiede zu den umgebenden Tönen sein müssen, um eine bestimmte Stärke der Akzentwahrnehmung zu erreichen. Reicht eine Verdoppelung der Tondauer oder ein Sprung von einer kleinen Terz aus, damit eine Akzentuierung wahrgenommen wird? Ebenfalls abhängig vom musikalischen Kontext und von der musikalischen Erfahrung des Hörers ist die Frage, welche der aufgeführten Akzentuierungstypen für die musikalische Wahrnehmung entscheidend sind, wie die Akzentuierungen zusammenwirken und wie sich die Akzentuierungstypen untereinander gewichten lassen.

Daniel Müllensiefen (2004) hat aufgrund der statistischen Auswertung eines Hörexperiments ein kombiniertes Akzentuierungsmaß vorgeschlagen. In seinem Hörversuch sollten musikalische Experten, d.h. sehr reliabel urteilende Hörer, die Ähnlichkeit zwischen je zwei Versionen von einstimmig präsentierten Pop-Melodien einschätzen. Anhand der Daten seines Experiments überprüfte er die zahlreichen in der musikpsychologischen Literatur genannten Akzentuierungsregeln auf ihre empirische Stichhaltigkeit. Mül-

lensiefen (2004, S. 311-319) kommt aufgrund statistischer Auswertungsmethoden zu dem Schluss, dass sich die Urteile seiner Versuchsteilnehmer mit einer Kombination der folgenden fünf Akzentuierungsregeln optimal in Übereinstimmung bringen lassen:²³

1. Ein Klangereignis, das mindestens doppelt so lange dauert, wie der vorhergehende Klang, wird als akzentuiert empfunden;
2. ein Klangereignis nach einem Intervallsprung, der größer als eine große Terz ist, wird als akzentuiert empfunden;
3. Klangereignisse nach einer Umkehr der Tonhöhenkontur erscheinen als akzentuiert;
4. Klangereignisse an Phrasenenden werden als akzentuiert wahrgenommen, wobei Phrasenenden durch Interonsetintervalle gebildet werden, die mindestens viermal so lang sind wie der vorherrschende Notenwert (Modalwert) der Melodie;
5. Klangereignisse, deren Tonhöhe im Dreiklang der implizierten Harmonie enthalten ist, erscheinen als akzentuiert.

Für die Stärke des Akzentuierungsgrads sind demnach neben der Klangdauer und der Position innerhalb einer Gruppierung mehrere tonale Kriterien wichtig: die Stellung einer Tonhöhe in der melodischen Kontur (Intervallsprung und Extrempunkte des Tonhöhenverlaufs) sowie im harmonischen Kontext. Da die Melodien in einer einheitlichen Lautstärke und Klangfarbe gespielt wurden, konnten Akzentuierung durch diese beiden Parameter nicht berücksichtigt werden.

Wechselwirkungen zwischen Akzentuierung und Gruppierung

Die Akzentuierung bestimmter Klangereignisse steht mit der wahrgenommenen Gruppierung in enger Wechselwirkung. Einerseits haben Unterschiede im Akzentuierungsgrad der Klangereignisse – sei es in der Lautstärke, der Tondauer, der Tonhöhe oder in der Stellung innerhalb der melodischen Kontur – Rückwirkungen auf die Gruppenbildung. Andererseits wird durch die Gruppierung der wahrgenommene Grad der Akzentuierung, vor allem die wahrgenommene Lautstärke, beeinflusst. Handel hat die wichtigsten experimentellen Ergebnisse zu diesen Interaktionen zusammengestellt und graphisch veranschaulicht (Handel 1989, S. 386ff).

Beim Phänomen der sog. subjektiven Rhythmisierung isochroner Klangfolgen führt, wie bereits erwähnt, die Gruppenbildung zugleich zur

23 Aufgrund des experimentellen Paradigmas der Ähnlichkeitsurteile ließ sich eine Gewichtung zwischen den fünf Regeln nicht ermitteln, vgl. Müllensiefen 2004, S. 319.

Akzentuierung des Elements nach der wahrgenommenen Gruppengrenze. Umgekehrt kann eine höhere Lautstärke einzelner Klangelemente die Gruppenbildung und sogar die wahrgenommenen Timing-Eigenschaften beeinflussen. Bei einer Erhöhung der Lautstärke jedes zweiten oder dritten Klangereignisses einer isochronen Klangfolge wird das Intervall vor dem laueren Element als verlängert wahrgenommen, nach dem laueren Klang dagegen als verkürzt. Das lautere Element erscheint als akzentuiert und steht am Beginn einer Gruppierung (Abb. 9). Wird jedes zweite oder dritte Ereignis der Klangfolge verlängert, so erscheint das Intervall vor dem längeren Element als verkürzt, das Intervall nach dem Element dagegen als verlängert. Das längere Element wird als akzentuiert empfunden und steht am Ende der Gruppierung (Woodrow 1909, vgl. Abb. 10). Verändert man die Zeitintervalle zwischen den Elementen einer Klangfolge, so werden die verlängerten Intervalle zu Gruppierungsgrenzen. Bei kurzen Pausen zwischen den Klangpaaren wird das Element nach der Pause als akzentuiert empfunden und steht am Beginn einer Gruppe. Bei längeren Pausen erscheint jedoch das Element *vor* dem verlängerten Pausenintervall als akzentuiert; es steht am Ende der Gruppierung (Povel/Okkerman 1981, vgl. Abb. 11).

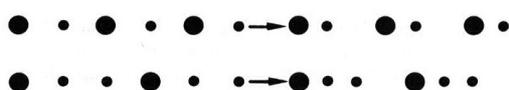


Abb. 9: Wechselwirkungen zwischen Akzentuierung durch Lautstärke und wahrgenommener Gruppierung (nach Handel 1989, S. 387).

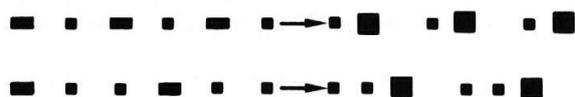


Abb. 10: Wechselwirkungen zwischen Tondauer, Lautstärkeempfinden und wahrgenommener Gruppierung (nach Handel 1989, S. 387).

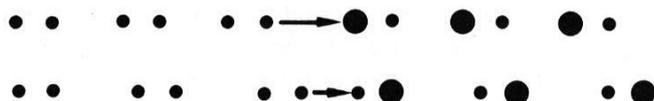


Abb. 11: Wechselwirkungen zwischen unterschiedlichen Interonsetintervallen und Lautstärkeempfinden (nach Handel 1989, S. 387).

Tonhöhenveränderungen können die Wahrnehmung isochroner Klangfolgen auf mehrere Weisen beeinflussen. Erklingen zwei alternierende Tonhöhen, so wird zumeist der höhere Ton als der akzentuierte Ton zu Beginn einer Gruppe wahrgenommen. Die Gruppengrenzen werden durch eine subjektive Verlängerung der entsprechenden Zeitintervalle betont. Treten die beiden Tonhöhen unterschiedlich häufig auf, so steht der seltenere Ton am Anfang einer Gruppierung (Abb. 12).

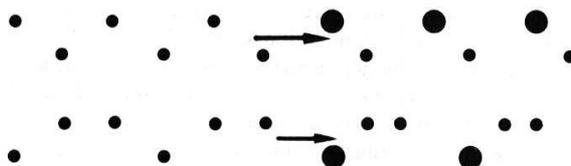


Abb. 12: Gruppierung und Lautstärkeempfinden bei isochronen Klangfolgen mit wechselnden Tonhöhen (nach Handel 1989, S. 387).

Bei isochronen Klangfolgen mit mehreren unterschiedlichen Tonhöhen werden aufgrund der Tonhöheninformationen Gruppierungen gebildet. Handel (1989, 408f) nennt hierfür eine Reihe von Präferenzregeln:

1. Elemente mit der gleichen Tonhöhe werden zu Gruppierungen zusammengefasst;
2. Tonfolgen mit Schritten in gleicher Richtung werden bevorzugt zu Gruppierungen zusammengefasst;
3. Wechselnoten („Triller“) werden zusammengruppiert;
4. Gruppengrenzen entstehen durch Tonhöhen sprünge.

Außerdem können Strukturähnlichkeiten und Parallelismen zwischen den einzelnen Gruppierungen wichtig für die Gruppenbildung sein (z.B. Transpositionen oder Umkehrungen). Am Beginn der Gruppierungen entsteht vielfach eine subjektiv stärkere Akzentuierung, die sich in einem Anstieg der wahrgenommenen Lautstärke sowie der Länge der Gruppengrenzen äußert (sog. strukturelle Akzentuierung).

In vielen Fällen gibt es mehrere Möglichkeiten, die Gruppierung von Klangfolgen mit unterschiedlichen Tonhöhen wahrzunehmen. So führen im folgenden Beispiel Regel 2 (Tonhöhenbewegung in gleicher Richtung) und Regel 3 (Wechselnoten) zu unterschiedlichen Gruppierungen (Abb. 13). Dieselben Klangfolgen können demnach als unterschiedliche Klanggestalten wahrgenommen werden.



Abb. 13: Alternative Möglichkeiten der Gruppierung einer isochronen Klangfolge mit unterschiedlichen Tonhöhen.

Die wahrgenommene Gruppierung und Akzentuierung von Klangereignissen wird somit von einer Vielzahl von klanglichen Variablen beeinflusst. Nicht nur Dauer und Interonsetintervalle, sondern auch Unterschiede einzelner Klangelemente in puncto Lautstärke, Tonhöhe (Stellung in der melodischen Kontur und tonal-harmonische Gewichtung) und Klangfarbe spielen hierbei eine Rolle. Durch die Veränderung einer der genannten Variablen kann sich die wahrgenommene Organisation der gesamten Klangfolge komplett verändern bzw. neu strukturieren. Veränderungen an einem bestimmten Element der Klangfolge (z.B. Veränderungen der Akzentuierung oder des Interonsetintervalls) können die Wahrnehmung aller anderen Elemente beeinflussen. „Given the restructuring of the sequence“, so Handels Schlussfolgerung, „we must conceptualize rhythm as relative timing. The timing and accentuation of any single element is determined relative to the timing and accentuation of all other elements, adjacent ones as well as nonadjacent ones“ (Handel 1989, S. 390). Stellt man in Rechnung, dass zudem eine willentliche Steuerung der Aufmerksamkeit möglich ist, so erweist sich die Rhythmuswahrnehmung bereits auf einer grundlegenden Ebene als ein aktiver und konstruktiver Prozess.

Hinzu kommen gelernte Gruppierungsmechanismen, die aufgrund von Strukturähnlichkeiten, Parallelismen (z.B. parallele Tonbewegungen auf geänderter Tonstufe) oder der exakten Wiederholung von rhythmisch-melodischen Mustern die Wahrnehmung lenken. Wenn Klangmuster oft wiederholt werden, können sie als Ganzes memoriert werden und bei erneutem Auftreten die spontane Segmentierung beeinflussen, auch wenn dies mit anderen „primitiven“ Gruppierungsmechanismen im Widerstreit steht. Bei den auf Wiederholung basierenden Gruppierungsprozessen, bei denen Lernprozesse (Memorieren, Erinnern und Wiedererkennen) und somit Verarbeitungsstrategien des Langzeitgedächtnisses involviert sind, unterscheidet Snyder (2000, S. 45) zwischen stückspezifischen Gruppierungseffekten („objective sets“) und stilspezifischen Gruppierungseffekten („subjective sets“). Stückspezifische Gruppierungseffekte entstehen aufgrund von Prototypen oder Schemata, die beim Hören eines bestimmten Stückes entstehen und durch welche die Wahrnehmung gleicher oder ähnlicher Strukturen an späterer Stelle des Stückes geprägt wird. So werden etwa prägnante melodische oder rhythmische Motive auch an anderen Stellen

innerhalb eines Stückes wiedererkannt – selbst dann, wenn sie in einem völlig anderen Kontext oder mit leichten Variationen wiederholt werden. Stilspezifische Gruppierungseffekte entstehen aufgrund von Erwartungen, die sich durch Hören der Stücke eines bestimmten musikalischen Stiles aufbauen.

3.3 Metrumswahrnehmung

Musikalisches Metrum und metrische Hierarchie

Ein zentrales Merkmal der musikalischen Zeitgestaltung ist die Regelmäßigkeit der Abstände zwischen den Anfängen (Onsets) der einzelnen Klänge (Interonsetintervalle, IOI). Gleich große Interonsetintervalle gibt es nicht nur zwischen direkt nacheinander erklingenden, sondern auch zwischen weiter auseinander liegenden Klangereignissen. Außerdem können sich bestimmte Akzentuierungen und Akzentuierungsgrade der Klänge (durch Lautstärke, bestimmte Klangfarben oder Tonhöhen usw.) in regelmäßigen Zeitabständen wiederholen. Diese Regelmäßigkeiten oder Periodizitäten werden in der Musik gemeinhin mit dem Begriff des Metrums, wörtlich: Maß, in Verbindung gebracht, wobei sich das „Maß“ hier auf die normierten Interonsetintervalle bezieht. Metrische Periodizitäten erzeugen einen schematischen Erwartungsrahmen, der die Aufmerksamkeit auf bestimmte zukünftige Zeitpunkte bzw. auf die dann erklingenden Ereignisse lenkt. Durch diese zukunftsorientierte Aufmerksamkeitslenkung werden Prozesse der musikalischen Informationsverarbeitung sowie der körperlichen und sozialen Synchronisation erleichtert. Dabei werden zumeist leichte Abweichungen der Klang-Onsets von den metrisch exakten Zeitpunkten toleriert. Versteht man unter einem Metrum ein zeitliches Bezugssystem oder kognitives Schema, so kann man dieses „Zurechthören“ als eine Form der schemegeleiteten kategorialen Wahrnehmung beschreiben (vgl. Abschnitt 3.4).

Der einfachste Fall eines metrischen Bezugsrasters ist eine periodische, isochrone Folge von Zeitpunkten. In Musik liegt das Tempo dieser isochronen Folge von Zeitpunkten oder Schlägen innerhalb eines Bereichs von Interonsetintervallen zwischen ca. 200 ms und 2 s (30 bis 300 bpm), bevorzugt jedoch in einem Bereich zwischen ca. 500 und 700 ms (85 bis 120 bpm) dem sog. Grundschatz oder Grundpuls (tactus) zugrunde (London 2002, S. 536). Normalerweise spricht man jedoch erst dann von einem musikalischen Metrum, wenn *mindestens zwei* Periodizitäten oder Pulsfolgen unterschiedlichen Tempos miteinander kombiniert und aufeinander bezogen werden: „A metrical pattern“, so die Musikpsychologen Caroline Palmer und Carol L. Krumhansl, „usually contains nested hierarchical levels, in

which at least two levels of pulsation are perceived at once, and one level is an integer multiple of the other“ (Palmer/Krumhansl 1990, S. 728).

Wird z.B. jeder zweite, dritte oder vierte Zeitpunkt einer Pulsfolge akzentuiert, so entsteht ein zweites, langsames Raster von akzentuierten Schlägen, deren Tempo dann halb, ein Drittel oder ein Viertel so schnell ist; weitere langsamere Akzentuierungsebenen sind möglich. Auf diese Weise entsteht eine sog. metrische Hierarchie mit unterschiedlich schnellen Pulsfolgen, bei der die akzentuierten Zeitpunkte einer höheren hierarchischen Ebene zu den Elementen der nächsttieferen Ebene werden.²⁴ Den Fall eines streng hierarchisch organisierten Metrums zeigt Abb. 14.

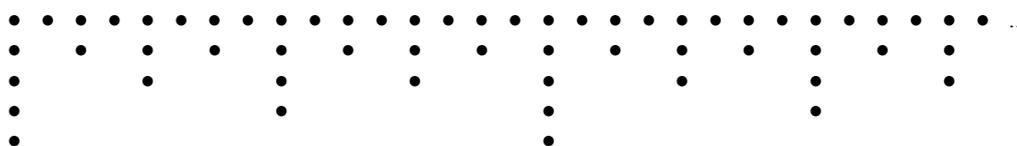


Abb. 14: Metrische Hierarchie mit fünf Ebenen, wobei die Interonsetintervalle jeder tieferen metrischen Ebene doppelt so lang sind wie die Interonsetintervalle der jeweils höheren Ebene.

Mit der Zahl der periodisch organisierten, aufeinander bezogenen Ebenen wächst der Grad an metrischer Strukturiertheit oder metrische „Tiefe“ eines Musikstückes. Besteht eine metrische Hierarchie aus nur wenigen Ebenen, so ist sie relativ „flach“. Der Extremfall einer „flachen“ metrischen Hierarchie wäre eine metrische Organisation, die z.B. neben der Referenzebene des Grundschlags nur noch periodische Grundschlagsunterteilungen aufweist. Existiert allerdings nur eine einzige periodische oder isochrone Ereignisfolge, z.B. der Grundschlag, so sollte nach der oben gegebenen Metrums-Definition nicht von einem musikalischen Metrum gesprochen werden.

Natürlich gibt es in der Musik auch Regelmäßigkeiten, bei denen die in Abb. 14 dargestellte strikte Hierarchie der Dauerverhältnisse durchbrochen wird. So ist denkbar, dass die Einheiten auf einer bestimmten Hierarchieebene unterschiedlich lang sind und etwa durch Addition verschieden langer Zeiteinheiten erzeugt werden. Hierdurch verringert sich die Anzahl der streng periodischen Ebenen und damit der Grad der metrischen Organisa-

²⁴ Allerdings gelangt die Metrumshierarchie bei Periodizitäten längerer Zeitintervallen in einen Übergangsbereich, in dem sich die Metrumserwartung mit bestimmten Formschemata deckt - so bei musikalischen Formmodellen, die sich aus regelmäßigen Taktgruppen zusammensetzen, z.B. der 32taktigen Standardform des American Popular Song (vgl. 5.2).

tion. Werden solche Patterns, wie z.B. das Habanera-Pattern oder ein westafrikanisches Time Line-Pattern (Abb. 15), zyklisch wiederholt, so erklingen genau zwei streng periodische Ebenen: die Interonsetintervalle des schnellsten Pulses und die Länge des gesamten Zyklus.

