

# Über strukturelle Gemeinsamkeiten zwischen sprachlichen und musikalischen Melodien

## Einführung

Sprachliche und musikalische Tonverläufe werden gleichermaßen mit dem Begriff Melodie belegt. Das mag verwundern und ist doch wohlbegründet. Es handelt sich zwar dem Höreindruck nach um voneinander völlig verschiedene Phänomene, entgegen allen intuitiven Analogien: So scheint die Sprechmelodie mit musikalischen Kategorien schlecht greifbar, sie hat in ihrer Handhabung von Tonhöhe und Dauer meßtheoretisch ein einfacheres Niveau und wirkt dadurch verglichen mit Musik amorph. Aber gerade dies trägt, Phonetik und Phonologie lehren, daß die sprachliche Prosodie sehr differenzierten Gebrauch von melodischen und Akzentmustern macht. Sie ist in detaillierten Regeln fixierbar und läßt sich sogar anhand dessen über sogenannte Text-to-Speech-Programme resynthetisieren (CARLSON, GRANSTRÖM & HUNNICUTT 1990). Wenn »Melodie« heißt, Tonhöhen regelhaft und strukturiert im Zeitkontinuum anordnen, dann haben Musik und Sprache Melodie. Es bleibt aber nicht bei dieser eher äußerlichen formalen Übereinstimmung. Musikalische Melodie und sprachliche Intonation haben – darum soll es im Folgenden gehen – vermutlich auch substantielle Gemeinsamkeiten, für die ihr Melodiecharakter nur der Rahmen ist.

Diese Gemeinsamkeiten hängen an jenen Merkmalen der Sprechmelodie, die als Universalien an so gut wie allen und insbesondere sehr verschiedenartigen Sprachen nachgewiesen werden können. Universalie ist, daß die Abgeschlossenheit sprachlicher Äußerungen durch fallenden, die Unabgeschlossenheit durch nicht-fallenden, Interrogativität durch steigenden Tonverlauf beziehungsweise hohe Tonlage angezeigt werden (abgesehen von Ausnahmefällen wie den Fragewortfragen) (vgl. BOLINGER 1964, 1986, 1989 und CRUTTENDEN, 1981). BOLINGER (1978) hat zahlreiche Sprachen bezüglich der Intonation von terminalen Aussprüchen und von

Fragesätzen untersucht: Abweichungen von dem allgemeinen Prinzip sind selten und meistens erklärbar. Nach OHALA (1983, 1984), BOLINGER (1986) und anderen ist der hier behandelte Aspekt der Sprechmelodie im Ausdrucksverhalten verankert, OHALA postuliert aufgrund dessen einen genetisch fixierten »Frequenzcode« ( $f_0$ -Code, s. u.). Wegen dieser ursprünglichen Bindung an Ausdrucksphänomene kann der Gebrauch der wenn auch sprachlich konventionalisierten Intonationsmuster als »gestisch« betrachtet werden, so daß auch im Kontext ihrer linguistischen Verwendung der Terminus Intonationsgeste gerechtfertigt ist. Den vielfältigen Gestalten, die die Sprechtonkurve annehmen kann, liegt, wenn man dem folgt, ein einfaches und augenscheinlich fest verankertes gestisches Signalsystem als konstruktives Element zugrunde. Es darf erwartet werden, daß dies Faktum - eben im Sinne einer substantiellen Gemeinsamkeit - auch für musikalische Tongebilde Bedeutung hat (DOMBROWSKI 1987, 1992).

Hierfür sprechen weitere bekannte Sachverhalte: Die Entwicklung der europäischen Musik ist maßgeblich von der Auseinandersetzung mit Sprache geprägt worden (vgl. GEORGIADES 1954). Das schloß auch den sprachlichen Gestus und damit die Sprechmelodie mit ein (Nachvollziehen kann man dies am gregorianischen Choral, an der Figurenlehre des Barock und weiter bis zu Schönbergs musikalischer Prosa und in die neueste Zeit). Mit der Kadenz hat ein Strukturschema in die abendländische Musik Eingang gefunden, das ursprünglich und im Wortsinn mit der sprachlichen terminalen Intonation zusammenhängt.

Ferner setzen sich historische musiktheoretische und ästhetische Schriften verschiedentlich mit der Bedeutung des Tons der Sprache für Musik auseinander (vgl. zum Beispiel MATTHESON 1739, HERDER 1769, Quellen hierzu enthält GRUHN 1978).

In der phonetischen Literatur hingegen werden Beziehungen zwischen Intonation und Musik zwar an einigen Stellen hergestellt, aber nicht systematisch ausgearbeitet (vgl. etwa BOLINGER 1986, CRYSTAL 1987, WINKLER 1979).

Die musikpsychologische Forschung hat (abgesehen von Arbeiten wie denen von DEUTSCH 1991, 1994 und DOLSON 1994) der Sprechmelodie bisher wenig Beachtung geschenkt, namentlich nicht der Rolle der Satzintonatorischen Kategorien. Kommt hier die Prosodie zur Sprache, geht es wesentlich um Akzentstrukturen, Timing, Metrik (letzteres beispielsweise bei LERDAHL & JACKENDOFF 1983, siehe auch CARLSON, FRIBERG, FRYDÉN, GRANSTRÖM & SUNDBERG 1989).

Zweck dieses Beitrags ist nun, zu zeigen, daß die Betrachtung der Sprechmelodie und ihrer Kategorien wichtige Aufschlüsse für die psychologische Konzeptualisierung musikalisch melodischer Gebilde geben kann.

## Zwei Ebenen der Sprechmelodie

Die sprachliche Prosodie läßt sich in die Aspekte Akzentuierung und Intonation einteilen (siehe BOLINGER 1986, S. 24, KOHLER 1991b). Akzent als Satzakkent hebt aus der Folge der sprachlichen Einheiten einer Äußerung solche hervor, die im buchstäblichen Sinn betont werden sollen. Die Auswahl trifft im Deutschen und anderen Sprachen, die außerdem einen Wortakkent haben (etwa Englisch), fast ausschließlich die wortakzentuierten Vokale.