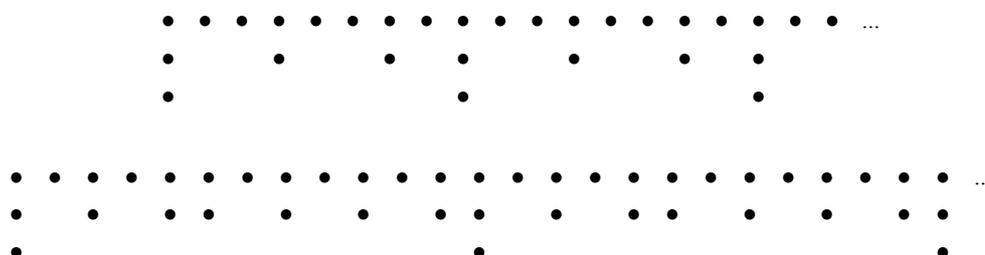


Abb. 15: Zwei Beispiele für eine flache metrische Organisation mit additiver Ebene (Ebene 2); oben: Habanera-Rhythmus (Tresillo, vgl. 5.6), unten: westafrikanisches Time Line-Pattern (vgl. 4.2).

Nicht jede Musik ist periodisch organisiert. Fehlt dem Hörer jedoch jegliches zeitliche Bezugssystem, auf das er die gehörten Klangfolgen beziehen könnte, so wird ein zukunftsorientierte Aufmerksamkeitsmodus und damit der Aufbau von Erwartungen über die Zeitpunkte zukünftiger Klangereignisse unmöglich. Stattdessen ist nur eine analytische, detailorientierte Aufmerksamkeit möglich, wodurch kognitive Leistungen, z.B. Erinnern oder Unterscheiden von Rhythmusgestalten, die Synchronisation von Körperbewegungen mit der Musik sowie eine Koordination der verschiedenen Stimmen innerhalb eines Ensembles, stark erschwert werden.

Metrische und figurale Verarbeitung von Rhythmen

Dirk-Jan Povel und Peter Essens (1985) haben in mehreren Experimenten untersucht, inwieweit der Grad der metrischen Organisation einfacher Klangmuster eine Rolle bei der musikalischen Informationsverarbeitung spielt. Je besser ein rhythmisches Zeitmuster mit einem metrischen Schema übereinstimmt und je höher daher der Grad seiner metrischen Organisation ist, so ihre Hypothese, desto stärker werden kognitive Verarbeitungsprozesse erleichtert. In einem Hörexperiment erhielten die Versuchsteilnehmer die Aufgabe, einfache Rhythmen nachzuklopfen, die mehrmals hintereinander vorgespielt wurden. Gemessen wurde zum einen die Dauer, bis sie den jeweiligen Rhythmus auswendig gelernt hatten und mit dem Nachklopfen begannen, zum anderen der Grad der Übereinstimmung ihrer Repro-

duktion mit der Vorlage. Lernzeit und Reproduktionsgenauigkeit wurden als Indizien für kognitive Gedächtnisleistungen angesehen.

In ihrem Experiment wählten Povel und Essens 35 Rhythmen mit unterschiedlichen Graden an metrischer Organisation. Die Gesamtlänge jedes rhythmischen Patterns betrug 16 Zeiteinheiten in einem Tempo von 300 bpm, also insgesamt 3,2 Sekunden. Povel und Essens gingen davon aus, dass in einem 16er-Zyklus nur eine symmetrische Unterteilung in 4er-Gruppen (4/4-Takt) zu einem sinnvollen metrischen Schema führen kann. Der Grundschlag (Viertelnote) entspräche hierbei einem moderaten Tempo von 75 bpm. Von den 16 möglichen Zeitpunkten wurden acht mit kurzen Klangimpulsen belegt (830 Hz-Rechteckimpulse von 50 ms Dauer), die anderen acht Schläge waren Pausen. Bei den Schlägen 1 und 13 erklang immer ein Impuls; am Schluss aller Rhythmen entstand durch eine drei Zeiteinheiten dauernde Pause eine Gruppierungsgrenze. Die restlichen sechs Klänge verteilten sich auf die Schläge 2 bis 12, wobei Pausen maximal zwei Schläge lang waren. Da die Rhythmen aus identischen Klangergebnissen bestanden, konnten Akzentuierungen nur aufgrund der Timing-Eigenschaften entstehen. Hierbei stützten sich Povel und Essens auf drei von Povel/Okkermann (1981) formulierte Akzentuierungsregeln, wonach ein isolierter Ton, der zweite Ton einer Zweitongruppe sowie der Anfangs- und Schlusston einer Gruppierung von drei oder mehr Tönen als akzentuiert wahrgenommen werden. Aufgrund der Anzahl der akzentuierten Ereignisse, die mit dem Grundschlag des 4er-Metrums (auf 1, 5, 9 und 13) zusammenfallen, bestimmten Povel und Essens sodann in neun Abstufungen den Grad der metrischen Organisation der 35 Rhythmen (vgl. Abb 16).

Die gemessenen Lernzeiten und Reproduktionsfehler korrelierten in hohem Maße mit dem Grad der metrischen Organisation der einzelnen rhythmischen Muster. Der positive Zusammenhang zwischen metrischer Ordnung und kognitiver Verarbeitung wurde in einem zweiten Experiment durch das Hinzufügen eines tiefen Tones auf die metrisch starken Zählzeiten 1, 5, 9 und 13, die das metrische Bezugschema der Rhythmen verdeutlichen sollte, wie zu erwarten weiter verstärkt.

Povel und Essens räumen allerdings ein, dass sich aufgrund der künstlichen Natur des verwendeten Stimulus-Materials die Ergebnisse ihrer Versuche nur mit Einschränkungen auf einen musikalischen Kontext übertragen lassen. Werden Rhythmen von Musikern gespielt, so verdeutliche der Musiker seine Metrums-Interpretation zusätzlich durch weitere Akzentuierungsmöglichkeiten (Lautstärke, Artikulation, Klangfarbe) sowie durch die mikrorhythmische Gestaltung, wodurch Wahrnehmungsprozesse und Gedächtnisleistungen womöglich weiter erleichtert werden.

Außerdem räumen die Autoren ein, dass vermutlich nicht alle Patterns von allen Versuchsteilnehmern automatisch mit einem metrischen Bezugsrahmen in Verbindung gebracht werden. Vielmehr können Muster auch auf

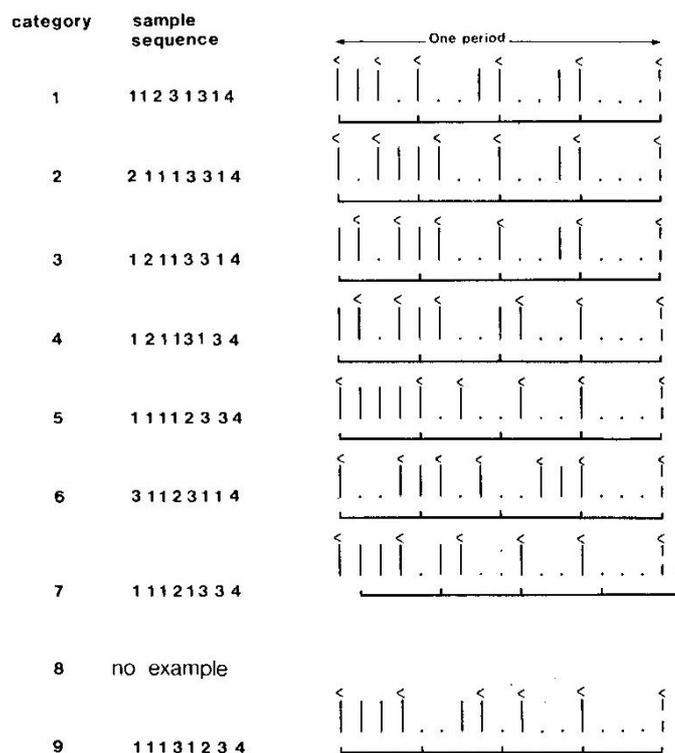


Abb. 16: Beispiele für Rhythmen mit unterschiedlichen Graden an metrischer Organisation (1-9) im Experiment von Povel und Essens (1985); beim 17. Schlag beginnt das Pattern erneut. Über den Klängen (senkrechte Balken) sind die wahrgenommenen Akzentuierungen (<) vermerkt, darunter das beste metrische Bezugsschema. Auffällig ist die Phasenverschiebung des Metrums bei Rhythmus 7 (aus Povel/Essen 1985, S. 420).

andere Weise ohne Zuhilfenahme von metrischen Rastern verarbeitet werden. Aufgrund eines Vergleichs zwischen denjenigen Versuchsteilnehmern, die bei der Reproduktionsaufgabe die besseren Ergebnisse erzielten und denen, die schlechter abschnitten, vermuten die Autoren, dass es neben der metrischen Repräsentation auch die Möglichkeit einer figuralen Repräsentation der Rhythmen gibt. Bei der figuralen Wahrnehmung wird ein Klangmuster nicht auf ein Zeitraster bezogen, sondern als Gruppierung von Klangelementen wahrgenommen, die aufgrund der *relativen* Länge oder Kürze der Interonsetintervalle eine charakteristische rhythmische Gestalt besitzen. Daher können leicht Verwechslungen zwischen rhythmischen Mustern auftreten, bei denen Pausen – also Gruppierungsgrenzen – zwar an derselben Stelle im Muster auftreten, jedoch unterschiedlich lang sind (Abb. 17).



Abb. 17: Zwei figural gleiche zeitliche Muster (acht Ereignisse verteilt auf zwölf Zeitpunkte); beim zweiten Muster ist die Pause zwischen Schlag 3 und 4 verlängert, die Pause zwischen Schlag 1 und 2 verkürzt.

Handel (1998) widmet sich in drei Experimenten der Frage, welcher Stellenwert der figuralen bzw. metrischen Kodierung bei der kognitiven Verarbeitung von Zeitmustern tatsächlich zukommt. Seine Versuchspersonen sollten den Grad der Ähnlichkeit bzw. Unterschiedlichkeit zweier Klickfolgen einschätzen. Aufgrund der Untersuchungsergebnisse kommt Handel zu dem Schluss, dass bei der Beurteilung der Ähnlichkeit von Rhythmuspaaren entscheidend ist, ob die beiden Rhythmen figural gleich oder figural unterschiedlich organisiert sind, ob sie also aus einer Abfolge derselben Gruppierungen – wobei die Pausen zwischen den Gruppierungen nicht gleich lang sein müssen – oder aber aus unterschiedlichen Gruppierungen bestehen. Wenn zwei Rhythmen eine unterschiedliche figurale Organisation aufweisen, so wird dies von den Hörern leicht erkannt. Erstaunlicherweise hat der Grad an metrischer Organisation in diesen Fällen keinen Einfluss auf die Ähnlichkeitseinschätzung. Nur wenn die figurale Struktur beider Rhythmen identisch ist, helfen Unterschiede in der metrischen Organisation, die Rhythmen zu unterscheiden – allerdings nur dann, wenn sich bereits der erste Rhythmus metrisch gut einordnen lässt. Da bei einer identischen figuralen Organisation die Unterschiedlichkeit nur aufgrund der Länge der Pausen zwischen den Gruppierungen bestimmt werden kann, ist in diesem Fall der Bezug auf ein metrisches Raster das entscheidende Kriterium der Unterscheidung – ähnlich wie ein Tonalitätsraster bei der Unterscheidung ähnlicher Melodiekonturen helfen kann. Vom Standpunkt der kognitiven Verarbeitung aus betrachtet, ist somit die metrische Struktur eines rhythmischen Musters ein Spezialfall der grundlegenderen figuralen Verarbeitung. „From this perspective“, so Handels Fazit, „the better reproduction of metric rhythms occurs not because they are metric, but because they are simpler, and ‚simpler‘ would be the preferred explanation because it is applicable in a broader domain“ (Handel 1998, S. 1561). Regelmäßige zeitliche Klangstrukturen werden demnach von vielen Hörern nicht unbedingt als „metrischer“ wahrgenommen, sondern vermutlich als „einfacher“. Möglichkeiten einer figuralen und einer metrischen Kodierung bestehen vermutlich parallel nebeneinander, wobei die Personen mit größeren musikalischen Vorkenntnissen, z.B. aufgrund eigenen Musizierens, eher auf eine metrische Orientierung zurückgreifen.

Entwicklungspsychologische Studien weisen darauf hin, dass es sich bei metrischen Zeitrastern um Schemata handelt, die im Laufe der Kindheit erlernt werden und über die nicht alle Menschen im selben Ausmaß verfügen. Zu dieser Schlussfolgerung kommt Jeanne Bamberger aufgrund ihrer Analyse der musikalischen Notationsgraphiken von Kindern (Bamberger 1982, 1991). Während bei den figuralen Notationsformen der vier- bis fünfjährigen Kinder die Bewegung des Rhythmus frei zeichnerisch nachgeahmt wird, setzen erst die etwas älteren Kinder regelmäßige Zeitintervalle in gleich große Abstände zwischen den Symbolen in ihren Zeichnungen um. Auch die Fähigkeit zur Produktion von regelmäßigen Schlagfolgen bzw. zur Synchronisation mit regelmäßigen Klangstrukturen ist dem Menschen nicht angeboren, sondern wird in den ersten Lebensjahren erlernt. Kleinkinder können im Alter von anderthalb bis zweieinhalb Jahren ihre Bewegungen nur für kurze Dauer mit Musik synchronisieren; drei- bis vierjährige Kinder können zumeist noch nicht in einem konstanten Tempo gehen (vgl. Gembris 1998, S. 285, s. auch Bamberger 1991). Minkenberg (1991, S. 89f) stellte fest, dass Kinder im Alter von ca. fünf Jahren metrisch regelmäßig singen können. Interessanterweise lässt sich erst im Zeitraum zwischen dem fünften und zehnten Lebensjahr eine Präferenz der Kinder für metrisch organisierte Rhythmen gegenüber metrisch ungebundenen Rhythmen feststellen (Zenatti 1993, S. 182f).

Metrische Akzente beeinflussen die Rhythmuswahrnehmung

In vielen populären Musikstilen ist ein metrischer Bezugsrahmen häufig bereits durch das Spiel der Begleitinstrumente – Schlagzeug, Bass, Akkordinstrumente – eindeutig vorgegeben, sodass die melodisch-rhythmischen Linien z.B. von Gesang oder Soloinstrument im Kontext dieses Erwartungsrahmens gehört werden. In diesem Fall wird durch den bestehenden metrischen Bezugsrahmen die wahrgenommenen Akzentuierungen und damit die Rhythmusstruktur der Melodiestimme beeinflusst.

Povel und Essens (1985) konnten experimentell nachweisen, dass ein erklingender metrischer Bezugsrahmen die Wahrnehmung von Rhythmen maßgeblich prägt. Dazu ließen sie die Gleichheit bzw. Unähnlichkeit von zwei aufeinander folgenden Rhythmen einschätzen. Verwendet wurden jeweils dieselben Zeitmuster von 12 Schlägen Länge, die jedoch einmal mit einem 3er-Metrum (tiefer Klangimpuls auf die Zählzeiten 1, 4, 7 und 10), das andere Mal mit einem 4er-Metrum (tiefer Klangimpuls auf die Zählzeiten 1, 5 und 9) unterlegt waren (Abb. 18).



Abb. 18: Rhythmus (Balken oben) mit zwei unterschiedlichen metrischen Bezugsrahmen (Balken unten) (nach Povel/Essens 1985, S. 434).

Die Ähnlichkeit bzw. Gleichheit der beiden in ihrer zeitlichen Struktur identischen Patterns wurde aufgrund der unterschiedlichen metrischen Bezugsrahmen bei allen zwölf Rhythmuspaaren nicht erkannt. Der metrische Kontext prägt somit die Art und Weise, wie zeitliche Muster als Rhythmen wahrgenommen und kognitiv verarbeitet werden, ganz entscheidend mit. Ein aktiviertes metrisches Schema lässt verschiedene Klangereignisse in unterschiedlichem Ausmaß als akzentuiert erscheinen. Klangereignisse, die mit metrisch stark gewichteten Zeitpunkten zusammen fallen, werden als besonders akzentuiert wahrgenommen – sie besitzen einen sog. metrischen Akzent. Die Akzentuierungsstruktur wiederum prägt maßgeblich die Wahrnehmung des Rhythmus der Klangfolge.

Modelle der Grundschlags- und Metrumsinduktion

Wie entsteht nun ein metrisches Orientierungsraster während des Musikhörens? Wie kann aus klanglichen Strukturen, die ja nicht nur Periodizitäten, sondern auch viele Unregelmäßigkeiten enthalten, ein metrisches Schema abgeleitet werden? In den vergangenen Jahrzehnten ist eine fast unüberschaubare Forschungsliteratur zur Frage der Induktion von Grunds Schlag (Grundpuls, Beat) und Metrum entstanden (vgl. hierzu im Überblick Clarke 1999, S. 482-489, Snyder/Krumhansl 2001, S. 455f und 483ff, Todd/Lee/O'Boyle 2002, S. 26f, Langner 2002, S. 17-24, Temperley 2004). Das große Interesse der Psychologen an diesem Phänomen hängt damit zusammen, dass Fragen des sog. „time tracking“ oder der „beat induction“ in vielen Lebensbereichen relevant sind. Am Beispiel der Musik kann untersucht werden, welche grundlegenden Mechanismen dem menschlichen Koordinations- und Synchronisationsvermögen zugrunde liegen. Viele Erklärungsansätze führen zu formalisierten Modellen, die als Computerprogramme realisiert werden. Durch Vergleiche zwischen dem Verhalten des Modells und dem Verhalten von Hörern, denen in Hörexperimenten dieselben Inputs präsentiert werden, können die Modelle schrittweise verbessert werden. Das wissenschaftliche Erkenntnisziel, durch Modellbildung und empirische Überprüfung der Modelle den Prozessen menschlicher Wahrnehmung und Informationsverarbeitung auf die Schliche zu kommen, verbindet sich dabei mit einer praxisorientierten Zielsetzung: Es sollen Com-

puterprogramme entwickelt werden, die Grundschat und Metrum eines vorgegebenen Musikstückes verlässlich bestimmen können. Wird dieses Ziel erreicht, so ist dies freilich noch kein Beweis dafür, dass die neuronalen und kognitiven Vorgänge im menschlichen Gehirn auf dieselbe Art und Weise funktionieren wie das konstruierte Computermodell. Um die Modelle möglichst realitätsnah zu gestalten, wird jedoch mitunter versucht, sie auf neurophysiologische Erkenntnisse aufzubauen. Im Folgenden soll auf eine Reihe von Gemeinsamkeiten, Unterschiede und Probleme der Modelle hingewiesen werden. Sie unterscheiden sich vor allem hinsichtlich des verwendeten Inputs, der Prinzipien der Modellierung und des Outputs.

Als Input der Modelle werden in den meisten Fällen ausschließlich Informationen aus Notentexten verwendet. Dies stellt bereits eine starke Vereinfachung der Wahrnehmungsprozesse dar, da Unterschiede in Lautstärke oder Klangfarbe sowie Mikrotiming-Schwankungen nicht berücksichtigt werden. Dagegen verwenden nur wenige Modelle als Input die Audiosignale der Musikaufführungen (so die verschiedenen Modelle von Neil Todd, s.u., und Toiviainen 1997), wobei hier ein Mechanismus zur Bestimmung der Tonanfänge vorgeschaltet wird.

Die meisten Modelle der Metrumsinduktion basieren auf einer Analyse der Regelmäßigkeiten von Interonsetintervallen. Informationen aus Melodik, Harmonik und Klangfarbe und deren Auswirkungen auf die Akzentuierung der einzelnen Klangereignisse werden dagegen nicht berücksichtigt. Es werden nicht nur Interonsetintervalle zwischen benachbarten Tönen untersucht, sondern auch die Abstände zwischen Ereignissen, die weiter voneinander entfernt sind. Aufgrund dieser zeitlichen Informationen versuchen die Modelle, zeitliche Regelmäßigkeiten und damit eine metrische Struktur abzuleiten. Wenn zwei oder mehr der erkannten Periodizitäten miteinander harmonisieren, d.h. in einem ganzzahligen Verhältnis (1 : 2, 1 : 3, 1 : 4) zueinander stehen und außerdem in ihrer Phasenlage übereinstimmen, so ergibt sich hieraus das Metrum der Tonfolge.

Die verschiedenen Modellierungsansätze unterscheiden sich allerdings hinsichtlich der Art und Weise, wie die Periodizitäten aus den Input-Daten abgeleitet werden. Todd/Lee/O'Boyle (2002, S. 26f) und Snyder/Krumhansl (2002, S. 483f) unterscheiden übereinstimmend drei Typen der Modellierung: Die regelgeleiteten Modelle (z.B. Steedman 1977, Longuet-Higgins/Lee 1982, Povel/Essens 1985, Lee 1991) formulieren Regelsysteme, nach denen aufgrund der zeitlichen Strukturen der Klangfolgen ein möglicher Grundschat bzw. ein mögliches Metrum hergeleitet wird. Das Modell von Parncutt (1994) berücksichtigt zusätzlich einen bevorzugten Tempobereich. Povel/Essens (1985, S. 413f) vertreten die Idee einer inneren Uhr bzw. einer mehrschichtig gegliederten Uhrenhierarchie, deren Tempo sich nach bestimmten Regeln den Klangfolgen anpasst. Das Uhrenmodell stößt aller-

dings angesichts mikrorhythmischer Temposchwankungen an seine Grenzen; außerdem fehlen Hinweise zur neurophysiologischen Existenz solcher inneren Uhren.²⁵ Ein zweiter Modellierungsansatz verwendet sog. adaptive, non-lineare Oszillatoren, die sich an Tempoveränderungen anpassen können (Large/Kolen 1994, Large/Jones 1999 u.a.). Dieser Ansatz knüpft an das Modell der dynamischen Aufmerksamkeit von Mari Riess Jones (vgl. 3.1) an; die Oszillatoren entsprechen dabei dem Aufmerksamkeitssystem, das sich den Ereignisfolgen dynamisch anzupassen vermag.²⁶ Ein dritter Modellierungsansatz ermittelt die Periodizitäten der eintreffenden Audio-Signale aufgrund eines linearen Filterungsprozesses (Todd 1994, Scheirer 1998, Todd/O'Boyle/Lee 1999, Todd/Lee/O'Boyle 2002 u.a.) Anders als die Uhren- und Oszillatorenmodelle deckt sich diese Modellierung mit neurophysiologischen Charakteristika der Sinneswahrnehmung. Todd, O'Boyle und Lee berücksichtigen zudem die Rolle bestimmter Bewegungsschemata bei der Grundschlagsinduktion (vgl. 3.5).

Snyder und Krumhansl kommen bei ihrem Vergleich der zahlreichen Modelle zu folgender Bewertung: „Since filtering models and oscillator models can both track temporal variability in music and represent pulse and meter, determining which approach best describes human pulse-finding mechanisms will likely require detailed computational analysis and experimentation“ (Snyder/Krumhansl 2002, S. 484.)

Empirische Ergebnisse zur Metrumswahrnehmung

Die meisten Autoren fordern von ihrem Modell der Metrumsinduktion einen Output in Form eines eindeutigen Grundschlags oder einer eindeutigen metrischen Hierarchie. Das Modell soll demnach nur eine einzige richtige Lösung ermitteln. Laut Langner (2002, S. 22ff) werden Mehrdeutigkeiten der metrischen Struktur von manchen Autoren zwar erwähnt, jedoch nur selten wird anerkannt, dass es sich hierbei um eine besondere Qualität der musikalischen Wahrnehmung handeln könnte. Ausnahmen stellen die Ansätze von Povel/Essens (1985), Longuet-Higgins/Lee (1984) sowie von Parncutt (1994) dar, bei denen die Möglichkeit metrischer Mehrdeutigkeiten ausdrücklich gewürdigt wird. So gewichtet Parncutt (1994) in seinem Modell verschiedene Möglichkeiten der metrischen Interpretation einfacher rhythmischer Muster: „A central assumption of the present model“, so Parncutt, „is the inherent ambiguity of the underlying pulse (tactus) and meter of a rhythm. The model does not output a single solution, but instead

25 Dieser Einwand betrifft ebenfalls die konnektionistischen Ansätze, welche die Metrumsinduktion als sog. neuronale Netzwerke zu modellieren versuchen (Desain 1992, Desain/Honing 1999).

26 Dagegen bevorzugt Langner (2002) in seinem Modell eine Oszillatorenbank mit 4080 Oszillatoren, die fest auf bestimmte Frequenzbänder eingestellt sind.

considers many possible pulse and meter sensations, estimating the relative importance or salience of each“ (Parncutt 1994, S. 433).

Parncutt legt seinem Modell der Grundschlagsinduktion die empirischen Ergebnisse zweier Hörversuche zugrunde. Im ersten Experiment sollten seine Versuchsteilnehmer zu vorgegebenen Rhythmusmustern, die in mehreren Tempi gespielt werden, einen Puls klopfen; im zweiten Experiment mussten sie den Akzentuierungsgrad eines durch einen tiefen Ton markierten Zeitpunkts innerhalb des Rhythmus einschätzen. Parncutt verwendete sechs Klickfolgen: eine isochrone Schlagfolge sowie fünf Zeitmuster mit verschiedenen Interonset-Verhältnissen, die er teilweise mit genretypischen Grundrhythmen in Verbindung bringt: Walzer 2 : 1 (♩ ♪), Marsch 2 : 1 : 1 (♩ ♪ ♪), Swing 3 : 2 : 1 (♩ ♩ ♪) sowie „Skip“ 3 : 1 : 2 (♩ ♪ ♪) und „Cross“ 2 : 1 : 1 : 2 (♩ ♪ ♪ ♪). Aufgrund der Schlagfolgen, welche die Versuchsteilnehmern geklopft haben, wird bei jedem rhythmischen Muster für jeden Zeitpunkt eine Gewichtung errechnet. Auf diese Weise entsteht aus den empirischen Daten (den Klopfhäufigkeiten zu jedem Zeitpunkt des Patterns) das statistische Modell einer Metrumshierarchie.

Die Versuchsteilnehmer klopften bei den sechs Patterns isochrone Schlagfolgen, die, abhängig vom Tempo, bei den meisten Rhythmen in einem hierarchischen oder geschachtelten Verhältnis zueinander standen, so etwa 1 : 3 : 6 Schläge beim Walzer oder 1 : 2 : 4 Schläge beim Marsch. Allerdings stellen einige der Beobachtungen eine eindeutige, streng hierarchische Metrumsorganisation in Frage. So wurden die vier Schläge des Marsch-Rhythmus im schnellen Rhythmus von manchen Personen als 2 : 1 : 1 (♩ ♪ ♪), von anderen jedoch als 1 : 1 : 2 (♩ ♪ ♪) zusammengefasst. Insbesondere die beiden letzten Rhythmen („Skip“ und „Cross“) führten zu zwei verschiedenen, gegeneinander phasenverschobenen metrischen Orientierungen. Somit lassen rhythmische Muster vielfach mehrere Deutungsmöglichkeiten zu, insbesondere aufgrund von Phasenverschiebungen zwischen den metrischen Hauptakzenten.

In musikalischen Kontexten sind vergleichbare Mehrdeutigkeiten der metrischen Organisation weit verbreitet. Carolyn Drake, Amandine Penel und Emmanuel Bigand (2000) forderten Versuchsteilnehmer auf, beim Hören verschiedener Musikstücke der Klavierliteratur (Bach, Scarlatti, Beethoven, Chopin, Ravel und Bartók) Schlagfolgen zunächst in einem ihnen angenehmen Tempo mitzuklopfen. Anschließend sollten sie schnellere und langsamere Tempi mitklopfen. Es stellte sich heraus, dass die 36 Versuchsteilnehmer, 18 Musiker und 18 Nicht-Musiker, sich neben dem bevorzugten Tempo in durchschnittlich zwei bis drei weiteren Tempi mit den Musikstücken synchronisieren konnten, wobei sie eher schnellere als lang-

samere Tempi wählten.²⁷ Eine willkürlicher Wechsel des Aufmerksamkeitsfokus von der Referenzebene des Grundschlags auf andere metrische Ebenen ist somit unproblematisch. Die gewählten Schlagfolgen entsprachen in ca. 80 % der Fälle der hierarchischen Metrumsstruktur der Musikstücke; allerdings kam es auch hier, ähnlich wie im Versuch von Parncutt, zu Phasenverschiebungen: Musiker wählten in 29 % der Fälle, Nicht-Musiker sogar in 49 % beim Mitklopfen eine Pulsfolge, die nicht mit dem metrischen Hauptakzent laut Takteinteilung in der Notenvorlage übereinstimmte. Die metrische Hierarchie verliert somit auf tieferen metrischen Ebenen an Eindeutigkeit. Das Finden des „richtigen“ Metrums hängt zudem stark vom Grad der musikalischen Ausbildung ab. Nicht-Musiker sind in ihrer metrischen Interpretation weitaus flexibler als Musiker.

Handel untersuchte in einer Reihe von Experimenten die Wahrnehmung der Kombinationen von zwei oder mehr unterschiedlich schnellen Pulsfolgen, die mit einem gemeinsamen Anfangspunkt beginnen und deren Geschwindigkeiten in ganzzahligem Verhältnis zueinander stehen, z.B. 3 : 4 : 5 (Handel/Oshinsky 1981, Handel/Lawson 1983, vgl. auch Handel 1984). Hiervon erhoffte er sich Aufschlüsse über die kognitive Verarbeitung mehrstimmiger musikalischer Strukturen, die nicht oder nur in geringem Maße im Sinne einer metrischen Hierarchie organisiert sind. Experimentell variiert wurden das Tempo, die Konfiguration der Pulsfolgen sowie die Tonhöhen, Dauern oder Lautstärken einzelner Elemente einer oder mehrerer Pulsfolgen. Die Versuchsteilnehmer wurden gebeten, „[to] respond to any rhythmic structure that they perceive and prefer“ (Handel 1984, S. 471), ohne Vorgaben, *was genau* sie mitklopfen sollten. Dadurch sollten die Reaktionen der Probanden nicht von vornherein auf isochrone Schlagfolgen eingeschränkt werden. Vielmehr ließ diese Versuchsanweisung verschiedene Strategien und Interpretationsweisen zu, die charakteristisch für bestimmte Rhythmen bzw. Hörer sind. Handel unterscheidet drei Möglichkeiten der Wahrnehmung der polyrhythmischen Strukturen:

1. Aus dem Zusammenspiel mehrerer Pulsfolgen entstehen sog. Kreuzrhythmen unterschiedlicher Komplexität.
2. Eine Pulsfolge wird zum metrischen Bezugsrahmen für die anderen Pulsfolgen.
3. Bei schnellen Tempi werden nicht mehr Schläge zur Referenzebene, sondern zyklische Einheiten vom gemeinsamen Anfangspunkt bis zum nächsten gemeinsamen Anfangspunkt der Pulsfolgen.

²⁷ Darüber hinaus stellten Drake, Penel und Bigand fest, dass eine expressive Interpretation der Musikstücke die Fähigkeit zur Synchronisation gegenüber einer mechanischen Aufführung nicht etwa erleichterte, wovon Povel und Essens (1985) ausgegangen waren, sondern etwas erschwerten.

Es zeigte sich, dass eine Reihe von Kriterien Einfluss darauf haben, welche der drei Strategien von den Versuchsteilnehmern gewählt werden und welche der Pulsfolgen (bei Strategie 2) als metrischer Bezugsrahmen wahrgenommen wird. Entscheidend ist das Tempo der Pulsfolgen, da in der Regel Pulsfolgen nur in einem Bereich von 200-800 ms IOI (75 bis 300 bpm) mitgeklopft werden. Daneben haben Veränderungen in der Lautstärke und Tonhöhe, bisweilen auch der Dauer der Klänge sowie vor allem die Art der polyrhythmischen Konfiguration starken Einfluss auf die rhythmische Interpretation der Versuchsteilnehmer. Auffällig groß waren die individuellen Unterschiede zwischen den Probanden.

Handel kommt zu dem Schluss, dass ein polyrhythmisches Pattern mehrere Rhythmen enthält und verschiedene Möglichkeiten der Wahrnehmung zulässt. Die Interpretation von Rhythmen ist abhängig von einer Reihe klanglicher Faktoren, die sich untereinander beeinflussen und beeinträchtigen. Eine Vorhersage der rhythmischen Interpretation allein aufgrund der zeitlichen Struktur ist nicht möglich: „The prediction of rhythmic interpretation may be impossible in novel contexts due to the interacting, emergent nature of the rhythmic levels“ (Handel 1984, S. 481).

Empirische Studien zur Wahrnehmung polyrhythmischer Strukturen in musikalischen Kontexten sind nach wie vor selten. Uwe Seifert, Fabian Olk und Albrecht Schneider (1995) befragten Studierende am musikwissenschaftlichen Institut der Universität Hamburg hinsichtlich ihrer Wahrnehmung zweier am Sequencer programmierter Patterns. Die Patterns kombinierten ein westafrikanisches Time Line-Pattern (Glocke) mit regelmäßigen Schlagfolgen (tiefe Trommeln), was zu einer Hemiolenbildung und zu regelmäßigen, aber auch zu unregelmäßigen Akzentuierungen führte. Die Beispiele wurden bewusst metrisch mehrdeutig konstruiert, was sich in den Aussagen der Versuchspersonen zum Metrum bzw. zur Takteinteilung der Musikbeispiele niederschlug. Zudem wurde aufgrund der unregelmäßigen Einwürfe die Zyklizität der Patterns mitunter nicht erkannt, insbesondere dann nicht, wenn die Gesamtlänge des Zyklus mit neun Sekunden den Umfang der psychologischen Gegenwart überschritt.

Weitere Aspekte der Metrumsinduktion

In musikpsychologischen Untersuchungen und Modellen zur Wahrnehmung von Grundschat und Metrum werden vorwiegend die Timing-Informationen, vor allem die Interonsetintervalle berücksichtigt, wobei mit Vorliebe Folgen von identischen Klangereignissen, z.B. Klickfolgen, verwendet werden. Akzentuierungen durch Lautstärke werden dagegen nur selten, Akzentuierungen aufgrund der Harmonik oder Melodik so gut wie nie berücksichtigt. Clarke (1999, S. 489) kritisiert diese Eindimensionalität

der meisten Untersuchungen und Modelle zur Metrumsinduktion. Obwohl ein Metrum als Ergebnis sowohl von zeitlichen als auch von nicht-zeitlichen Strukturmerkmalen, sowohl von Dauern bzw. Pausen als auch von verschiedenen Akzentuierungen angesehen wird, werden nur die zeitlichen Informationen als Einflussgröße für die Induktion einer metrischen Hierarchie herangezogen. Eine Übertragung auf musikalische Kontexte ist daher nur unter Vorbehalten möglich.

Jones (1993) und Monahan (1993, S. 144ff) beschreiben dagegen Modelle, in denen, wie in europäischer Musik üblich, melodische und zeitliche Akzentuierungen übereinstimmen. Wenn sich die Timing-Struktur mit der melodisch-harmonischen Struktur in Einklang befindet, werde die Wahrnehmung eines eindeutigen metrischen Schemas gefestigt: „Usually, the combination of pitch-level and natural temporal accenting will clarify the position of the bar or measure“ (Monahan 1993, S. 145); „[...] accents from pitch and time patterns combine into a common dynamic structure called a Joint Accent Structure“ (Jones 1993, S. 76). Studien von Dawe, Platt und Racine (1993, 1994, 1995) weisen darauf hin, dass bereits bei relativ einfachen Akkordfolgen die zeitliche Abfolge von Harmoniewechseln nachhaltigen Einfluss auf die Wahrnehmung der gesamten Rhythmusstruktur ausübt. Bei einem Vergleich zwischen Musikern und Nicht-Musikern zeigte sich, dass bei Musikern die Harmoniewechsel der entscheidende Faktor der Metrumswahrnehmung sind – noch vor verschiedenen zeitlichen und melodischen Akzentuierungsmöglichkeiten.