So könnten in dem Satz »Die Blätter fallen von den Bäumen« je nach Aussageabsicht die Silben »Blät« und/oder »fal« und/oder »Bäu« hervorgehoben werden.

Die mit Akzent ausgezeichneten Silben - und insbesondere dominierende Akzente - werden mit einer Melodiebewegung versehen, mit Intonation also. Sie erhält dabei die doppelte Aufgabe, unterstützt von Quantität (=Silbendauer) und Intensität (=Lautheit) (1) die Akzentuierung und durch Wahl einer melodischen Kategorie (2) sich selbst anzuzeigen: Im obigen Beispielsatz würde etwa eine deutlich fallende Intonation auf »Bäu« zugleich auf einen Akzent und auf Terminalität verweisen. Intonationsmodelle bieten für die Realisierung von Intonationsmustern je eine bestimmte Zahl von solchen möglichen melodischen Kategorien an, die Steigen und Fallen in verschiedenen Graden, sowie Kombinationen von steigendem und fallendem Verlauf einschließen. Diese Kategorien unterscheiden sich von Sprache zu Sprache, sind aber letztlich auf den genannten Grundgegensatz von Steigen und Fallen reduzierbar (vgl. wiederum KOHLER 1991b für das Deutsche, zum Beispiel HALLIDAY 1967 für das Englische).

Ferner variiert auf den tontragenden Vokalen der Einsatzzeitpunkt der Tonveränderung - mit entsprechenden Konsequenzen für die Interpretation der so entstehenden Muster (siehe hierzu KOHLER, 1991a,b, LADD 1983; im Deutschen kommt es dabei zu sogenannten frühen, mittleren und späten Gipfeln, frühen und späten Tälern, KOHLER 1991b, 1992). Durch die Bestimmung von auslösenden Akzenten und die Positionierung von

Gipfel- und Talkonturen werden Beginn und Gestalt sprachlicher Tonverläufe präzise festgelegt.

Festzulegen ist auch das Ende der Tonbewegung, sei es auf dem Niveau der einzelnen Akzentstellen oder auf dem nächsthöheren von Phrasen und Äußerungen. Ein Phrasenende tritt prosodisch durch finale Dehnung und/oder anschließende Signalpause hervor. Phrasengrenzen bewirken zudem eine stärkere Ausprägung steigender und fallender Tonverläufe (KOHLER 1992, S. 246f.) – wobei aber nicht eindeutig klärbar ist, ob das der Phrasengrenze selbst oder noch dem davor liegenden Akzent zuzuschreiben ist.

Die einzelnen Details und Einzelmerkmale der Sprechtonkurve fügen sich zu komplexen Melodiekonturen zusammen (wie etwa dem bekannten Hutmuster, vgl. KOHLER 1991b).

Je nach Auswahl, Lokalisation und Konfiguration der melodischen Elemente ergibt sich für die intonierte Äußerung ein anderer Sinn beziehungsweise eine andere Konnotation.

Pointierend zusammengefaßt heißt das: Durch Akzentuierung und Intonation werden ein Punkt oder mehrere Punkte im sprachlichen Kontinuum Auslöser von (zeitlich) abgegrenzten Tonbewegungen, die als phonologisch distinktiv verstehbar sind (vgl. KOHLER 1991a, S. 20).

Der Verlauf der Sprechtonkurve wird in Intonations- oder, anders ausgedrückt, Phrasierungseinheiten gegliedert und anhand solcher Einheiten beschrieben. Die Vorstellung einer Intonationseinheit legt nahe, nach einem sie insgesamt kennzeichnenden Moment zu suchen. Wohl auch deshalb wird vielfach (aber nicht unwidersprochen) davon ausgegangen, daß eine Intonationseinheit nur *einen* dominierenden Akzent aufweist. Dieser zieht die sie charakterisierende Melodiebewegung nach sich. Gebräuchliche Begriffe für derartige akzentgebundene melodietragende Strukturen sind *nucleus* (CRUTTENDEN 1981, KOHLER 1977), *tonic* (HALLIDAY 1967), *nuclear stress* (CHOMSKY & HALLE 1968), *primary stress* (TRAGER & SMITH 1951) und andere. Am *nucleus* setzt demnach der *nuclear tone* (CRUTTENDEN 1981) an, der von dort bis an das Ende der Einheit reicht, inklusive der dort befindlichen  *juncture* (TRAGER & SMITH 1951). Einige Autoren heben weniger den dominierenden (fokalen) Akzent sondern mehr den Endbereich der sprachlichen Äußerung als Bezugspunkt distinktiver Intonationsmerkmale hervor, etwa v. ESSEN (1964), der von einer »Endphasenbewegung« spricht, MEINHOLD & STOCK (1982) oder BIERWISCH (1971), der auf den mit dem letzten Akzent verbundenen Konturabschnitt verweist. BIERWISCH nimmt überdies (in Anschluß an KATZ &

POSTAL, 1964) ein tiefenstrukturell verankertes »Fragemorphem« an, das die Intonation von Fragesätzen determiniert. PIERREHUMBERT (1987) teilt Intonation auf: in Tonhöhenakzente, den Satzakkzent und den boundary tone.

Das Bemühen um die Gesamtkennzeichnung von Intonationseinheiten entspricht einer mehr ganzheitlichen Betrachtung sprechmelodischer Phänomene. Einen anderen Weg beschreitet das neu entwickelte Kieler Intonationsmodell (=KIM; KOHLER 1991a,b, 1992), indem es jeden einzelnen in einer Signalfolge identifizierten Akzent anhand mehrerer phonologisch distinktiver Merkmale (features) etikettiert - und so seinen Grad und das mit ihm assoziierte Melodiesegment festlegt. Hier wird eine Beschreibung von Tonhöhenverläufen ausgehend vom molekularen Niveau einzelner phonologischer Etikettierungen möglich, so daß das Ganze der Intonationseinheit auf dem Wege der Zusammensetzung erscheint. Dies Modell ist in Einzelaspekten verwandt mit den Ansätzen von LADD (1983), PIERREHUMBERT (1987), BRUCE (1977) etc. Die für die Beurteilung der Intonationseinheiten (im Sinne der hier erörterten Frage) schließlich maßgebliche Tonbewegung ist im KIM wiederum die zwischen dem letzten melodisch signierten Akzent und dem Phrasenende.