Ein weiterer möglicher Faktor für die Metrumsinduktion sind Parallelismen, d.h. ähnlich strukturierte („parallele“) Klangfolgen werden bevorzugt demselben metrischen Schema zugeordnet (Temperley/Bartlette 2002). Neben metrischen Schemata aufgrund von Parallelismen innerhalb bestimmter Stücke spielen vermutlich auch stilspezifische Schemata bei der Metrumswahrnehmung eine Rolle, insofern ähnliche rhythmisch-metrische Muster in verschiedenen Stücken eines Stilbereichs gehäuft auftreten.

Vermutlich werden beim Hören komplexerer musikalischer Strukturen die gehörten Rhythmen mit gelernten, metrisch organisierten Prototypen verglichen („pattern matching“). Je besser die rhythmische Textur eines Stückes mit dem gelernten Prototyp übereinstimmt, um so leichter können die Rhythmen kategorial verarbeitet werden. Nach Christopher S. Lee verknüpft der Hörer jedes Metrum mit bestimmten Patterns von starken und schwachen Schlägen, sog. „canonical accent-pattern“ (Lee 1991, S. 121). Diese metrischen Prototypen werden möglicherweise kulturspezifisch gelernt.²⁸ Auch Palmer und Krumhansl weisen auf die Möglichkeit hin, dass ein Metrum nicht aus auditiven Informationen abgeleitet wird, sondern dass

28 Einen vergleichbaren Ansatz vertritt Johnson-Laird mit seinen schematisch gelernten Rhythmusfamilien (Johnson-Laird 1991).

vielmehr den Klangstrukturen gelernte metrische Strukturen quasi übergestülpt werden: „[...] meter may be independent of sensory functions; instead, abstract knowledge of the periodic temporal functions that operate in Western tonal music may play a role in meter perception“ (Palmer/Krumhansl 1990, S. 729). Auf diese Weise könnte ein vertrautes metrisches Schema als Hörstrategie eingesetzt werden, um unvertraute musikalische Strukturen mental zu organisieren. Denkbar wäre außerdem, dass bestimmte rhythmische Muster auch dann einen Erwartungsrahmen bilden, wenn sie nicht oder nur eingeschränkt metrisch gebaut sind. Allerdings ist auch hier die Frage berechtigt, ob alle Menschen unabhängig vom Grad ihrer musikalischen Vorbildung und Erfahrung bei der Wahrnehmung gleichermaßen auf metrische Schemata zurückgreifen oder ob nicht auch eine figurale, gruppierende Organisation von auditiven Stimuli praktiziert wird, insbesondere bei musikalisch wenig erfahrenen Hörern.

3.4 Mikrotiming

Bei der Aufführung von Musik kommt es immer wieder zu Abweichungen von einem metronomisch exakten Zeitraster und zu Schwankungen des lokalen Tempos. Diese Phänomene werden als Mikrotiming oder Mikrorhythmik bezeichnet. Musiker machen immer wieder kleine unbeabsichtigte „Fehler“ im Sinne von kleinen Abweichungen vom isochronen Grundschlag und dessen exakter Unterteilung oder von graduellen Tempoänderungen der Schlagfolgen. Diese Fehler können zufällig verteilt sein oder sie können aufgrund von spieltechnischen Schwierigkeiten systematisch erfolgen.²⁹ Bei vielen Mikrotiming-Abweichungen handelt es sich allerdings nicht um unbeabsichtigte, sondern um intendierte Abweichungen, die systematisch und regelgeleitet vorgenommen werden. Daher liegt der Schluss nahe, dass sie im Prozess der musikalischen Kommunikation zwischen Musikern und Hörern bestimmte Funktionen erfüllen.

Rasch (1978, 1979) hat darauf hingewiesen, dass leichte Asynchronizitäten (bis ca. 30 ms und mehr) die Wahrnehmbarkeit unterschiedlicher Tö-

29 Letzteres zeigt ein Versuch von Krampe et al. (2000), bei dem verschiedene Pianisten aufgefordert wurden, auf einem digitalen Flügel einen „4 über 3“-Polyrhythmus in 17 verschiedenen Tempi zu spielen. Die graphische Darstellung der Interonsetintervalle der linken Hand (4er-Pattern) bzw. rechten Hand (3er-Pattern) zeigt bei langsameren Tempi eher zufällige Abweichungen. In schnelleren Tempi kommt es jedoch zu einem regelmäßigen Abweichungsmuster, das einen systematischen Wechsel zwischen zu langen und zu kurzen Intervallen dokumentiert. Krampe et al. werten dies als Beleg dafür, dass bei schnelleren Tempi ein Fehlerkorrekturmechanismus in Gang kommt, bei dem das motorische System der beiden Hände regelmäßig um die metronomisch exakte Lösung herum oszilliert.

ne und Klangfarben in mehrstimmiger Musik fördern (vgl. hierzu auch Gabrielsson 1999, S. 543f). Indem durch die asynchronen Einsätze einer möglichen Maskierung von Frequenzen entgegengewirkt wird, lassen sich verschiedene Klangströme und damit verschiedene Klangquellen besser voneinander unterscheiden.

Typen der mikrorhythmischen Gestaltung

Es lassen sich drei mikrorhythmische Gestaltungsweisen unterscheiden (vgl. Abb. 19):

1. Ein Musikstück besitzt in allen Stimmen lokale Temposchwankungen.
2. Einzelne Stimmen einer musikalischen Textur, vor allem die Melodiestimmen, weichen von einem gleichmäßigen Tempo ab, während andere Stimmen, vor allem die Begleitstimmen, ein metrisch exaktes Tempo beibehalten.
3. Die Interonsetintervalle der Klangereignisse einer metrischen Ebene werden systematisch verändert, während die IOIs der nächsttieferen metrischen Ebene exakt gleich groß sind.

1. Lokale Temposchwankungen aller Stimmen	2. Mikrotiming-Abweichungen einer Stimme gegenüber der streng metrischen Begleitung	3. Asymmetrische Unterteilung einzelner Ebenen der metrischen Hierarchie
Freies Tempo Rubato	Gebundenes Tempo Rubato	Mikrorhythmische Asymmetrie
Klassisch-romantische Musik	Musik des Barock, romantische Klaviermusik (z.B. bei Chopin), Gesang und Solisten in populärer Musik	Volksmusik, populäre Musik: - Swing-Achtel - Walzer-Viertel

Abb. 19: Typen der mikrorhythmischen Gestaltung.

Lokale Temposchwankungen aller Stimmen und Mikrotiming-Schwankungen einzelner Stimmen werden gemeinhin als Tempo Rubato (wörtlich: gestohlene Zeit) bezeichnet. Mit diesem Ausdruck wird darauf hingewiesen, dass einzelne Klänge benachbarten Klängen etwas von ihrer Dauer (genauer: von den Interonsetintervallen) abziehen oder „stehlen“. In der abendländischen Musikgeschichte kann eine frühe Praxis des Tempo Rubatos, bei dem die Stimme des Solisten gegenüber der Ensemble-Begleitung mikrorhythmische Freiheiten besitzt, von einem seit dem 19. Jahrhundert verbrei-

teten Tempo Rubato mit Temposchwankungen der gesamten klanglichen Textur unterschieden werden (Hudson 1994). Die frühe Praxis des Tempo Rubato wird auch als gebundenes Tempo Rubato, die spätere als freies Tempo Rubato bezeichnet. Beim dritten Fall der mikrorhythmischen Gestaltung sind die Interonsetintervalle z.B. auf der Ebene unterhalb des Grundschlags ungleichmäßig lang, während die Grunds Schlagfolge selbst von diesen Schwankungen unbeeinflusst bleibt; oder die Interonsetintervalle der Grunds chlagebene sind unterschiedlich lang, während größere metrische Einheiten (z.B. Takte) gleich lang sind. Zur Bezeichnung dieser Gestaltungsweise möchte ich den Ausdruck „mikrorhythmische Asymmetrie“ vorschlagen. Bekannte Beispiele hierfür sind der ungleichmäßig unterteilte Grunds chlag im Jazz (sog. Swing-Achtel, vgl. 6.2) und die ungleich langen Takteile (Viertelnoten) im Walzer (vgl. 5.1).

Methoden der Mikrotimingforschung

Die musikpsychologische Erforschung der mikrorhythmischen Gestaltung basiert auf der exakten Messung von Toneinsätzen bei Musikaufführungen. Dabei wurden verschiedene Vorgehensweisen entwickelt (vgl. Kopiez 1996, 507ff, Gabrielsson 1999, 525ff). In Anknüpfung an Untersuchungen von Carl Seashore aus den 1930er Jahren (Seashore 1938) untersuchte in den 1960er und 1970er Jahren eine Forschungsgruppe von schwedischen Musikwissenschaftlern und Psychologen um Ingmar Bengtsson und Alf Gabrielsson, ob es bei der Aufführung von Musik zu systematischen und somit regelgeleiteten mikrorhythmischen Abweichungen gegenüber der Notentextvorlage kommt (Bengtsson 1975, Bengtsson/Gabrielsson 1980, 1983, vgl. auch Gabrielsson 1988, S. 42ff). Nicht nur Aufnahmen westlicher Kunstmusik, sondern auch populäre Musik und schwedische Volksmusik wurden auf ihre mikrorhythmischen Abweichungen hin untersucht.

Die Mikrotiming-Messungen erfolgen normalerweise aufgrund des Lautstärkeverlaufs (Amplitudenhüllkurven) von Musikaufnahmen mit einer zeitlichen Auflösung im Millisekundenbereich.³⁰ Seit den 1980er Jahren wurde durch die MIDI-Technik eine Erhebung und Auswertung von großen Datenmengen sehr erleichtert, da die mikrorhythmischen Strukturen von Einspielungen auf MIDI-Pianos direkt in den Computer eingelesen werden können. Bei Aufnahmen komplexerer, mehrstimmiger Musik lassen sich die Toneinsätze der einzelnen Stimmen allerdings anhand von Amplitudenhüllkurven nur schwer bestimmen, da die Hüllkurven stets den Laut-

30 Der genaue Zeitpunkt des Onsets wird bei dieser Messmethode in der Regel per Augenschein bestimmt. Die Reliabilität der Messungen wird durch Messwiederholung ermittelt; abhängig von der Art des akustischen Stimulus liegt die Unschärfe der Messung bei Bengtsson und Gabrielsson im Bereich von +/- 5 Millisekunden (Bengtsson/Gabrielsson 1980, S. 259, vgl. auch Gabrielsson 1987, S. 87).

stärkeverlauf aller Stimmen abbilden, wobei in der Regel die Stimmen einander überdecken. Durch Spektrogramme, die Onsets frequenzabhängig darstellen, können die Messmöglichkeiten bei klanglich komplexeren Aufnahmen jedoch erweitert werden (vgl. Friberg/Sundström 2002).

Bengtsson und Gabrielsson vergleichen die tatsächlich gemessenen Interonsetintervalle mit den aufgrund der Gesamtdauer des untersuchten Stückes (bis zum Beginn des Schlusstones) errechneten durchschnittlichen Dauern der entsprechenden Noten- und Pausenwerte. Auf diese Weise werden sowohl die absoluten Abweichungsdauern als auch die in Prozentwerten ausgedrückten relativen Abweichungen von einer mechanischen Ausführung bestimmt. Die Abweichungswerte können graphisch in einer sog. Tempokurve dargestellt werden, bei der Tonverlängerungen oberhalb, Tonverkürzungen unterhalb der x-Achse notiert sind (vgl. Abb. 20).

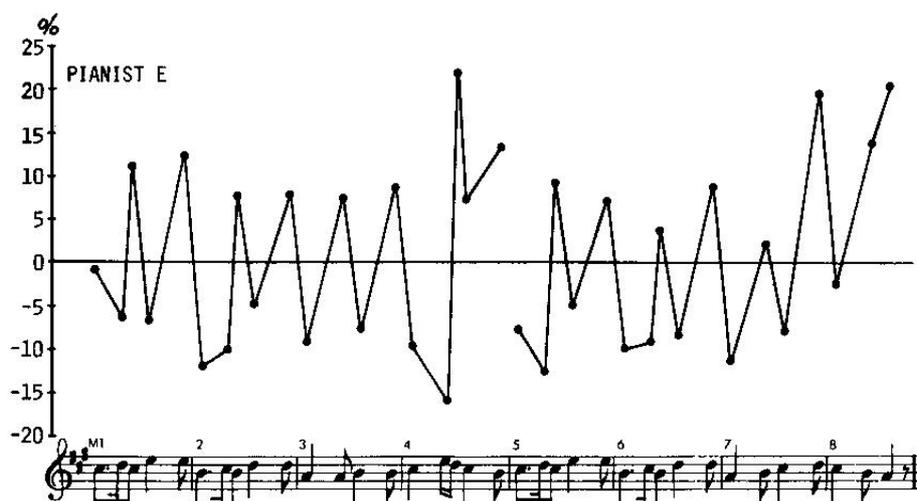


Abb. 20: Beispiel einer Tempokurve: Interpretation des Themas des ersten Satzes aus Mozarts Klaviersonate A-Dur (KV 331) durch Alfred Brendel (Aufnahme Philips 9500 025, 1975). Die Abweichungen der Tondauern vom Notentext sind in Prozenten oberhalb (verlängert) und unterhalb (verkürzt) der x-Achse notiert (nach Gabrielsson 1987, S. 91).

Bei der statistischen Auswertung der Mikrotiming-Messungen werden Tendenz- und Streuungswerte für verschiedene Einheiten des Stückes (Takte, Phrasen, Taktgruppen, Abschnitte), Verhältniswerte für Nachbartöne usw. ermittelt – innerhalb einer Aufnahme sowie bei verschiedenen Interpretationen eines Musikstückes. Mit einer Faktorenanalyse der Abweichungswerte bei mehreren Aufführungen eines Stückes können schließlich Zusam-

menhänge zwischen verschiedenen Mikrotiming-Abweichungen und charakteristische Interpretationsweisen ermittelt werden.

Werden aufgrund der Messungen bestimmte Regelmäßigkeiten festgestellt, so muss überprüft werden, ob diese Regeln tatsächlich von Bedeutung für die Musikwahrnehmung sind. Dies kann durch Synthese von Klangbeispielen aufgrund der ermittelten Regeln und deren Bewertung in Hörversuchen erfolgen. Auch hier stellt die MIDI-Technik, die eine feine Manipulation von Tondauern und Interonsetintervallen (sowie der Lautstärke) zulässt, eine große Erleichterung dar (vgl. Busse 2002).

Mikrorhythmische Nuancen und kategoriale Wahrnehmung

Wenn in Klangfolgen die Klang-Onsets von den isochronen Zeitpunkten des einmal etablierten metrischen Rasters leicht abweichen, werden sie in der Wahrnehmung vielfach auf diese Zeitpunkte zurechtgerückt. Dies zeigt sich bereits in den Experimenten von Paul Fraise, bei denen Versuchsteilnehmer gebeten wurden, Tonfolgen mit unterschiedlich großen Interonsetintervallen nachzuklopfen. Die Interonsetintervalle werden von den Versuchsteilnehmern derart verkürzt oder verlängert, dass schließlich alle geklopfen Interonsetintervalle in einem Verhältnis von 2 : 1 stehen oder gleich lang sind (Fraise 1982).

Der englische Musikpsychologe Eric F. Clarke konnte in einem Hörversuch zeigen, dass mikrorhythmische Abweichungen einen einmal etablierten metrischen Wahrnehmungsrahmen nicht außer Kraft setzen (Clarke 1987). Vielmehr werden die Mikrotiming-Veränderungen toleriert, in den metrischen Bezugsrahmen eingegliedert und somit „zurechtgehört“. In einem Hörexperiment präsentierte Clarke seinen Versuchsteilnehmern verschiedene Versionen einer einfachen Melodie. Die Melodieanfänge wurden in zwei unterschiedlichen metrischen Kontexten präsentiert: einmal in einem 3er-Metrum (6 Töne mit einer Dauer von je 320 ms), das andere Mal in einem 2er-Metrum (5 Töne mit einer Dauer von 480 ms bzw. 240 ms) (vgl. Abb. 21). Die Dauer zweier Töne am Schluss der Melodie wurde stufenweise verändert. Das Verhältnis ihrer Dauern betrug einmal 2 : 1 (640 ms : 320 ms), dann 1 : 1 (480 ms : 480 ms) sowie schließlich IOI-Verhältnisse in sieben weiteren Abstufungen (620 : 340 ms, 600 : 360 ms usw. in Schritten à 20 ms). Die Versuchspersonen sollten nun entscheiden, ob die beiden Interonsetintervalle in einem Verhältnis von 2 : 1 oder 1 : 1 stehen. Es bereitete ihnen keine Schwierigkeiten, die unregelmäßigen Tondauern in den bestehenden metrischen Rahmen einzupassen. Beim Kontext des 3er-Metrums neigten sie eher zur Einschätzung, es handle sich um ein Längenverhältnis von 2 : 1, beim 2er-Metrum eher zu 1 : 1.

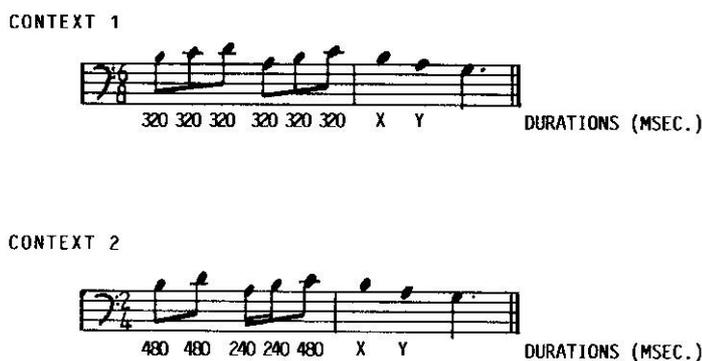


Abb. 21: Melodie in unterschiedlichen metrischen Kontexten; die Länge der Töne X und Y wird stufenweise verändert (aus Clarke 1987, S. 23).

In einem zweiten Versuch sollten die Probanden den Grad der Unterschiedlichkeit von jeweils zwei Versionen der Melodie beurteilen, bei denen die Dauern der beiden Töne gegenläufig um 20 ms verkürzt bzw. verlängert wurden. Unterschiede der beiden Versionen wurden innerhalb des unscharfen Zwischenbereichs von 560 : 400, 540 : 420 und 520 : 440 ms relativ gut wahrgenommen; bis zu 80 % der Versuchspersonen erkannten hier die Abweichungen. Dagegen wurden Unterschiede zwischen zwei Melodieversionen, die nur wenig von den regulären metrischen IOI-Verhältnissen (2 : 1 bzw. 1 : 1) abwichen, von weit weniger Versuchspersonen (ca. 60 %) erkannt. Die Bereitschaft zum Zurechthören wächst also mit der Nähe zu einem metrisch interpretierbaren Kontext.

Clarke bringt das Phänomen des Zurechthörens mit dem Konzept der kategorialen Wahrnehmung in Verbindung. Wenn bei der Wahrnehmung von Musikstücken bestimmte metrische Bezugsschemata aktiviert werden, so handelt es sich bei den Längen der Interonsetintervalle um Kategorien, die aufgrund des Grundtempos bestimmte Standardwerte besitzen. Durch die kategorialen Wahrnehmung werden mikrorhythmische Abweichungen innerhalb eines bestimmten Toleranzspielraums durch die Standardwerte der metrisch exakten Interonsetintervalle ersetzt.

Die Toleranz gegenüber nuancenhaften Veränderungen der Tondauer innerhalb eines bestehenden metrischen Rahmens sollte jedoch nicht darüber hinwegtäuschen, dass sehr feine Abweichungen in der Größenordnung von wenigen Millisekunden dennoch auch von ungeübten Hörern wahrgenommen werden können. Dass Hörer einerseits Mikrotiming-Schwankungen kategorial „zurechthören“, sie andererseits dennoch wahrnehmen können, lässt sich durch die unterschiedlichen Aufmerksamkeitsmodi erklären. Während beim Hören metrisch organisierter Musik eine zukunftsorientierte Aufmerksamkeit leichte Abweichungen vom metrischen Bezugsraster

tolerabel erscheinen lässt, lösen die Teilnehmer des zweiten Versuchs von Clarke ihre Aufgabe vermutlich in einem detailorientierten Aufmerksamkeitsmodus.

Clarke verwendet zur Bezeichnung aller messbaren Mikrotiming-Variationen den Ausdruck „expressive timing“ (Clarke 1987, 1991 sowie 1999, 489f). „After the temporal information for rhythm has been perceptually categorised“, so Clarke, „any ‚remainder‘ (i.e. any deviations from a perfect categorical fit) is considered to be expressive information, or perhaps accidental inaccuracy“ (Clarke 1987, S. 30). Alle Informationen zu den exakten Tonlängen, die aus dem kategorialen Rahmen der metrischen Periodizität herausfallen, werden demnach als expressive Nuancierungen wahrgenommen. Die Hörer können, so Clarke, oft nicht genau bestimmen, wie diese Abweichungen beschaffen sind (ob zu lang oder zu kurz, wie groß usw.) Die Präferenz für regelmäßige, normierte Klangdauern, die in einem ganzzahligen Verhältnis (1 : 1, 1 : 2, 1 : 3) zueinander stehen, bestimmt zwar die strukturelle, nicht jedoch die expressive Dimension der Musikwahrnehmung. In diesem Zusammenhang vergleicht Clarke die kategoriale Rhythmuswahrnehmung mit der visuellen Wahrnehmung von Objekten im Raum, da in beiden Fällen mehrere Wahrnehmungsdimensionen in eine konsistente Übereinstimmung gebracht werden müssen. In der Musikwahrnehmung handelt es sich um die beiden Dimensionen metrische Struktur und expressives Mikrotiming. Abbildung 22 veranschaulicht die unterschiedliche expressive Wahrnehmung von IOIs abhängig vom metrischen Kontext; während im Bezugsschema eines 3/8-Metrums in der Mitte der Phrase ein Ritardando wahrgenommen wird, erscheint im 2/4-Metrum die erste Note durch die Verlängerung als besonders akzentuiert.

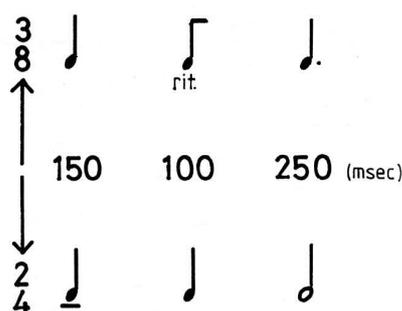


Abb. 22: Zwei mögliche Wahrnehmungen einer Klangfolge mit den Interonsetintervallen 150, 100 und 200 ms in unterschiedlichen metrischen Kontexten (aus Clarke 1987, S. 32).

Problematisch ist in diesem Zusammenhang allerdings der Ausdruck „expressiv“. Wenn Clarke darunter alle „[...] continuously variable parameters of a performance that are used by a player to convey an interpretation of the music“ bzw. alle „systematic departures from the indications of the score“ (Clarke 1991, S. 185) versteht, so widerspricht dies dem üblichen Gebrauch des Terminus „Ausdruck“ („expression“), nach dem „of course, the structure of the piece is in itself a basic correlate of experienced expression“ (Gabrielsson 1999, S. 522).

Mikrotiming in klassischer Musik: Strukturverdeutlichung

Clynes und Walker (1982) haben festgestellt, dass Pianisten ihre mikrorhythmische Gestaltungsweise bestimmter Klavierstücke mitunter über Jahre hinweg konstant beibehalten. Weshalb weichen nun aber klassisch ausgebildete Musiker in ihrem Spiel dermaßen absichtsvoll von den Notenvorlagen ab? Welchen Stellenwert und welche Funktionen besitzen diese Gestaltungsspielräume für das Musikhören? Nach Todd (1985) und Clarke (1991) lernen Musiker mikrorhythmische Gestaltungsweisen nicht einfach nur durch Nachahmung, sondern gelangen zu diesen „expressiven“ Interpretationseigenheiten aufgrund ihres Verständnisses der Struktur eines Musikstückes: „[...] the expressive properties of the performances are symptomatic of the performers' representations of the structure of the music“ (Clarke 1991, 186). Eine zentrale Funktion der mikrorhythmischen Gestaltung ist es demnach, die strukturellen Eigenschaften eines Musikstückes, z.B. musikalische Gruppierungen und formale Gliederungen, klanglich zu verdeutlichen. Nach Kopiez ist es die Theorie vom Zusammenhang zwischen mikrorhythmischem Ausdruck und Struktur, welche „[...] mit ihrer Verschränkung von expressiven Mitteln und Tonsatzstruktur die momentan beste Erklärung für die Funktion ausdrucksbedingter Abweichungen von aller mechanischen Regelmäßigkeit des Spiels bietet. Eine definitive Bestätigung der Ausdruck-Struktur-Theorie unter Verwendung komplexer Musikstücke steht allerdings noch aus“ (Kopiez 1996, S. 566). Clarke vermutet, dass es eine begrenzte Anzahl von Interpretationsregeln gibt, mit denen die Musiker ihre strukturelle Repräsentation der Musik in die mikrorhythmische Gestaltung ihrer Aufführungen transformieren. Veränderungen und Unterschiede in der Interpretation eines Stückes korrespondieren demnach mit Mehrdeutigkeiten der musikalischen Struktur (Clarke 1991, S. 186).

Die grundlegende Regel der Strukturverdeutlichung durch mikrorhythmische Gestaltung bezieht sich auf systematische Verlängerungen von Notenwerten an Gruppierungsenden, z.B. am Schluss von melodischen Motiven, Phrasen und längeren Abschnitten. Systematische Vergleiche

zwischen der musikalischen Gruppierungsstruktur und entsprechenden Verzögerungen in verschiedenen Aufführungen hat Neil Todd (1985) durchgeführt. Aufgrund der Musiktheorie von Fred Lerdahl und Ray Jackendoff (1983, vgl. 4.1) bestimmte er die hierarchische Gruppierungsstruktur mehrerer Klavierkompositionen und versah die verschiedenen Gruppierungsgrenzen mit unterschiedlichen Gewichtungen. Auf der Grundlage dieser theoretischen Annahmen errechnete er lokale Tempoverlangsamungen, die bei den untersuchten Aufnahmen der Stücke meist gut mit den von Musikern tatsächlich vorgenommenen Temposchwankungen übereinstimmen.

Bruno Repp (1992) konnte zeigen, dass sich das Strukturverständnis der Hörer in ihrer Erwartung von lokalen Temposchwankungen niederschlägt. In zwei Versuchen prüfte er, mit welcher Sicherheit Hörer einzelne Verlängerungen von Achtelnotenwerten in einer ansonsten mechanischen Aufführung³¹ an unterschiedlichen Stellen der musikalischen Phrasenstruktur erkennen. „The hypothesis was that *listeners would find it more difficult to detect lengthening in places where they expect it to occur*, particularly at the ends of structural units, in strong metric positions, and at points of harmonic tension“ (Repp 1992, S. 344).

Die Verlängerung jeweils einer einzigen Achtelnote in einer Folge von 47 Achtelnoten der mechanischen Aufführung (Achtel = 341 Millisekunden) um 13,3 %, 10 % bzw. 6,7 % (um 45,4 ms, 31,1 ms bzw. 22,7 ms) wurde von den 20 Versuchspersonen in durchschnittlich 25,7 % der Fällen erkannt; 22,6 % aller Versuchspersonen bezogen die Verlängerung allerdings auf die Achtel nach dem tatsächlich verlängerten Ton. Nimmt man beide Angaben zusammen, so kommt man auf einen Prozentsatz richtig erkannter Verlängerungen, der mit knapp 50 % weit über der Wahrscheinlichkeit von ca. 2 % für zufällige Angaben (bei 47 möglichen Achteln) liegt. Dabei gab es abhängig von der Position der Achtel jedoch erhebliche Unterschiede (zwischen 0 und 90 %) der richtigen Treffer. Verlängerungen am Beginn und Schluss von Phrasen oder Motiven wurden weit schlechter erkannt als in der Mitte von Phrasen und Gruppierungen. Die Ergebnisse des Hörtests (gemittelt über alle Versuchsteilnehmer) wurden außerdem mit dem Mikrotiming-Profil einer exemplarischen Aufführung des Stückes durch einen Konzertpianisten (Murray Perahia) verglichen. Zwischen den erkannten Achtelverlängerungen und den Timing-Schwankungen der Aufführung ergab sich eine signifikante, hohe negative Korrelation ($r = -.59$, bei $p < .001$). An den Stellen, an denen der Konzertpianist sein Spieltempo verlangsamte, wurden verlängerte Achtelnoten von der Versuchsteilnehmern demnach nur sehr schwer erkannt. Die Ergebnisse von Repps Unter-

31 Bei dem Experiment wurde ein Ausschnitt aus dem dritten Satz der Sonate Nr. 18, Es-Dur, op. 31 Nr. 3, von Ludwig van Beethoven gewählt. Der Ausschnitt wurde nach der Notenvorlage von einem MIDI-Piano interpretiert.

suchung sprechen für die Annahme, dass die Wahrnehmung von Zeitintervallen durch strukturell bedingte Erwartungen beeinflusst wird. Wo aufgrund der Gruppierungsstruktur Tonverlängerungen erwartet werden, lassen sich tatsächlich erfolgte Tonverlängerungen nur schwer erkennen. Da den meisten Hörern das verwendete Musikbeispiel nicht bekannt war, muss zudem davon ausgegangen werden, dass die Hörer ihre Erwartungen bezüglich der Gruppierungsstrukturen nicht aufgrund ihrer Vorkenntnis eines bestimmten Musikstückes entwickeln, sondern aufgrund ihrer generellen musikalischen Kompetenz.³²

Die Strukturverdeutlichungs-Theorie wird der Vielschichtigkeit der musikalischen Mikrorhythmik-Phänomene allerdings nicht in vollem Umfang gerecht. Repp (1992) gibt zu bedenken, dass der Musiker einen Spielraum für das Ausmaß der mikrorhythmischen Gestaltung besitzt, den er nach persönlichen Vorlieben gestalten kann. Auch Clarke räumt ein, dass die strukturellen Eigenheiten eines Stückes nicht die einzige Determinante für das Mikrotiming in der musikalischen Aufführung sind. Als weitere Einflussfaktoren nennt er Möglichkeiten des Instruments und der Raumakustik, bestimmte Ideologien der Aufführungspraxis, weitere Aspekte der Kommunikation zwischen Musiker und Publikum im Konzertsaal sowie Stimmungen und Intentionen des Musikers (vgl. Clarke 1991, S. 189).

Gabrielsson (1988) weist auf die motivationalen und emotionalen Aspekte des musikalischen Erlebens hin. Ihn interessiert der Einfluss des Mikrotimings auf den Ausdruck und die Wahrnehmung nicht nur der Strukturen eines Stückes, sondern ebenso auf dessen Bewegungscharakter und auf die damit verbundenen motivationalen und emotionalen Qualitäten des Musikerlebens. Viele Melodien und rhythmische Begleitfiguren besitzen einen je eigenen Bewegungscharakter. Besonders im Falle der Verlangsamung zum Ende einer musikalischen Struktureinheit und dem Ritardando am Schluss eines Stückes, bei dem die musikalische Bewegung zur Ruhe kommt, liegt eine Verbindung von Tempoveränderungen und Bewegungsempfindung auf der Hand.

Bislang wurde in erster Linie das Phänomen von lokalen Temposchwankungen der gesamten musikalischen Textur (freies Tempo Rubato) untersucht, da diese Form der mikrorhythmischen Gestaltung für die Tradition der klassisch-romantischen Musik zentral ist (vgl. Gabrielsson 1999, S. 540). In den meisten populären Musikrichtungen spielen lokale Tempoverlangsamungen, z.B. bei Phrasenenden oder formalen Einschnitten, jedoch kaum eine Rolle. Von Ausnahmen (z.B. Einleitungen von Aufnahmen, solistische Darbietungen, langsame Stücke) abgesehen wird der metrische Bezugsrahmen in den meisten populären Musikstilen vielmehr vergleichs-

32 Repp konnte zudem keinen Zusammenhang zwischen den Einschätzungen und demographischen Daten oder musikalische Vorkenntnissen feststellen.

weise exakt beibehalten. Stattdessen spielt das gebundenen Tempo Rubato, also die mikrorhythmischen Nuancierungen nur einer Stimme, zumeist des Gesangs oder des Melodieinstruments, in vielen Genres populärer Musik eine wichtige Rolle. Aber auch innerhalb der Begleitung kommt es zu systematischen Mikrotiming-Abweichungen. Außerdem wird in manchen Musikstilen systematisch auf niederer metrischer Ebene von einer Isochronizität der Schlagfolgen abgewichen. Auf Mikrotiming-Phänomene in populärer Musik wird im Zusammenhang der stilanalytischen Darstellungen zurückzukommen sein (vgl. Abschnitte 5.1, 5.2, 5.5, 6.2 und 7.2).

3.5 Bewegung, Rhythmus, Musik

Musik ist in mehrerer Hinsicht mit Bewegung verbunden (vgl. Shove/Repp 1995, Clarke 1999, S. 494ff, Clarke 2001). Den physikalischen Grundlagen von Musik liegen zahlreiche Bewegungen zugrunde, so die körperlichen Spielbewegungen der Musiker und die mechanischen Schwingungen der Musikinstrumente und Lautsprechermembrane, die sich wellenartig ausbreitenden Bewegungen der Luftteilchen und die Bewegungen der Basilarmembran im Innenohr. In fast allen Musikkulturen der Welt steht Musik zudem in enger Verbindung mit den Körperbewegungen bzw. dem Tanzen der Zuhörer; nicht selten tanzen auch die Musiker. Die europäische Kunstmusiktradition hat in dieser Hinsicht einen Sonderweg eingeschlagen. Hier hat sich seit wenigen hundert Jahren eine Kultur des passiven Zuhörens entwickelt (Bessler 1975). Doch auch innerhalb der europäischen Kunstmusiktradition gehört der Ausdruck „musikalische Bewegung“ zum Standardvokabular der Musikbeschreibung. Mit dem Topos der musikalischen Bewegung bzw. der „tönend bewegten Formen“, die „Inhalt und Gegenstand der Musik“ sind (Hanslick 1854, S. 32), werden allerdings oftmals nicht die genannten physikalischen Bewegungen umschrieben, sondern Bewegungsempfindungen, die mit den musikalischen Strukturen in Verbindung gebracht werden.

Es lassen sich demnach mehrere Aspekte des Zusammenhangs von Musik und Bewegung unterscheiden, die zu verschiedenen Auffassungen von der „Realität“ der musikalischen Bewegungen führen. Weit verbreitet ist die Auffassung, dass es sich beim Topos der musikalischen Bewegung um eine Metapher handelt, um die Übertragung der Erfahrung von Bewegungen physikalischer Objekte oder des eigenen Körpers auf das musikalische Klanggeschehen. Wahrnehmungspsychologische Ansätze versuchen, bestimmte musikalische Bewegungsempfindungen durch grundlegende Wahrnehmungsmechanismen oder durch assoziative Verknüpfungen zwischen musikalischen Strukturen und der Erfahrung physikalischer Bewe-

gungen zu erklären. Ein weiterer Ansatz bringt die Bewegungsempfindung mit bestimmten Körperbewegungen oder motorischen Schemata beim Musikmachen oder Musikhören in Verbindung. Die musikpsychologische Erforschung dieser Aspekte steht allerdings noch in den Anfängen. Im Folgenden sollen erste theoretische Ansätze und die spärlichen empirischen Untersuchungen zum Zusammenhang von Bewegung, Rhythmus und Musik zusammengetragen werden. Gefragt wird nach der Rolle von Bewegungsmetaphern für die Musikwahrnehmung, nach dem Zusammenhang zwischen Musik und Körperbewegungen sowie nach möglichen Grundlagen der musikalischen Bewegungsempfindung in der auditiven Bewegungswahrnehmung. Außerdem soll auf einen weiteren Aspekt der musikalischen Bewegungserfahrung hingewiesen werden: Die emotionalen Erfahrungen, die beim Hörer durch den Rhythmus und das Tempo der Musik ausgelöst werden, können als inneres „Bewegt-Werden“ und „Bewegt-Sein“ erfahren werden.