Das Kieler Intonationsmodell und verwandte Modelle haben für den hiesigen Gedankengang den Vorteil, daß sie sehr explizit sind. Sie erlauben, die im akustischen Signal für den Eindruck der Terminalität, Progredienz, Interrogativität und weiterer Qualitäten wichtigen Abschnitte genau zu lokalisieren, und können Tonbewegungen auf verschiedenen hierarchischen Niveaus beschreiben. Wie später an einem Beispiel gezeigt wird, ermöglicht dies präzise Vergleiche mit musikalischen Sachverhalten.

Das Bisherige erlaubt folgende Ableitungen und Interpretationen:

(1) Wenn Intonationsmodelle auch über eine Vielzahl verschiedenartiger möglicher melodischer Muster Rechenschaft ablegen müssen, gehen sie doch auf die erwähnte elementare Dichotomie von »steigend-fallend« oder »hoch-tief« zurück – mit ihrer charakteristischen Semantik, wie CRUTTENDEN (1981) und später OHALA (1983) sie beschreibt. Auch KOHLER (1991, S. 312) hat dies nochmals herausgestellt.

(2) Intonationsmodelle grenzen im sprachlichen Kontinuum Bereiche ab, in denen Melodieverläufe im Sinne der genannten fallenden terminalen und der nicht-fallenden oder steigenden progredienten bzw. interrogativen Muster interpretiert werden, in denen – um diesen Begriff wieder ins Spiel zu bringen – linguistisch relevante Gesten erscheinen. Dafür prädestiniert ist

die Strecke, zurückgreifend vom Phrasenende bis zum letzten lexikalischen, letzten realisierten oder zum nuklearen Akzent; Akzente zeichnen sich durch hohe Prominenz (das heißt Auffälligkeit) aus, Phrasengrenzen durch geringe Ereignisdichte. Den so markierten Bereichen stehen solche gegenüber, die nicht oder nur in geringerem Maße auf diese Art gedeutet werden oder die überhaupt keine größeren Melodiebewegungen aufweisen.

(3) Prosodie wird deshalb auf zwei Ebenen beschrieben: auf der intonatorischer Muster und auf einer Ebene, die ich als erlebte sprachliche Zeit bezeichnen möchte. Abstrakte Grundlage für sie ist das Aufeinander-Folgen der Silben im Sprachsignal, genauer: der Vokale. Sie bildet eine »elastische« Zeitskala auf »aktionalem«, vom physikalischen Standpunkt aus niedrigem metrischen Niveau. Auf dieser Basis wird ein Pulsieren von Intensität und Ereignisdichte erzeugt (durch Akzente, Pausen, Tempovariation), welches die Lokalisation von phonologisch beziehungsweise – wie oben angedeutet – gestisch relevanten Melodieabschnitten ermöglicht. So werden abgeschlossene und weiterweisende Tonverläufe erkennbar. – Diese Überlegungen stützen sich, das sei einschränkend hinzugefügt, auf die Evidenz vorhandener Intonationsmodelle und auf die These der Universalität der Verwendung der Grundfrequenz, sie bedürfen ungeachtet ihrer Eignung für sprachlich-musikalische Analogien experimenteller Fundamentierung.

In dem folgenden musikalisch transkribierten gesprochenen Satz »Die Blätter fallen von den Bäumen« wurden die (je nach Auflösungsgrad der Betrachtung) für die Interpretation der Melodie im Sinne von Terminalität und Progredienz besonders relevanten Abschnitte eingetragen (Abb. 1).

Die Kontur kann (1) als ganze betrachtet werden: Dann ist sie ein terminales Muster. Die zugrundeliegende Äußerung kann aber auch (2) in Teilphrasen (Nominal- und Verbalphrase) zerlegt und die Tonkurve entsprechend eingeteilt werden: So steht ein auftaktig steigendes einem abtaktig fallenden Motiv gegenüber. (3) Von den drei realisierten Akzentpositionen der Äußerung – als kleinsten Einheiten – bildet schließlich die erste ein »Tal«, die zweite ist partiell deakzentuiert und an die dritte angeschlossen, sie ist ebenfalls ein »Tal«. Die kennzeichnende dritte Position ist ein »früher Gipfel«. Je nach Feinheit der Analyse entstehen in dem Beispiel also ein, zwei oder drei melodisch relevante Bereiche. Je größer die gewählte Einheit, desto mehr »irrelevante« Kontur muß in den folgenden »relevanten«, distinktiven Bereich integriert und ihm untergeordnet werden. Abweichend

von üblichen phonetischen Analysen wird hier die Interpretation mit Hilfe steigender vs. fallender »Tongesten« auf allen denkbaren hierarchischen Niveaus versucht.

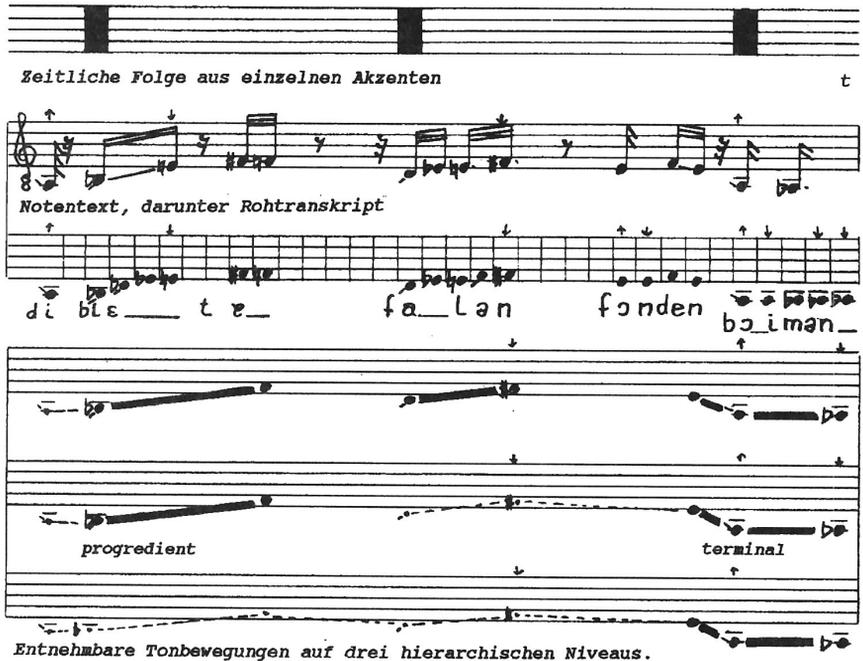


Abb. 1: Tonverlauf des Satzes »Die Blätter fallen von den Bäumen« in musikalischer Transkription, gewonnen anhand einer Bandaufnahme durch Abgreifen von Zeitintervallen von ca. 100 ms Dauer. Die Darstellung zeigt die zeitliche Ebene der Akzentfolge (– Blöcke) und die Ebene melodischer Gesten (– Balken). Die Melodiebewegung wurde auf drei Niveaus im Sinne des Gegensatzes »steigend-fallend« interpretiert. Gestrichelte Linien zeigen zur Charakterisierung der Tonmuster zusätzlich heranziehende Töne an.

Nach diesem Beispiel ist der weitere Verlauf der Darstellung absehbar: Er enthält den Versuch, die eben skizzierte Struktur, die der Sprechmelodie zugrunde gelegt wurde, auf den musikalischen Gegenstand zu übertragen. Zunächst soll in der nun folgenden argumentativen Schleife die Lücke näher beschrieben werden, die der hier vertretene Ansatz ausfüllen will. Dies geschieht in Form von Kommentaren zu einigen Konzepten aus Emotionspsychologie, Musikpsychologie und Phonetik, Konzepten, die für seine Herleitung und Einordnung wichtig sind.

## Vom Gefühlsausdruck zur sprechgestischen Struktur

### I. Einige musikpsychologische Ansätze

1. Die prosodischen Merkmale sprachlicher Signale (wie Tonhöhenverlauf, Dauer, Lautheit, spektrale Charakteristika, Tempo etc.) werden auch als paralinguistische oder nicht-linguistische Phänomene analysiert – etwa im Rahmen von Ausdruck/Eindruck-Studien. Dazu sind sie – was nicht unproblematisch ist – zunächst aus dem je konkreten sprachlich-semantischen Kontext herauszulösen, oder sie müssen technisch synthetisiert werden. So hat die Gruppe um SCHERER entlang dem Konzept der »Vokalen Kommunikation« (SCHERER & BERGMANN 1984) und des »component process-Modells« der Emotion (SCHERER 1984, s. SCHERER 1991) unter anderem untersucht, welchen Ausdruckswert die Ausprägungen isolierter akustischer Parameter haben (hohe vs. niedrige, steigende vs. fallende, ferner Variabilität der Grundfrequenz, Schallintensität, Frequenzspektrum und weiteres mehr). Ziel dieser Arbeiten ist, analog zu Forschungen zum sichtbaren nonverbalen Verhalten (Mimik, Gestik), grundlegende vokale expressive Signale zu identifizieren (für Angst, Freude, Ärger, Selbstsicherheit, aber auch für überdauernde Merkmale wie Depression, vgl. die Übersicht in SCHERER 1991). Dabei machen SCHERER & OSHINSKY (1977) in ihren Experimenten auch den Versuch, von ihnen gefundene auditive Parameter expressiver Signale auf musikalische Tonfolgen zu übertragen. Sie halten es demnach für gangbar, musikalische Gebilde mit Hilfe vokal-kommunikativer gestischer Muster zu beschreiben. SCHERER & OSHINSKY entwickeln allerdings keine Vorstellungen darüber, wie diese Prozesse des Ausdrucks in musikalische Strukturen eingehen: Musik wird als Ausdrucksphänomen verstanden, aber in struktureller Hinsicht auf eher diffuse Weise. Die Unterteilung prosodischer Phänomene in linguistische und nicht-linguistische – wenn sie sinnvoll möglich ist – müßte auf Seiten der Musik die Unterscheidung von musikalisch und paramusikalisch nach sich ziehen. Damit wäre das Problem, wie vokaler Ausdruck in das System Musik transformiert werden kann, überhaupt erst formuliert und diskutierbar. Die Reflexion der Rolle vokaler expressiver Signale in Musik sollte diesen Schritt der Differenzierung eigentlich einschließen.

2. In unmittelbarer Weise ist das Moment der Strukturiertheit von Ausdrucksprozessen in dem expressiven Emotionsmodell von CLYNES (1977) repräsentiert. Dies Modell ist wie kaum ein anderes als gestisch zu bezeich-

nen (vgl. auch BEHNE 1982); es ist in der Auseinandersetzung mit Ausdrucksbewegungen zu Musik entstanden, Merkmalen der Schlagbewegung beim Dirigieren, wie sie BECKING (1928) untersucht hatte (sog. Becking-Kurven, *composer pulses*). Das von CLYNES schließlich entwickelte operationale Kriterium bleibt der physischen Geste nahe: CLYNES überführte den motorischen Impuls in einen Tastendruck und erfaßte nun Gefühlsäußerungen in Form von Tastendruckverläufen (*essentic forms*) bei der Betätigung des hierzu konstruierten Gerätes »Sentograph«. Einbezogen wurden sieben als funktionell unabhängig angesehene Gefühle: Ärger, Haß, Trauer, Liebe, Sex, Freude, Ehrfurcht.