Was sind rhythmische Bewegungen?

Ähnlich wie der Begriff der Zeit ist Bewegung einerseits Gegenstand der menschlichen Alltagserfahrung, andererseits ein Konzept der Physik, genauer der Mechanik. Bewegungen im Sinne der klassischen Mechanik sind kontinuierliche und gerichtete Ortsveränderungen im dreidimensionalen Raum. Aus der zurückgelegten Strecke pro Zeitintervall resultiert die Geschwindigkeit sowie gegebenenfalls eine Veränderung der Geschwindigkeit (Beschleunigung oder Verlangsamung). Sowohl Geschwindigkeiten als auch Beschleunigungen sind immer im Raum gerichtet und werden daher in der Physik durch Vektoren ausgedrückt. Auf diese Weise wird es möglich, in unterschiedliche Richtungen verlaufende Bewegungen zu addieren.

Bei vielen Bewegungen ist allerdings noch kein Rhythmus, keine rhythmische Struktur zu erkennen. Will man rhythmisch strukturierte von ungeordneten, nicht-rhythmischen Bewegungen unterscheiden, so ist es naheliegend, innerhalb der Bewegung nach einer fasslichen Regelmäßigkeit oder Ordnung zu fragen, nach bestimmten Markierungen, die eine Segmentierung in Bewegungseinheiten ermöglichen, sowie nach deren Akzentstruktur, d.h. nach jenen besonders markierten Zeitpunkten, an denen z.B. eine Geschwindigkeits- oder Richtungsänderung vonstatten geht. Eine einförmige, kontinuierliche und daher unstrukturierte Bewegung durch den physikalischen Raum ist nicht Gegenstand der menschlichen Erfahrungswelt. Vielmehr sind Bewegungen in der alltäglichen Erfahrung in bestimmte Abläufe eingebunden. Bewegungen sind Vorgänge mit einem Anfang und einem Ende, wobei die Grenzen einer Bewegungseinheit durch bestimmte Ereignisse markiert werden. Zugleich beinhalten natürliche Bewe-

gungen in der Regel eine kontinuierliche, oder aber eine sprunghafte Änderung der Geschwindigkeit oder Bewegungsrichtung – so etwa bei ruckartigen Änderungen der Richtung und Geschwindigkeit aufgrund des Aufpralls des bewegten Körpers auf ein Hindernis oder aufgrund von neu einwirkenden Kräften. Nicht-rhythmische Bewegungen wären dagegen alle kontinuierlichen Bewegungen ohne erkennbare Struktur oder mit kontinuierlichen Richtungs- und Geschwindigkeitsänderungen, so etwa der freie Fall eines Gegenstandes (kontinuierliche Geschwindigkeitsänderung), oder die Flugbahn eines geworfenen Balles (als Addition von gerichteter Bewegung und freiem Fall, woraus eine kontinuierliche Änderung von Bewegungsrichtung und Geschwindigkeit resultiert). Schließlich spielt bei der Erfahrung rhythmischer Bewegungen die Fasslichkeit ihrer zeitlichen Ausdehnung, die Gabrielsson (1993) in seiner Definition der musikalischen Rhythmuserfahrung ins Feld führt, ebenfalls eine Rolle.

Die Erfahrung rhythmischer Bewegungen kann somit analog zur musikalischen Rhythmuserfahrung mit den Kriterien der Gruppierung, der Akzentuierung, der Regelmäßigkeit und der Grenzen der psychologischen Gegenwart in Verbindung gebracht werden. Diese Analogie bildet eine Grundlage für die Übertragung der Erfahrung von physikalischen Bewegungen auf die musikalische Rhythmuserfahrung.

Musikalische Bewegung: Metapher und Assoziation

Eine Metapher ist eine assoziative Verbindung zwischen zwei unterschiedlichen Bereichen der menschlichen Erfahrung. Wenn sich ein menschlicher Erfahrungsbereich nur schwer in Worte fassen lässt, so werden gerne Beschreibungsweisen aus anderen Erfahrungsbereichen metaphorisch auf den fraglichen Bereich übertragen. Grundlegend sind dabei physikalische und physiologische, also körpernahe Vorstellungen. Als körperliche Wesen, die in einer räumlich geordneten Umwelt leben, erfahren Menschen Oben und Unten, Innen und Außen, Zentrum und Peripherie, Verbindungen zwischen Elementen und Teil-Ganzes-Beziehungen, sie kennen Wege, die zu einem Ziel führen, und erfahren die Linearität der Bewegung von Objekten im Raum. Diese grundlegenden Vorstellungsschemata oder „Image Schemas“, so die These des Linguisten George Lakoff und des Philosophen Mark Johnson, strukturieren mentale Konzepte und werden metaphorisch auf die unterschiedlichsten Bereiche der menschlichen Erfahrung übertragen (vgl. Lakoff 1987, Lakoff/Johnson 2000).³³ „Image schemas“, so Bob Snyder,

33 Der kognitive Metaphernbegriff unterscheidet sich vom literaturwissenschaftlichen Verständnis einer literarischen Metapher. Lakoff schreibt: „Thus, schemas are not understood in terms of meaning postulates and their interpretations. Rather, meaning postulates themselves only make sense given schemas that are inherently meaningful because they structure our direct experience“ (Lakoff

„represent the most stable constancies and structures we all share as human beings, derived from dynamic patterns of interaction with our environment“ (Snyder 2000, S. 108). Metaphorische Beschreibungen von musikalischen Phänomenen sind nichts Ungewöhnliches. Metaphern im Sinne grundlegender Vorstellungsschemata prägen nicht nur das Sprechen über Musik, sie formen die Musikwahrnehmung bereits auf einer grundlegenden Ebene (vgl. Snyder 2000, S. 107-119). Ein Musikstück hat Grenzen; es „enthält“ Klänge, Melodien, Rhythmen (und schließt Umweltgeräusche in der Regel aus); ein ganzes Musikstück besteht aus einzelnen Teilen. In der Musik „fällt“ eine Linie von den hohen Tönen hinab zu tiefen Tönen, die vielfach zugleich eine Art tonales Gravitationszentrum und ein Ziel darstellen, das auf verschiedenen Wegen und mit unterschiedlichen Bewegungen erreicht werden kann. Mit der physikalischen Realität der Schallwellen haben diese metaphorischen Vorstellungen und Redeweisen nur wenig zu tun, denn musikalische Klänge sind keine Objekte, die sich wo auch immer bewegen, sondern in den meisten Fällen Ereignisse, die nacheinander erklingen. Dennoch werden Klangfolgen als Bewegungen musikalischer Objekte wahrgenommen, die linear oder auf Umwegen in der Zeit voranschreiten.

Wie vermag nun eine Folge zeitlich klar gegeneinander abgegrenzter Klänge, die Empfindung einer kontinuierlichen musikalischen Bewegung auszulösen? Clarke (2001) und Gjerdingen (1999) versuchen, diese Frage unter Hinweis auf das ähnlich gelagerte Phänomen der visuellen Wahrnehmung von Scheinbewegungen („apparent motion“) zu erklären (vgl. dazu Goldstein 2002, S. 190f, 287f). Leuchten in schneller Folge benachbarte Lichter auf, so wird innerhalb bestimmter zeitlicher Grenzen eine kontinuierliche Bewegung vom einem zum anderen Licht wahrgenommen. Ist der zeitliche Abstand zwischen dem Aufblinken der Lichter zu klein, so wird das Leuchten der Lichter als gleichzeitig, ist er zu groß, als einander abwechselnd wahrgenommen.³⁴ Robert O. Gjerdingen (1999) hat das Modell eines neuronalen Netzwerkes, das die Empfindung visueller Scheinbewegungen auf neuronaler Ebene rekonstruiert, auf die Musikwahrnehmung übertragen. Ungeachtet seiner recht komplexen internen Organisation vermag das Modell die prinzipielle Möglichkeit einer musikalischen Bewegungswahrnehmung auf der Grundlage diskreter Klangereignisse zu belegen. Gegen sein Modell lässt sich allerdings einwenden, dass eine unmittelbare Verschmelzung verschiedener Töne zu einer Klangbewegung nur in sehr engen zeitlichen Grenzen stattfindet, nämlich nur dann, wenn der Abstand zwischen den Klang-Onsets kleiner als ca. 120 ms ist (Zwicker/Fastl

1987, S. 273). Nach Lakoff und Johnson (2000) liegen die grundlegenden Vorstellungsschemata oder „Image Schemas“ der Logik unseres gesamten Denkens und Schlussfolgerns zugrunde.

34 Vergleichbare Bewegungsempfindungen wurden auch im Bereich der taktilen Sinnesmodalität nachgewiesen (vgl. Goldstein 2002, S. 321).

1999, S. 274f). Erst bei diesem Tempo (500 bpm) verschwimmen melodische Linien zu einem kontinuierlichen Glissando, zu einem „sich bewegenden Klangobjekt“.

Ein besonderer Fall der Übertragung von Erfahrungen physikalischer Bewegung auf die Musik liegt beim Schlussritardando vor, bei dem die musikalische Bewegung eines Stückes zum Stillstand kommt. Die Verlangsamung des Tempos am Schluss eines Musikstückes oder, in weniger ausgeprägtem Maße, am Schluss einer melodischen Phrase oder eines formalen Abschnitts sind mit verschiedenen physikalischen Modellen erklärt worden. Sundberg und Verillo (1980) errechneten auf der Grundlage der Schlussritardandi von 24 Klavieraufnahmen, vorwiegend von Kompositionen Johann Sebastian Bachs, eine durchschnittliche Ritardandokurve, mit der die zunehmende Verlängerung der Notenwerte am Schluss der Stücke veranschaulicht wird. Während die erste Phase dieses Schlussritardandos von den verschiedenen Pianisten relativ flexibel gestaltet wird, weisen die Interpretationen der zweiten Phase einen regelmäßigen Verlauf auf, der sich mathematisch darstellen lässt. In einer Folgestudie (Kronberg/Sundberg 1987) wurde der empirisch ermittelte Ritardando-Verlauf mit der physikalischen Gesetzmäßigkeit für die Verlangsamung physikalischer Bewegungen verglichen, wie sie auch für das Gehen oder Rennen charakteristisch sind.³⁵ Beide Modelle stimmen gut miteinander überein; allerdings besitzt beim Schlusston die „musikalische Bewegung“ noch ca. 30 % der Geschwindigkeit vor Beginn des Ritardandos. Erst wenn der Verlangsamungsprozess quasi in den Schlusston hinein verlängert wird, ergibt sich die erwartete Übereinstimmung. Die musikalische Bewegung eines Schlussritardandos kommt demnach nicht mit Beginn des Schlusstones, sondern erst kurze Zeit später zum Stillstand.

Neil Todd (1995) hat unter Rückgriff auf Gleichungen der elementaren Mechanik versucht, den Ansatz von Sundberg, Verillo und Kronberg mathematisch weiter zu präzisieren und auf musikalische Accelerandi auszuweiten, wobei er sich auf die Tempoveränderungen in drei Auführungen eines Prélude von Frédéric Chopin stützt. Todd stellte eine weitgehende Übereinstimmung zwischen physikalischem Modell und empirischen Daten fest, weist jedoch zugleich auf die Grenzen der Analogie von physikalischer und musikalischer Bewegung hin:

35 Friberg und Sundström (1999) haben Ritardandokurven mit dem Schrittverlauf von Sprintern verglichen, die erst in einem auditiv vorgegebenen Tempo (340 ms und 250 ms pro Schritt) rennen und dann bis zum Stillstand abbremsen sollten. Auch hier ergab sich eine gute Übereinstimmung; allerdings setzte der Bremsvorgang der Sprinter abrupt ein als die musikalischen Schlussritardandi.

„There is an essential difference between the motion of an inanimate particle and the motion of musical expression. In one case the motion is completely determinate but in the other the accelerations vary from segment to segment in a structured but nondeterminate manner for purpose of communication. This is the difference between physics, where forces are given and the problem is to account for an observed trajectory, and cybernetics where we are concerned with the inverse problem of attempting to infer the goals or intentions underlying a given trajectory“ (Todd 1995, S. 1945).

Neben der metaphorischen Verknüpfung von physikalischer und musikalischer Bewegung kommt es vermutlich zu unmittelbaren assoziativen Verknüpfungen zwischen rhythmischen Klangstrukturen in der Musik und Klangstrukturen, die bestimmte Alltagsbewegungen begleiten. Die Bewegungen von Objekten und die Empfindung der eigenen Körperbewegung (Gehen, Springen) sind in vielen Lebensbereichen mit bestimmten Klängen und Geräuschen verbunden. Ein bewegtes Objekt wird zu bestimmten Zeitpunkten seiner Bewegung selbst zur Klangquelle, z.B. immer dann, wenn es in seiner Bewegung durch ein Hindernis gehemmt wird. So bezeugen die regelmäßigen Geräusche des Aufsetzens der Füße beim Gehen oder der Bodenkontakte eines hüpfenden Balls die zeitliche Struktur der hierdurch ausgelösten Richtungsänderungen. Die Verknüpfung von Bewegungsstrukturen und Klangstrukturen wird nun assoziativ auf ähnliche musikalische Strukturen übertragen. Entscheidend für die assoziative Verknüpfung von Musik und den rhythmischen Bewegungen realer Objekte sind drei Aspekte: das Tempo der Bewegung, der Klang des Bewegungsgeräusches (aufgrund der beteiligten Kräfte und Materialien) sowie ein charakteristisches Bewegungsmuster, eine Bewegungsstruktur, ein Bewegungsrhythmus. Decken sich diese Aspekte in ausreichendem Maße mit wahrgenommenen musikalischen Klangfolgen, so können durch Musik die entsprechenden Bewegungsassoziationen oder -empfindungen ausgelöst werden.

Es gibt eine Reihe von Partizipien, die sich zur Beschreibung der musikalischen Bewegungsstrukturen gut eignen, z.B. „pulsierend“, „gehend“, „springend“, „schwingend“, „schaukelnd“, „hüpfend“, „stampfend“, „hämmernd“, „klopfend“, „pochend“, „hinkend“, „stolpernd“ (vgl. Gabrielsson 1973c). Die meisten dieser Bewegungsarten besitzen eine klar identifizierbare rhythmische Struktur und einen bevorzugten Tempobereich.³⁶ Die entsprechenden Verben beschreiben entweder menschliche Fortbewegungsarten oder Bewegungsarten von physikalischen Objekten. Durch die Verwendung von Partizipien wird allerdings die problematische Frage umgangen, *was* sich da in der Musik bewegt.

³⁶ Ein Sonderfall ist das Adjektiv „fließend“, da es sich hier offensichtlich nicht um eine rhythmische, sondern um eine kontinuierliche Bewegung handelt.

Körperbewegungen von Musikern und Hörern

Wenn Geräusche als Hinweise auf bestimmte Bewegungen der Klangquellen wahrgenommen werden, so liegt es nahe, dass musikalische Klangfolgen als auditive Hinweise auf die mechanischen Bewegungen von Musikinstrumenten und auf die Spielbewegungen der Musiker dienen. Den musikalischen Bewegungsempfindungen liegen daher, so die These von Patrick Shove und Bruno Repp, in letzter Konsequenz immer die Spielbewegungen der Musiker zugrunde:

„[...] the performer is the ultimate source of the movement perceived. True, the abstract image of a sonic object rising and falling in pitch space bears little, if any, resemblance to a pianist's fingers and hands moving across the surface of the keyboard. And if one takes into account the physical and mechanical differences between instruments, the resemblance becomes even more remote. Yet, without the performer's rhythmic movements (real, imagined or computer-simulated), the perception of melodic motion (for that matter, any motion) would be impossible. Simply put, patterned articulatory movements create patterned sequences of tones. In other words, articulatory movements are sound-structuring movements. The motion one attributes to a succession of tones - including its pacing, its character, even its directionality - belongs first and foremost to the performer. That a listener reports hearing a sonic object in motion, rather than a performer, reflects the listener's perceptual attitude towards the musical event“ (Shove/Repp 1995, S. 61).

Auch wer beim Musikhören nicht mehr die Spielbewegungen der Musiker vor Augen hat, nimmt demnach musikalische Linien als Bewegungen in einem „musikalischen Raum“ wahr, weil in ihnen Informationen zu Eigenschaften und Bewegungen der Klangquellen enthalten sind. Allerdings hängt die Verknüpfung von rhythmischen Klang- oder Geräuschfolgen und Bewegungen bzw. Bewegungsempfindungen von Lernprozessen ab. Erst wenn ein Hörer weiß, dass immer jemand Tasten drückt, wenn ein Klavierstück erklingt, kann diese assoziative Verknüpfung entstehen.

Der Zusammenhang zwischen den Spielbewegungen der Musiker und den Eigenschaften der produzierten Klangfolgen sowie deren rhythmischen Strukturen liegt auf der Hand. Schnellere Bewegungen führen in der Regel zu schnelleren Klangfolgen und stärkere Bewegungen zu lauterem Klängen. Zusätzlich zu den Spielbewegungen klopfen viele Musiker einen periodischen Grundschlag mit Händen oder Füßen oder wiegen ihren Oberkörper oder Kopf im Rhythmus der Musik. Ähnlich wie die Musiker bewegen sich auch viele Hörer zur Musik, sie schlagen die Grundschlagfolge, seltener konkrete rhythmische Muster mit dem Fuß oder den Händen, schnippen mit

den Fingern oder führen weiter ausgreifende Bewegungen mit Kopf, Armen und Oberkörper aus. Diskrete Informationen über die Timing-Struktur der Rhythmen werden so in strukturierte Körperbewegungen umgesetzt.