Wenn man Gefühle anhand von charakteristischen zeitlich abgegrenzten Ausdrucksmustern betrachtet, liegt es nahe, Sequenzen der entsprechenden Ausdruckshandlungen zu erheben – dies auch abgesehen von dem Ursprung des CLYNESschen Verfahrens in iterativen Bewegungen zu Musik. Indem CLYNES seine Versuchspersonen auffordert, ihre Ausdrucksreaktionen bei der Betätigung des Sentographen zu reihen, entstehen solche Sequenzen und dabei unvermeidlich eine syntaktische Verbindung zwischen den einzelnen Gesten. Jede dieser Sentographen-Gesten hat die in der Sequenz typische Verlaufsform; gleichzeitig macht das Auf und Ab der Intensität des aufgewendeten Fingerdrucks und machen die entstehenden Pausen die Reaktionen einzeln erfaßbar. Somit ergibt sich mit den Ausdrucksreaktionen zugleich zeitliche Struktur. Diese kann man – wie oben in dem Zwei-Ebenen-Modell zur Intonation - als ein durch Intensitätsverlauf und Pausen gegliedertes Pulsieren beschreiben.

Es entspricht der Logik und dem Ursprung ihres Ansatzes, daß CLYNES & NETTHEIM (1982) die für einzelne Gefühlsqualitäten spezifischen *essentic forms* des Sentographen in Tonhöhengebilde und (ganz analog zu SCHERER & OSHINSKY, 1977) in Melodien transformieren, um sie so erneut zu testen. Und es ist genauso folgerichtig, daß CLYNES (1986) seine im Umgang mit Musik gewonnenen Ideen auf Musik rückbezieht und nun Musik in ihrer Struktur aus der Mikrostruktur der *pulses*, dem *amplitude shaping* und den melodisierten *essentic forms* herleiten will. –

Kern von CLYNES Denkweise bleibt dabei – mag der Anspruch auch weiter greifen – die Reihung von gleichartigen gestischen Elementen, also eher ein rhythmisches Prinzip. Die form- oder strukturbildende Dynamik, die im Verlauf einer Tongeste selbst liegen kann, ist nicht Gegenstand seines Modells.

3. Das Verhältnis zwischen expressiver Geste und musikalischer Struktur

ist auch Gegenstand der Forschungen zur Interpretation und Reproduktion von Musik (Performance-Forschung). Hier steht die expressive Geste, der Ausdruck, dem Notentext als Kommentar gegenüber und soll sich aus der Wahrnehmung der musikalischen Struktur ergeben (vgl. zum Beispiel CLARKE 1987). So erscheinen beispielsweise wahrgenommene Phrasengrenzen in der Interpretation als expressive Abweichungen von der metronomgemäßen Spielweise. Ausdruck wird hier emotional nicht näher qualifiziert, er ist mehr ein Vehikel. Ausgedrückt wird die Struktur eines Stückes, das heißt Aspekte der Form (zum Beispiel die Abstufung wahrgenommener Grenzen zwischen Abschnitten). Das scheint auf den ersten Blick ganz etwas anderes zu sein als das, was weiter oben Ausdruck hieß. Wenn nun Gegenstand der expressiven Gesten die Struktur ist und nicht etwa eine bestimmte Emotion, ist es nur ein kleiner Schritt, die Struktur selbst als gestisch aufzufassen - und damit auch Tonhöhenphänomene der interpretierten Musik unter diesem Gesichtspunkt einzuordnen: Die im Notentext enthaltene musikalische Substanz wäre dann selbst und unmittelbar auf ihren gestischen Charakter zu befragen und nicht nur die interpretatorische Abweichung vom Text, die schließlich Geste des »Wiedererkennens« ist.

Die Phänomene, die die Performance-Forschung mißt, können in sprachlicher Analogie als prosodisch angesehen werden (vgl. CARLSON, FRIBERG, FRYDÉN, GRANSTRÖM & SUNDBERG 1989): Temposchwankungen, Signalpausen, Dauernveränderungen an Einschnitten, Phrasengrenzen und akzentuierten Stellen werden bereits an Sprache untersucht. Es handelt sich also, wenn man so will, um prosodische Zutaten zur Musik. In sprachlichen Signalen ist die zeitliche Strukturierung durch solche prosodischen Faktoren funktionell mit sprechmelodischen Mustern verknüpft. Etwas Entsprechendes darf für Musik erwartet werden. Und – um die Analogie komplett zu machen – wenn Interpretation durch ihre Zutaten auf in der Musik selbst liegende gestische Strukturen verweist, könnten diese aus sprechmelodischen Mustern abgeleitet sein.

Daß der melodische Aspekt der Prosodie von der Interpretationsanalyse bisher ausgeschlossen blieb, hat auch sachlogische Gründe: Der Notentext liegt bei der Interpretation von Musik fest. Der melodische Aspekt wäre in Studien zur musikalischen Produktion leichter faßbar (wie sie in Zusammenhang mit dem hier entwickelten Konzept durchgeführt wurden, DOMBROWSKI 1987, 1992, in Vorbereitung).