Die bevorzugte Frequenz von Körperbewegungen korrespondiert in vielen Fällen mit der Dauer bzw. den Interonsetintervallen bestimmter musikalischer Einheiten. Beide Vorgänge spielen sich grob in folgenden zeitlichen Bereichen ab (vgl. Fraisse 1982, S. 151f, Iyer 2002, S. 392f):

1. Im Bereich von einer bis zehn Sekunden (0,1 bis 1 Hz) liegt die Dauer von Atembewegungen, Armbewegungen, dem Wiegen des Oberkörpers sowie die Länge einer musikalischen Phrase. Die Länge einer musikalischen Phrase hängt beim Gesang und bei Blasinstrumenten unmittelbar mit dem Lungenvolumen der Musikern bzw. Sänger zusammen; aber auch andere Instrumentalisten „atmen“ in Rhythmus der musikalischen Phrasen.
2. Im Bereich von einer Drittel bis einer ganzen Sekunde (333 ms bis 1 s, 1 bis 3 Hz oder 60 bis 180 bpm) spielen sich das Tempo des menschlichen Herzschlags, Saug- und Kaubewegungen, die wichtigsten körperlichen Fortbewegungsarten (Gehen, Rennen) sowie die Nickbewegung des Kopfes ab. Dieser Tempobereich korrespondiert mit dem bevorzugten Tempo des Grundschlages in der Musik.
3. Im Bereich von 100 bis 333 ms Dauer (3 bis 10 Hz oder 180 bis 600 bpm) spielen sich Sprech- und Zungenbewegungen sowie Handgesten ab. Dem entspricht die schnellste Unterteilung des Grundschlag bzw. der schnellste notierte Rhythmuswert.³⁷
4. In noch kleineren Zeitbereichen (10 bis 60 Hz oder 17 bis 100 ms) ereignen sich schnellste Körperbewegungen: Die Dauer mancher Phone-me beim Sprechen und die schnellen Wirbel zwischen Fingern entsprechen Mikrotiming-Abweichungen und Verzierungen in der Musik.

Van Noorden und Moelants (1999) verknüpfen den von ihnen ermittelten bevorzugten Tempobereich von ca. 120 bpm (ca. 500 ms IOI) mit dem physikalischen Modell der gedämpften Resonanz. Ihrer Studie legten sie die Daten aus verschiedenen Mitklopfstudien (Parncutt 1994, Handel/Oshinsky 1981) sowie aus einer eigenen umfangreichen Erhebung zugrunde, bei der Versuchsteilnehmer zu verschiedenen Musikausschnitten (flämische Polyphonie des 16. Jahrhunderts, französische Barockmusik, romantische Klaviermusik, Jazz und Pop) sowie zur Musik von zufällig ausgewählten Radiostationen mitklopfen sollten. Die Häufigkeitsverteilung der Tempi aller Klopffolgen korrelierte in hohem Maße ($r = .971$) mit einer Re-

³⁷ Eine Analogie von schnelleren Klangeinheiten und den Bewegungen beim Sprechen belegt die Metapher, nach der manche Musiker mit ihren Händen und Fingern „sprechen“.

sonanzkurve, deren Resonanzspitze bei $f_0 = 2,193$ Hz (also 456 ms bzw. 132 bpm) und deren Dämpfungsfaktor, durch den die Steilheit der Kurve bedingt wird, bei $\beta = 0,5$ lag. Ob und auf welche Weise es tatsächlich zu einer körperlichen „Resonanz“ der Musik beim Menschen kommt, bleibt allerdings bei diesem Ansatz offen.

Um den Zusammenhang von musikalischen Klangsstrukturen und Körperbewegungen empirisch nachzuweisen, schlagen Shove und Repp vor, den Hörer als einen Übersetzer (transducer) zu begreifen, der aus den diskreten Ereignissen der Musik eine kontinuierliche räumliche Bewegung rekonstruiert. „The principal methodology for demonstrating that music does convey movement information is the reconstitution of an analogous spatial movement by a human listener. The listener’s body thus act as a transducer, a kind of filter for the often impulse-like coding of musical movement“ (Shove/Repp 1995, S. 65).

Shove und Repp verweisen auf entsprechende Selbstversuche von Eduard Sievers (1924) und Gustav Becking (1928) sowie auf Untersuchungen von Alexander Truslit (1938) und Manfred Clynes (1977). So versuchte Truslit aufgrund von umfangreichen empirischen Untersuchungen den Nachweis zu erbringen, dass es Bewegungen sind, die beim Komponieren form- und melodiebildend wirken (Truslit 1938, S. 34 und III.4).³⁸ Dabei wählte er den indirekten Weg, die musikalischen „Urbewegungen“ durch Mitbewegungen seiner Versuchsteilnehmer zu ermitteln, die er in sog. Bewegungskurven festhielt. Truslit schreibt: „So selten wir das Phänomen des spontanen Umsetzens einer Bewegung in Musik bewusst beobachtet und klar geschildert sehen, so oft finden wir den umgekehrten Vorgang der Auslösung einer Bewegungsempfindung durch einen Ton oder eine Melodie beschrieben“ (Truslit 1938, S. 34). Die Mitbewegungen werden bei Truslit zu Musikbeispielen mit gestreckten, parallelen Armen ausgeführt, so dass der ganze Oberkörper beteiligt ist. Sie folgen zumeist dem Tonhöhenverlauf sowie der metrischen Gestaltung der Stücke. Dabei gibt es neben den lokalen Bewegungen der Hände (kleinere Kurven) größere Bewegungen des gesamten Oberkörpers. Die musikalische Bewegung verläuft zumeist in Kurven; eckige Bewegungen sind untypisch. Truslit unterscheidet dabei offene, geschlossene und gewundene Bewegungskurven.

38 Truslits Untersuchungen stehen im Kontext der musikpsychologischen Synästhesie-Forschungen in der ersten Hälfte des 20. Jahrhunderts. Neben Bewegungen thematisiert er auch die visuelle Wahrnehmung von Musik, die er „Synopsis“ nennt (Truslit 1938, S. 35).

Die auditiv-motorische Theorie der Rhythmuswahrnehmung

Der englische Musikpsychologe Neil Todd sieht den Zusammenhang zwischen dem Rhythmus von Klangfolgen und Körperbewegungen in einem sog. „sensory-motor“-Mechanismus, einem Mechanismus der durch die Sinne geleiteten Bewegung (Todd 1999, Todd/O’Boyle/Lee 1999, Todd/Lee/O’Boyle 2002). Da sich Todds Modell explizit auf neurophysiologische Erkenntnisse stützt, stellt es einen vielversprechenden Ansatz dar, die unmittelbare Verbindung von Musikwahrnehmung und Bewegungsempfindung zu erklären.

Todd geht davon aus, dass akustische Stimuli unmittelbar bestimmte Bewegungsschemata auslösen, auf deren Grundlage Rhythmen wahrgenommen werden. Die kortikale Bewegungsplanung erhält, so seine These, einen Input vom auditorischen Kortex, so dass eine zeitliche Synchronisation mit auditorischen rhythmischen Sequenzen erfolgt. Es besteht demnach ein vergleichbarer Zusammenhang zwischen Hören und Eigenbewegung wie zwischen visueller Bewegungswahrnehmung und Eigenbewegung, z.B. bei visuell gesteuerten Zielbewegungen.

Der Zusammenhang von Zielbewegungen, z.B. dem Ergreifen eines Gegenstandes mit der Hand, und einem kontinuierlichen visuellen Input ist bereits relativ gut untersucht. Bei der Bewegungssteuerung und -ausführung wirken zahlreiche Instanzen des zentralen Nervensystems zusammen, die jeweils sensorische Informationen über die Körperstellung und über die Umwelt erhalten (vgl. Birbaumer/Schmid 2003, S. 252-254, 263-276). So arbeitet die zielorientierte Bewegungsplanung im seitlichen Kortex eng mit den für die Stützmotorik zuständigen Bewegungsprogrammen im Kleinhirn zusammen. Dabei kommt es zu antizipierenden Ausgleichsbewegungen der Stützmotorik. Bestimmte Neurone im primären visuellen Kortex reagieren nicht nur auf die räumliche Anordnung, sondern auch auf Reizmerkmale wie Ausrichtung, Bewegung und Größe. Bewegungserkennende Neurone besitzen dynamische, d.h. zeitlich veränderliche rezeptive Felder (Birbaumer/Schmid 2003, S. 396f, Goldstein 2002, S. 288ff). Neurone, die Bewegungen wahrnehmen können, lassen sich als Kombination eines räumlichen und eines zeitlichen Filters verstehen, welche durch einen bestimmten Raumausschnitt (räumliche Frequenz) und einen bestimmten Zeitausschnitt (zeitliche Frequenz) charakterisiert sind. Der Output eines Neuronenverbands von rezeptiven Feldern kann somit als ein räumlich-zeitliches Leistungsspektrum modelliert werden.

Der auditorische Kortex lässt sich nun laut Todd ganz ähnlich verstehen wie der visuelle Kortex. Auch hier gibt es Neurone, die dynamische

„räumlich-zeitliche“ rezeptive Felder besitzen.³⁹ Die „räumliche“ Dimension entspräche bei den auditiven rezeptiven Feldern der Tonhöhe bzw. der erregten Stelle auf der Basilarmembran; die zeitliche Dimension der rhythmischen Organisation. Auf diese Weise entsteht im auditorischen Kortex ein räumlich-zeitliches Energiespektrum des auditorischen Inputs („auditory image flow“). Da die Modellierung eines solchen komplexen Systems von dynamischen Neuronen extrem aufwendig ist, interessiert sich Todd ausschließlich für den in der Rhythmuswahrnehmung maßgeblichen zeitlichen Aspekt, den er mit Filterbänken modelliert.⁴⁰

„In sum then, according to the sensory-motor theory of rhythm perception, an external source is represented in the sensory system (auditory or visual cortex) as a power spectrum. An internal model (in the cerebellum) of the plant dynamics (a torsional pendulum in the case of beat tracking) enables the controller to interpret the sensory image in the form of a *sensory-motor* image“ (Todd/O’Boyle/Lee 1999, S. 122).

Einer Repräsentation der zeitlichen Informationen der Sinnesreize in Form eines Leistungsspektrums steht demnach eine Bewegungs-Repräsentation des Körpers („motor image“) gegenüber: „The first is a sensory representation (either auditory or visual) of temporal information in the stimulus signal in the form of a power spectrum of the sensory image over a range of temporal frequencies between about 0.5 Hz and 32 Hz. The second is a motor representation of the musculoskeletal system in the form of a motor image of the body“ (Todd/O’Boyle/Lee 1999, S. 119). Nach Todd, O’Boyle und Lee wird die musikalische Rhythmus Erfahrung somit von zwei Repräsentationen vermittelt. Beide Repräsentationen können im Konzept der sensorisch geleiteten Handlung miteinander verbunden werden. „A central principle of this sensory-motor theory is that if the spatiotemporal form of certain stimuli are matched to the dynamics of the motor system, then they may evoke a motion of an internal representation, or motor image, of the corresponding synergetic elements of the musculoskeletal system, *even if the musculoskeletal system itself does not move*“ (Todd 1999, S 119).

Todd geht also davon aus, dass bei der Rhythmuswahrnehmung immer der gesamte menschliche Körper beteiligt ist – selbst dann, wenn tatsächlich keine Bewegungen ausgeführt werden. Er beschränkt sich in seinen

39 Todd, Lee und O’Boyle (2002, 29f) führen für diese These Belege aus der neueren neuropsychologischen Forschungsliteratur an.

40 Das Modell geht zunächst von zwei Filterbänken aus, dem eine weitere Verarbeitungsstufe angeschlossen ist: 1. Low-Pass-Filter lokalisieren Ereignisse in der Zeit; sie richten sich auf die zeitliche Gruppierung; 2. Band-Pass-Filter richten sich auf bestimmte zeitliche Intervalle. Im sekundären auditorischen Kortex werden sodann die eintreffenden Informationen zu Merkmalen fusioniert (vgl. im Detail Todd/O’Boyle/Lee 1999 und Todd/Lee/O’Boyle 2002).

Modellen allerdings bislang aus pragmatischen Gründen auf den Zusammenhang zwischen Grundschlagsinduktion und der periodischen Gehbewegung, die ja auch grundlegend für das Tanzen ist. Da selbst das Gehen noch eine sehr komplexe Bewegung ist (vgl. Orlovskij/Deliagina/Grillner 1999), wird sie auf ein minimales System reduziert: auf die vor- und zurückpendelnde Bewegung des Beckens. Mit einigen Modifikationen ließe sich das Modell jedoch auch auf komplexere Körperbewegungen ausweiten.

Unwillkürliche Muskelbewegungen beim Hören lauter Musik

Nach dem Modell von Todd sind bei der Rhythmuswahrnehmung motorische Schemata involviert. Allerdings hat die Beteiligung der Bewegungsschemata nicht automatisch Körperbewegungen zur Folge. Zwar können auditive Rhythmen Bewegungen und Tanz stimulieren, sie müssen es jedoch nicht. Zudem können die unmittelbaren Körperbewegungen durch bewusste Bewegungsintentionen und gelernte Bewegungsschemata modifiziert werden. So ist es beim Tanz jedem Menschen freigestellt, eigenständige rhythmische Strukturen in die polyrhythmische Textur von musikalischen und körperlichen Bewegungen einzubringen. Eine relative Unabhängigkeit zwischen den Zeitstrukturen von Musik und Tanz ist im Ballett, aber auch im Volkstanz nichts Ungewöhnliches. Der Volksmusikforscher Felix Hoerburger spricht angesichts mancher volksmusikalischer Tanzformen gar von der Möglichkeit eines Tanzens „gegen“ die Musik (Hoerburger 1986, S. 43ff) – die Musik bleibt dabei jedoch Stimulation und Anregung für die körperliche Aktivität.

Dennoch reißen die Spekulationen um eine direkte und unmittelbare Stimulation von Körperbewegungen durch Musik nicht ab. Eine mögliche Option für einen Mechanismus, der die Rhythmuswahrnehmung zumindest bei hohen Lautstärken unmittelbar in Körperbewegungen umsetzt, wurde von Todd in die Diskussion eingebracht (Todd 1999, Todd/Cody 2000, Todd 2001). Bestimmte Muskeln, vor allem der Nacken- und Schultermuskulatur, lassen sich durch laute Töne reizen (Colebatch/Halmagyi/Skuse 1994). Die entsprechenden elektrischen Aktivierungspotentiale der Muskelnerven sind mit einem Elektromyogramm (EMG) messbar. Zur Überprüfung der Leistungsfähigkeit des Gleichgewichtsorgans werden Messungen des sog. myogenischen, vestibular erzeugten Potentials (myogenic vestibularly evoked potentials, MVEP) bei verschiedenen Nackenmuskeln durchgeführt. Da die unwillkürlichen Muskelpotentiale auch bei Patienten mit Nerventaubheit, nicht aber bei Patienten ohne Gleichgewichtsorgane gemessen werden konnten, wird davon ausgegangen, dass die Schallauf-

nahme nicht durch das Gehör (Cochlea), sondern durch das Gleichgewichtsorgan, insbesondere durch die Macula sacculi, erfolgt.

Todd (2001) sowie Todd und Cody (2000) kamen durch experimentelle Messungen von MVEPs zu folgenden Ergebnissen:⁴¹ Das MVEP auf kurze Klangimpulse ist frequenzabhängig mit einer besten Ansprache zwischen 300-350 Hz. Die Klänge müssen sehr laut sein: Erst ab einem Schallpegel von 90 dB⁴² treten MVEPs auf. Nicht nur kurze Impulse, sondern auch durchgehende Klänge können die Macula sacculi reizen. MVEPs können daher von verschiedenen musikalischen Stimuli ausgelöst werden, z.B. von der Bass Drum in lauter Dance Music, aber auch vom Klang der eigenen Stimme beim lauten Singen. Todds Theorie konnte allerdings noch nicht in musikalischen Hörzusammenhängen überprüft werden.

Auditive Bewegungswahrnehmung und musikalische Bewegung

Es ist fraglich, ob sich die musikalische Bewegungsempfindung auf die körperliche Eigenbewegung oder auf unwillkürlich ausgelöste Bewegungen reduzieren lässt. In der musikalischen Bewegungswahrnehmung wird ja nicht unbedingt der eigene Körper, sondern die musikalische Struktur oder wie auch immer geartete musikalische „Objekte“ als bewegt empfunden. Nach Clarke (2001) spielen grundlegende psychologische Mechanismen der visuellen und auditiven Bewegungswahrnehmung auch bei der Wahrnehmung musikalischer Strukturen eine Rolle. Clarke wendet sich vehement gegen die Ansicht, die musikalische Bewegungsempfindung lasse sich ausschließlich als Metapher verstehen. Vielmehr werde beim Musikhören auf jene Mechanismen zurückgegriffen, die dem Menschen die Identifikation von bewegten Klangquellen in seiner Umwelt erleichtern, um die fiktionalen Bewegungen der Musik wahrzunehmen.

„[...] since sounds in the everyday world specify (among other things) the motivational characteristics of their sources, it is inevitable that musical sounds will also specify movements and gestures of their actual physical production [...] and also the fictional movements of the virtual environment [...] which they conjure up. [...] this relationship is truly perceptual rather than metaphorical, symbolic or analogical“ (Clarke 2001, S. 222).

41 Todd/Cody (2000) untersuchten die MVEPs bei 10 Versuchspersonen, die Stimuli zwischen 90 und 120 dB (in 5 dB-Schritten) ausgesetzt waren. Die Stimuli bestanden aus einer Folge von perkussiven Klängen (Tempo 2,5 Hz, also 150 bpm), die einem Techno-Stück entnommen wurden. Der Reiz wurde aus Messgründen nur auf einem Ohr dargeboten.

42 90 dB entspricht dem Lärm einer einfahrenden U-Bahn.

Beim visuellen Wahrnehmen von Bewegungen stehen Veränderungen des Objektes im Raum in einem komplexen Wechselverhältnis mit den Bewegungen des eigenen Körpers (Augenbewegungen, Kopfbewegungen, körperliche Fortbewegung). Die Informationen zur Eigenbewegung und zur Fremdbewegung müssen ständig miteinander abgeglichen werden, um zu einer realistischen Einschätzung der Bewegung von Objekten im Raum zu gelangen. Der amerikanische Psychologe James Jerome Gibson hat sich intensiv damit beschäftigt, wie ein Beobachter, der sich bewegt, sich selbst und seine Umgebung visuell wahrnimmt. Der hierbei entstandene Ansatz einer ökologischen Wahrnehmungspsychologie untersucht das Sehen in der natürlichen Umwelt unter Berücksichtigung sowohl von Umgebungsinformationen als auch der Fortbewegung der wahrnehmenden Person (Gibson 1979). Eine Vorstellung des „Schnappschuss-Sehens“ lehnt Gibson als unangemessen ab. Stattdessen erforschte er sehr genau die Informationen, die aus dem Phänomen des sog. „optischen Fließens“ über die Umwelt gewonnen werden. Bei einer kontinuierlichen Bewegung des Beobachters bleibt nur der räumliche Zielpunkt der eigenen Bewegung als Zentrum des Auseinanderfließens unbewegt, während alle anderen optischen Informationen seitlich „wegfließen“.

Die auditive Wahrnehmung bewegter Klangquellen wird ebenfalls durch die Eigenbewegungen des Hörers stark verbessert. Grundlage der auditiven Bewegungswahrnehmung ist die Ortung von Klangquellen im Raum (vgl. Goldstein 2002, S. 418-429). Die Richtung, aus welcher der Schall auf die beiden Ohren trifft, wird aufgrund von Zeit- und Intensitätsunterschieden bestimmt; die Entfernung der Schallquelle vom Hörer wird vom Gehör aufgrund von Veränderungen der Lautstärke (bei einer Verdopplung des Abstands zur Schallquelle sinkt die Schallenergie auf ein Viertel des Ausgangswertes), durch die Klangfarbe (stärkere Absorption der hohen Frequenzen bei größerem Abstand) sowie in geschlossenen Räumen durch das Verhältnis zwischen der Lautstärke des Direktschalls und des reflektierten Schalls (mit frequenzabhängigen Absorptionen) ermittelt. Ob die Schallquelle vor oder hinter dem Hörer liegt, wird vorwiegend aufgrund von Unterschieden der Klangfarbe entschieden, die durch Resonanzen und Reflektionen an Kopf und Ohrmuschel entstehen. Beim Lokalisieren von Klangquellen helfen zudem Bewegungen des Kopfes, wodurch die Bereiche größter Empfindlichkeit genutzt, nicht eindeutige Richtungsinformationen in eindeutige geändert sowie vermutlich die Externalisierung der Hörereignisse unterstützt werden. Kopf- und Körperbewegungen stellen demnach ein wichtiges Mittel zur auditiven Lokalisierung von Klangquellen dar. Die auditive Bewegungswahrnehmung lässt sich somit wie die visuelle Bewegungswahrnehmung als Kombination von Fremd- und Eigenbewegungen verstehen.

Clarke weist anhand mehrerer Musikbeispiele darauf hin, dass Mechanismen der auditiven Bewegungswahrnehmung auch in Musik zu Bewegungsempfindungen führen können. Beim Näherkommen einer Klangquelle nimmt die Lautstärke zu, Hallanteile nehmen gegenüber dem Direktschall ab und durch die Zunahme hoher Frequenzen wächst das Empfinden der klanglichen Präsenz. Bewegt sich die Klangquelle hinter verschiedenen Hindernissen im Raum, so verändert sich die Lautstärke sowie aufgrund der frequenzabhängigen Dämpfung das Frequenzspektrum, wobei hohe Frequenzen stärker gedämpft werden. Beide Effekte können auch mit musikalischen Mitteln (Instrumentierung, Lautstärke) erzeugt werden.⁴³

Bei der auditiven Bewegungswahrnehmung bleibt bisweilen unklar, ob sich die Schallquelle bewegt oder der Hörer. Diese Unklarheit besteht, so Clarke, auch in der Musikwahrnehmung. Zusätzlich bleibt bei musikalischen Bewegungen weitgehend unklar, wie das Objekt beschaffen ist, das sich da bewegt, und wo die Bewegung stattfindet. Dies ist jedoch, so Clarke weiter, im künstlerischen Bereich nichts Ungewöhnliches. So werden in der bildenden Kunst vielfach Objekte dargestellt, die in der Realität nicht vorkommen. Bisweilen wird ein fiktionaler Raum eröffnet, der nur entfernt an die physikalische Realität erinnert oder die physikalischen Gesetze sogar negiert, wie z.B. in den Bildern von M.C. Escher. Auch Musik kann, so Clarke, eine virtuelle Umgebung schaffen, vergleichbar dem virtuellen Raum von Gemälden, in der manche Merkmale der physikalischen Realität außer Kraft gesetzt sind: „[...] sounds can specify a virtual domain that both abides and defies the normal laws of physics“ (Clarke 2000, S. 217).

Die Auffassung eines virtuellen Raumes, in dem die musikalische Bewegung zum Teil unabhängig von physikalischen Gesetzen vonstatten geht, wird durch die Ergebnisse eines Hörversuchs der israelischen Musikpsychologen Zohar Eitan und Roni Y. Granot (2004a, 2004b) gestützt. Ziel der Untersuchungen war es, genauere Aussagen über Analogien zwischen Bewegungsempfindungen und musikalischen Parametern zu gewinnen. Eitan und Granot beauftragten 77 Versuchsteilnehmer, sich einen menschlichen Cartoon-Charakter ihrer Wahl vorzustellen und diese Zeichentrickfigur in ihrer Vorstellung zu zahlreichen kurzen melodischen Mustern sich bewegen zu lassen. Sodann sollten die Versuchsteilnehmer Angaben darüber machen, wie sich ihr vorgestellter Charakter bei den Musikbeispielen jeweils bewegt hat. Im Einzelnen wurden abgefragt: der Bewegungstyp (z.B. gehen, rennen, springen), die vertikale und horizontale Richtung der Bewegung (nach oben oder unten, nach rechts oder links), den Wechsel des Ab-

43 Clarke stützt sich in seiner Argumentation auf folgende Musikbeispiele: 1. Interlude zwischen der 2. und 3. Szene des dritten Aktes von Alban Bergs „Wozzeck“; 2. Fatboy Slim „Build It up, Tear It down“ (aus: „You’ve Come a Long Way, Baby“, Skint Records 11, 1998); 3. Mozarts Streichquintett C Dur (KV 515), Durchführung des ersten Satzes (T. 152-205).

stands zum Betrachter (entfernd oder näherkommend), den Wechsel der Bewegungsgeschwindigkeit, die mögliche Einwirkung einer äußeren Kraft, die Art dieser Krafteinwirkung (unterstützend, hemmend, unabhängig) sowie das Energielevel der Bewegungen. Innerhalb der Klangbeispiele wurde jeweils einer der folgenden Parameter verändert: die Lautstärke (crescendo oder diminuendo), die Richtung der Tonhöhenbewegung (aufwärts oder abwärts), die Intervallgröße (größer bzw. kleiner werdend), die Länge der Interonsetintervalle (accelerando oder ritardando), die Geschwindigkeit der melodischen Muster sowie die Artikulation (tenuto bzw. staccato).

Erstaunlich ist die große Konsistenz der Angaben aller Versuchspersonen, so dass die Analogie zwischen musikalischen und kinästhetischen Wahrnehmungen weitgehend als intersubjektiv verbindlich angesehen werden kann. Es stellte sich heraus, dass jeder musikalische Parameter mit mehreren Bewegungseigenschaften in Zusammenhang steht. Die Veränderungen musikalischer Parameter stehen zudem in keinem streng symmetrischen Verhältnis zu analogen Veränderungen der Bewegungsempfindung, d.h. eine umgekehrte Veränderung beim musikalischen Parameter führt vielfach nicht zu umgekehrten Bewegungsassoziationen. So entspricht ein Diminuendo einer fallenden Bewegung, aber umgekehrt wird ein Crescendo nicht mit ansteigenden Bewegungen assoziiert; auch der Zusammenhang zwischen musikalischer Intensivierung und der empfundenen Veränderung des Bewegungstempos ist asymmetrisch.

„In general, while musical abatements (pitch descents, diminuendo, ritardando) are strongly associated with spatial descents, musical intensifications (crescendo, accelerando, even pitch rise itself) are not as strongly associated with intensifications. Similarly, while musical intensifications (pitch ascents, crescendo, accelerando, increasing motivic pace, sequential melodic progressions) are generally associated with increasing velocity, musical abatements (pitch descents, diminuendo, decreasing motivic pace) are not generally associated with decreasing velocity. [...] Opposite ‚directions‘ in this space do not present symmetrical contrasts - pitch rise is not the opposite of pitch fall, and crescendo not the opposite of diminuendo - but rather, each stresses different spatial and kinetic associations“ (Eitan/Granot 2004a, S. 7f).

Eine eindeutige Verknüpfung von musikalischer Bewegungsempfindung und den Bewegungen physikalischer Objekte ist angesichts dieser experimentellen Ergebnisse fragwürdig. Vielmehr wird Clarkes Auffassung gestützt, dass sich musikalische Bewegungen in einem fiktionalen, künstlerischen Raum ereignen, in dem andere Gesetze walten als in der physikalischen Realität.

Bewegung, Erregung, Emotion

Musik kann als bewegt empfunden werden und als Stimulans für Körperbewegungen und Tanz dienen. Darüber hinaus kann Musik bewegend sein im Sinne einer inneren, emotionalen Bewegung. Sie kann nicht nur ästhetisches Wohlgefallen oder Missfallen auslösen, sondern zugleich den Grad der Erregung oder Beruhigung beeinflussen, den ein Hörer in sich spürt, sowie dessen Stimmung prägen. Gibt es dabei einen direkten und eindeutigen Zusammenhang zwischen der Art der Rhythmusgestaltung und der emotionalen Erfahrung von Musik?

In der Psychologie werden die Ausdrücke „Erregung“, „Affekt“, „Gefühl“ und „Stimmung“ nicht einheitlich verwendet. Sloboda und Juslin (2001) schlagen folgende Unterscheidung vor: Der Ausdruck „Erregung“ („arousal“) wird zum einen auf die Intensität der emotionalen Erfahrung und des emotionalen Ausdrucks bezogen, zum anderen auf den Grad der Aktiviertheit im Nervensystem. Dagegen bezieht sich „Affekt“ auf die positive oder negative Wertigkeit emotionaler Erfahrung; vielfach wird Affekt als ein grundlegender, primitiverer Mechanismus angesehen, der den konkreten Emotionen zugrunde liegt. Gefühle sind eher kurz, lassen sich direkt auf einen Stimulus zurückführen und werden vielfach von bestimmten Gesichtsausdrücken begleitet. Dagegen ist eine Stimmung ein länger anhaltender Zustand, der nicht unbedingt auf einen direkten Stimulus zurückzuführen ist und nicht unbedingt einen expressiven Ausdruck findet. Bezüglich des umfassenden Begriffs der Emotion hat sich in der Psychologie inzwischen ein Definitionskonsens durchgesetzt (Kleinginna/Kleinginna 1981). Danach wird unter Emotion die komplexe Interaktion einer Reihe von subjektiven und objektiven Faktoren verstanden, durch die affektive Erfahrungen wie Erregung, Freude oder Missfallen erzeugt werden. Diese affektiven Erfahrungen führen zu kognitiven Zuschreibungsprozessen, zu einer physiologischen Anpassung an die erregenden Bedingungen sowie oftmals (nicht immer) zu einem expressiven, anpassenden oder zielgerichteten Handeln. Unklar bleibt die exakte Reihenfolge sowie der Anfangspunkt einer emotionalen Sequenz, die aus der Wahrnehmung bestimmter Ereignisse, deren kognitiver Bewertung, körperlichen Reaktionen, subjektiven emotionalen Erfahrungen und Ausdrucksverhalten besteht.

Es ist problematisch, wenn nicht unmöglich, die subjektiven emotionalen Erfahrungen beim Musikhören mit den musikalischen Strukturen in eine eindeutige kausale Verbindung zu setzen. Dagegen lässt sich ein Zusammenhang zwischen musikalischen Strukturen und den Veränderungen des momentanen Erregungszustands bzw. korrespondierenden physiologischen Reaktionen durchaus nachweisen. Ob Musik eine oder mehrere der elementaren Emotionen wie Freude, Ärger, Traurigkeit, Angst oder Ekel

auszulösen vermag, hängt von zahlreichen, weitgehend von der klanglichen Struktur unabhängigen Faktoren ab. Ob sie Anspannung und Energie vermittelt, oder aber Entspannung und Ruhe, steht dagegen unabhängig von individuellen Unterschieden der Hörer mit bestimmten musikalischen Parametern im Zusammenhang.

Waterman (1996) unterscheidet daher die Intensität des ausgelösten Affektes (sog. intrinsische Emotion) von dessen emotionalem Inhalt (sog. extrinsische Emotion), der dem Erleben bestimmter Objekte, Handlungen oder Ereignisse zugeschrieben wird. In einem Hörversuch stellte er fest, dass von Versuchsteilnehmern einheitlich bestimmte Zeitpunkte im Verlauf eines Musikstückes als sehr erregend erlebt werden. Diesen Passagen werden allerdings ganz unterschiedliche Emotionen zugeschrieben. Die erlebten intrinsischen Emotionen beziehen sich auf den zeitlichen Verlauf der musikalischen Strukturen, auf Höhepunkte, Spannungen, Entspannungen usw., aber auch auf ganz konkrete Merkmale wie Synkopen, enharmonische Wechsel u.a. Die Strukturen, an denen sich starke Gefühle – bis hin zu körperlichen Manifestationen wie Zittern, Gänsehaut und Frösteln (sog. Thrills, vgl. Sloboda 1991) – anknüpfen, hängen eng mit der Schaffung, Bestätigung oder Störung musikalischer Erwartungen zusammen. Sloboda und Juslin (2001) weisen darauf hin, dass jedes musikalische Detail bestimmte Erwartungen bestätigen oder stören kann, wobei die Ereignisdichte, die Stellung eines musikalischen Ereignisses innerhalb der Gesamtstruktur eines Stückes und eine mögliche Asynchronizität der verschiedenen musikalischen Ebenen berücksichtigt werden muss. Die Bestätigung einer Erwartung auf einer bestimmten musikalischen Ebene kann mit einer Erwartungsstörung auf einer anderen einhergehen. Da die Erwartungsbestätigungen bzw. -verletzungen vielfach unbewusst erfolgen, nutzen sie sich beim wiederholten Hören eines Musikstückes nicht ab, sondern tragen immer wieder neu dazu bei, dass Musik emotional ergreift.

Eine Studie von Nicola Dibben (2003) legt zudem den Schluss nahe, dass durch den körperlichen Erregungszustand, in dem Menschen Musik hören, Emotionen verstärkt werden. War der körperliche Erregungszustand ihrer Versuchsteilnehmer durch einen schnellen Spaziergang erhöht, erlebten sie Musik weit emotionaler als Mitglieder einer Kontrollgruppe, die vor dem Musikhören eine Entspannungsübung absolviert hatten. Vermutlich wird die emotionale Erfahrung also immer dann gesteigert, wenn sich Menschen zur Musik bewegen und durch diese Bewegung erregt sind.

Während intrinsische Emotionen fest mit bestimmten musikalischen Merkmalen verknüpft sind, besteht vermutlich kein eindeutiger kausaler Zusammenhang zwischen extrinsischen Emotionen und Musik. Emotionale Reaktionen beim Musikhören können sich auf Strukturähnlichkeiten von Musik und emotionaler Reaktion bzw. deren körperlichem Ausdruck grün-

den oder sich durch ein wiederholtes gemeinsames Auftreten von Musik und bestimmten emotionalen Empfindungen verfestigen. Sie bestehen mitunter jedoch auch aus rein assoziativen Verknüpfungen zwischen musikalischen Zusammenhängen und Emotionen, die auf einer Reihe von situativen und individuellen Faktoren beruhen und Ergebnis individueller oder kulturell bedingter Lernprozesse sind.

Vor diesem Hintergrund überrascht es nicht weiter, dass die Aussagen zur Verbindung von musikalischem Rhythmus und emotionalen Reaktionen, die Gabrielsson und Lindström (2001) aus verschiedenen empirischen Studien zusammengetragen haben, eher vage bleiben:

„Regular/smooth rhythm may be perceived as expressing happiness, dignity, majesty, and peace; irregular/rough rhythm, amusement, uneasiness, and anger; and varied rhythm, joy. Firm rhythm may be associated with expressions of sadness, dignity, and vigour, and flowing/fluent rhythm with expressions such as happy/gay, graceful, dreamy, and serene. Terminology varies among authors, which makes comparisons difficult“ (Gabrielsson/Lindström 2001, S. 241).

Die abschließende Bemerkung ist symptomatisch für die musikpsychologische Emotionsforschung, gerade in Bezug auf die terminologisch noch immer uneinheitlich beschriebene rhythmische Dimension.

Ein Zusammenhang zwischen emotionaler Wirkung und dem Tempo der Musik erscheint allerdings aufgrund zahlreicher Untersuchungen als gesichert (vgl. Gabrielsson/Lindström 2001, S. 235). Tempo wird vielfach als der wichtigste Faktor für die emotionale Wirkung und den emotionalen Ausdruck von Musik angesehen. Laut Gabrielsson und Lindström weisen die empirischen Untersuchungen darauf hin,

„that fast tempo may be associated with various expressions of activity/excitement, happiness/joy/pleasantness, potency, surprise, anger, and fear. Slow tempo may be associated with various expressions of calmness/serenity, dignity/solemnity, sadness, tenderness, boredom, and disgust. Which of the different expressions is perceived in each case is highly dependent on the actual context, that is, presence of other factors“ (Gabrielsson/Lindström 2001, S. 235, 239).

Da die im konkreten Fall wahrgenommenen Gefühle auch hier von individuellen und kontextuellen Faktoren sowie von anderen musikalischen Parametern abhängen, bleibt letztendlich nur die unmittelbar einleuchtende Aussage, dass mit dem musikalischen Tempo der Grad der Erregung und Aktivität wächst.