4. Ausdrücklichen Bezug auf Faktoren der Prosodie nehmen LERDAHL

& JACKENDOFF (1983) in ihrer *Generative Theory of Tonal Music*, sie bedienen sich als Vorlage für ihre Baumdarstellungen der *metric trees* von LIBERMAN & PRINCE (1977); es geht hier nicht um expressive Valenzen der Prosodie, sondern allein um ein adäquates Modell für die Repräsentation musikalischer Zusammenhänge und Gliederungen. Während LERDAHL & JACKENDOFF für die Modellierung der zeitlichen Segmentierung der Musik Elemente aus der Prosodie heranziehen, fehlt ein entsprechender sprachlicher Bezug (nämlich zur Intonation) in ihrer Konzeption der an Tonhöhe gebundenen musikalischen Phänomene: Hier ist für sie die Kadenz zentral, als harmonisches Schema, dessen Charakter vorerst aber nicht weiter analysiert wird - So bleibt etwa auch die oben angedeutete Möglichkeit ungenutzt, die Kadenz auf elementare sprachliche Intonationsmuster rückzubeziehen. Eine solche Deutung der Kadenz wäre immerhin eine psychologisch plausible Grundlage für die Erklärung der von LERDAHL & JACKENDOFF so genannten strukturellen Anfänge und Endungen (»b's« und »c's«). Bezüge zum gestischen Ausdruck finden sich jedoch in LERDAHL & JACKENDOFFs Vorstellung, daß den von ihnen strukturell beschriebenen musikalischen Einheiten Abläufe aus »Spannung und Lösung« entsprechen. Sie kommen in der »prolongational structure« zum Tragen. (Dieser Gedankengang wird bei JACKENDOFF (1991) weiterentwickelt). LERDAHL (1991) führt auch das Überwiegen der fallenden Melodiebewegung in musikalischen Schlüssen auf einen Entspannungsvorgang zurück – und nicht auf eine zeichenhafte Tongeste oder eine phonologische Kategorie (obwohl er die Analogie zur Sprache hier erwähnt).

Die Zusammenschau der erwähnten theoretischen Ansätze ergibt folgende Bilanz:

(1) Es gibt Versuche, vokale Signale der Form »hoch – tief«, »steigend – fallend« unter musikpsychologischem Gesichtspunkt zu untersuchen (SCHERER & OSHINSKY 1977). Dabei geht es nur um den Ausdruckswert dieser Signale, nicht ihre gleichzeitige sprachliche Funktion. Gerade die sprachliche Verwendung der elementaren sprechmelodischen Muster wäre aber für das Verständnis von Musik als Strukturgebilde von Interesse. Die Betrachtungsweise von SCHERER & OSHINSKY erlaubt jedoch, Tonhöhenverläufe im Sinne vokaler Kommunikation als Gesten aufzufassen.

(2) Wenn »Ausdruck« sich gestisch artikuliert, das heißt in Form von in der Zeit gestalteten und zeitlich begrenzten Mustern, kann er durch Reihung dieser Muster Strukturen oder Formen bilden (wie bei CLYNES),

kann ferner Strukturen »begleiten« (etwa bei CLARKE und auch bei CLYNES). Diese Eigenschaft ist unabhängig von der Modalität des Ausdrucks. Gestische Einheiten erscheinen so als eine mögliche Bedingung der Wahrnehmung und/oder Erzeugung musikalischer Strukturen. Musik selbst kann als geronnene gestische Struktur angesehen werden, konventionalisiert und schriftlich fixierbar. Die Muster der Satzintonation werden in dem zitierten Kontext einer gestischen Auffassung musikalischer Vorgänge bisher nicht erörtert - obwohl sie zu den uns geläufigsten Tongesten zählen und obwohl gerade sie gereimte Strukturen erzeugen.

(3) Werden – implizit oder explizit – Parallelen und Wechselbezüge zwischen sprachlicher Prosodie und Musik hergestellt, so beziehen sie sich auf ihre Eigenschaft, Zeitabläufe zu ordnen – die erste der oben genannten zwei Ebenen. Die Ebene der sprechmelodischen Muster, mit ihrer Dichotomie von Steigen und Fallen, bleibt aus den musikpsychologischen Überlegungen wiederum ausgeklammert (vgl. LERDAHL & JACKENDOFF 1983).

(4) An den behandelten musikpsychologischen Denkansätzen sollte zweierlei herausgestellt werden, eine spezifische Möglichkeit und ein spezifischer Mangel: Sie alle enthalten zum einen an untergeordneter oder zentraler Stelle Konzepte, die Einheiten des Ausdrucks (die als gestisch zu verstehen sind) und Einheiten der musikalischen Struktur miteinander verbinden. Sie eignen sich also in je unterschiedlichem Maße dazu, die Beschreibung der Musik als Ausdruck und als Form miteinander in Einklang zu bringen – wobei sie in diesem Versuch doch stets fragmentarisch bleiben.

Sie alle haben zum anderen den Mangel, daß sie sich nicht für die Sprechmelodie als musikalischen Gegenstand interessieren. Und es wurde bereits deutlich, daß von der Thematisierung der Sprechmelodie Anregungen für die musikpsychologische Theoriebildung zu erwarten sind.

## II. Intonation als Universalie

Das stärkste Argument für die Musikpsychologie, sich mit der Sprechmelodie zu beschäftigen, ist zweifelsohne der Aspekt Universalität. Intonation als Universalie legt dringend nahe, nicht nur in sprachlichen sondern auch in musikalischen Systemen nach der Semantik steigender und fallender Tonmuster zu suchen. Der eigentliche Vorteil der Integration der Sprechmelodie in das musikpsychologische Denken liegt aber in einem anderen Merkmal, das allerdings mit der Universalität verbunden ist: ihrer Ambivalenz zwischen Darstellungsfunktion einerseits und Appell- und Aus-

drucksfunktion andererseits (vgl. BÜHLER 1934). Gerade mit dieser Eigenschaft kann sie dazu beitragen, die am Emotionsausdruck und die an ihrer kognitiven Verarbeitung orientierte Sicht der Musik zusammenzuführen. Der Doppelcharakter der Intonationsphänomene läßt sich nicht darauf reduzieren, daß sie eine linguistische und eine paralinguistische Seite haben. Ambivalenz liegt bereits innerhalb der linguistischen Funktion. Die Intonation ist einerseits anhand von phonologisch definierbaren Kategorien beschreibbar, und ihre Muster haben die Eigenschaft sprachlicher Zeichen. Andererseits ist die Beziehung zwischen Zeichen und Bedeutung nicht frei, nicht arbiträr, sondern prädeterminiert – weshalb die Bedeutungsdimensionen wie »progradient/interrogativ - terminal«, »offen – geschlossen« und dergleichen (vgl. CRUTTENDEN 1981) fast ausnahmslos auf ein und dieselbe Weise mit den Kategorien »steigend – fallend«, »hoch – tief« verbunden werden. Diese festliegende Bedeutung der sprachlichen Intonationsmuster weist eher auf Ausdrucksphänomene hin, auf Gesten (Zum gestischen Charakter von Intonationsmustern äußert sich zum Beispiel BOLINGER, 1986).

OHALA (1983, 1984) hat versucht, mit seinem Konzept eines Frequenzcodes (frequency code,  $f_0$ -code) theoretische Rechenschaft über die Natur der intonatorischen Universalien abzulegen, einer Theorie über die universale Bedeutung des Parameters der Tonhöhe (und ihres Verlaufs), die gezielt eine Verbindung zwischen der Rolle der Tonhöhe in der vokalen Kommunikation und ihrer Aufgabe in echt sprachlichen Signalen herstellt. Damit wird im Grunde die Trennung zwischen paralinguistischen und linguistischen Phänomenen in dieser Hinsicht eingeebnet. OHALA übernimmt eine ethologische Position und argumentiert, daß es einen kulturell universalen und auch in der Tierwelt nachweisbaren Gebrauch tiefer und hoher, fallender und steigender Frequenzen gebe. Insbesondere fänden sich hohe und steigende Frequenzen bei der Signalisierung von Abhängigkeit, Unterlegenheit, Hilfsbedürftigkeit, tiefe und fallende seien mit Drohgebärden (»agonistic display«) und der Signalisierung von Dominanz verbunden. Die Sprache mache sich dieses Signalsystem zunutze – es schlägt sich in den Tonsprachen nieder (bei der Symbolisierung der Dimension »klein« vs. »groß«) und vor allem in den Intonationsmustern: Wer eine Frage stellt, signalisiert (abgesehen von charakteristischen Ausnahmefällen) intonatorisch, daß er vom Wohlwollen des Befragten abhängig ist. Eine Aussage wird mit einem Anzeichen von Bestimmtheit und Selbstsicherheit abgeschlossen.

Zwar existieren Befunde, die den Tonhöhenparameter auch mit anderen emotionalen Qualitäten in Verbindung bringen, ferner solche, die besagen, daß etwa für den Ausdruck von »Ärger« alternative Möglichkeiten bestehen (SCHERER 1991, 1989, OHALA 1983), die nur zum Teil oder vermittelt mit dem Frequenzcode koinzidieren. Für die hiesigen Überlegungen ist jedoch wichtig, daß es unter dem Gesichtspunkt der vokalen Kommunikation und analog auch in sprachlichen Signalen einen Dualismus »hoch – tief«, »steigend – fallend« gibt, in dem sowohl die eine wie die andere Ausprägung eigene Bedeutung haben. Der fallende, im Sinne sprachlicher Intonation »terminale« Tonverlauf ist demzufolge nicht allein Ergebnis etwa sinkenden Atemdrucks in der Expirationsphase (LIEBERMAN 1974), einer Entspannungsreaktion. Er ist nicht nur der negative Pol von Betonung, Hervorhebung, Akzent, sondern selbständiges intonatorisches Signal. Der Dualismus von »hoch« und »tief«, »steigend« und »fallend« auf zwei eigenständige vokale Gesten hin.

Die beschriebenen vokalen Gesten haben demnach sprachliche Funktion erlangt: Als technisches und zugleich im Ursprung expressives Instrument der Sprache dienen sie der Kennzeichnung der Satzmodi, der Grenzen von Intonationseinheiten, der Teilsatzgrenzen, von sprachlichen Über- und Unterordnungsverhältnissen, von Gegensätzen, bis hin zu Aspekten der Situativität (vgl. hierzu die Ausführungen von HELFRICH (1985) über Funktionen der Intonation; Beispiele bei KOHLER (1977) über Intonation unter bestimmten Redekonstellationen). Der Herkunft nach expressive Tonmuster wie das Heben und Senken der Stimme sind somit nun der Darstellungsfunktion der Sprache zuzuordnen, sie tragen zudem zum Prozeß der Sprachwahrnehmung bei (HELFRICH 1985).

Daneben bleiben ihre nicht-linguistische Ausdrucksfunktion und die Appellfunktion mit all ihren Facetten erhalten. Das macht ihre Ambivalenz aus.

Das eben Gesagte rechtfertigt, im Melodiekontinuum der Sprache eine Verkettung intonatorischer Gesten zu sehen, eine Verkettung gewissermaßen von ins Kognitive gewendeten Affektäußerungen, die den Sprechstrom tragen. Sprachliche Intonationsmuster bewirken insofern im Sprechen eine fortlaufende gestische Markierung.

Das Prinzip der Verkettung von Gesten kennen wir bereits aus CLYNES Gefühlstheorie und von seinem Musikbild (s. o.), das von dem Ideal einer geschichteten homogenen expressiven Struktur ausgeht, welche letztlich auf Reihung beruht. Die Dialektik der komplementären oder antagonistischen

schen sprechmelodischen Gesten trägt jedoch eine ungleich stärkere formbildende Dynamik in sich, als sich diese aus den Vorstellungen von CLYNES ergibt. Form und Ausdruck scheinen in der Sprechmelodik geradezu verschmolzen: Form und das, was sie »tönend bewegt«<sup>1</sup> (HANSLICK 1989/1854). Gerade diese Eigenschaft der Integration von expressiven und formbildenden Momenten macht das System der intonatorischen Gesten beziehungsweise der Intonation insgesamt als Vorlage für musikpsychologische Hypothesen interessant. Hier öffnet sich ein Zugang zu elementaren Techniken der Gestaltung und Verarbeitung von Melodien, ein Zugang zu ihren Verknüpfungsprinzipien (zur »musikalischen Syntax«).

## **Zwei Ebenen der musikalischen Melodie: Ein Intonationskonzept für Musik**

Wenn geklärt werden soll, ob intonatorische Grundmuster zur wahrnehmungsmäßigen Verarbeitung und zum Prozeß der Produktion musikalischen Materials beitragen, ist der nächste Schritt, ein auf Musik zugeschnittenes Intonationskonzept<sup>2</sup> zu entwickeln.

An den Anfang eines solchen Konzepts wären folgende Aussagen zu stellen: Sprachliche und musikalische Melodien können auf ähnliche Weise modelliert werden. Musik artikuliert sich wie Sprache in gestischen Grundeinheiten von gewisser Geschlossenheit. Beide bilden Ordnungsgefüge auf Grund der Reihung gestischer Elemente. In Musik liegen dem die Kategorien und Bausteine der musikalischen Syntax zugrunde, in Sprache Intonationseinheiten und ihre sprachlichen Träger.

Als Prosodie oben anhand eines Zwei-Ebenen-Modells beschrieben wurde, zielte dies bereits auf Anwendbarkeit in Musik und Sprache. Die dort benutzten sprachbezogenen Begriffe sind durch musikbezogene ersetzbar. Erfäßbar werden sollte mit dieser Darstellungsweise jenes Mehr an Gemeinsamkeit zwischen sprachlicher und musikalischer Melodie, welches darüber hinausgeht, daß beide Tonhöhen zeitlich strukturieren. Dieser nicht mehr nur äußerlich formale, sondern substantielle Bezug zwischen ihnen äußert sich in der psychologischen Kennzeichnung als gestische Struktur oder Struktur aus Gesten.

Ausgegangen wurde von der Auffassung, daß prosodische Phänomene zwei Aspekte haben: Akzent und Intonation. Der Akzentuierung, den Sprechpausen etc. wurde zugeschrieben, daß sie in einem Kontinuum erlebter sprachlicher Zeit Bereiche festlegen (Ebene 1), in denen Melodie-

verläufe im Sinne intonatorischer Gesten (Ebene 2) gedeutet werden. So gibt es im Verlauf der Sprechtonkurve einen Bezirk, der für die Abgrenzung des Satzes, die Beurteilung des Modus etc. besonders bedeutsam ist und einen solchen, der es weniger ist. Diese Vorstellung verband sich mit dem Konzept »Nukleus«.

Musikalische Verläufe können nun auf ähnliche Weise interpretiert werden: Metrische Schemata, rhythmische Gruppen, Akzente, Pausen, aber auch der Sprachrhythmus eines etwaigen zugehörigen Textes präjudizieren Stellen besonderer struktureller Betonung, legen entlang der Zeitachse Abschnitte fest, an denen musikalisch-syntaktisch relevante Wendungen zu setzen sind oder an denen melodische und harmonische Merkmale mit erhöhter Wahrscheinlichkeit syntaktisch interpretiert werden. Damit entsteht zum sprachlichen Nukleus ein musikalisches Äquivalent.

Besagte Wendungen und musikalische Strukturbestandteile vertreten mit musikalischen Mitteln intonatorische Gesten, das heißt sprachliche Intonationsmuster. Sie tun dies unmittelbar durch die melodische Linienführung an den ausgezeichneten Stellen, mittelbar harmonisch, etwa durch Kadenzen, also konventionell bestimmte Akkordfolgen, die mit melodischen Linienführungen (der Quintfallklausel im Baß zum Beispiel) funktional und bereits als Begriff verknüpft sind.

Es sind demnach in Musik kontinuierlich wiederkehrende Einheiten festgelegt, denen der Gegensatz »steigend – fallend« im Wahrnehmungsprozeß entnommen werden kann. Umgekehrt wird er in der musikalischen Produktion dort verschlüsselt. Im musikalischen Sinne handelt es sich dabei oft um Kadenzen und Klauseln.

Der gestische Wert fallender und steigender Melodieverläufe läßt sich an musikalischem Tonmaterial leicht nachvollziehen. So ist es vertraute musikalische Erfahrung, daß es eine Affinität zwischen fallender Melodieführung oder fallendem Intervall und der Schlußbildung gibt, wenn auch viele signifikante Abweichungen zu dem allgemeinen Prinzip auftreten. Mit der Sprechmelodie verhält es sich aber in der Hinsicht nicht anders.

Eine geradezu paradigmatische Darstellung des allgemeinen Grundsatzes liefert SCHENKER (1935) mit seinem Konzept der »Urlinie« und des »Ursatzes«. SCHENKER läßt die Musik, die er analysiert, auf eine simple fallende Linie im Diskant und eine kadenzierende steigend-fallende I-V-I-Fortschreitung in der Unterstimme zusammenschmelzen. Die Parallele zur terminalen Intonation ist evident. Das terminale Tonmuster wird geradezu musikalische Grundgestalt.

Die Anmutungsqualität steigender Tonverläufe andererseits spiegelt sich beispielsweise in der bekannten Metapher der »musikalischen Frage« (Es muß nicht erst auf die barocke musikalisch rhetorische Figur der interrogatio zurückgegriffen werden).

Verdichtung von Akzenten

Zeitlich-strukturelle Ebene      metrische und Dauernakzente

Notentext

progre dient      terminal      getilgt      terminal

v. Stufe

Terz: es'

terminal      terminal

Gestische Ebene

Abb. 2: Thema der Fuge c-Moll, BWV 847, von J. S. Bach, analysiert nach seiner intonatorisch-gestischen Struktur. Die zeitlich-strukturelle Ebene zeigt die angenommenen Akzentpositionen, die die Melodiebewegungen gestisch interpretierbar machen (– Blöcke); notiert ist ursprünglich 4/4-Takt. Die gestische Ebene zeigt die herausgelesenen »Intonationsmuster« (– Balken) bei halbtaktigen, ganztaktigen und doppel-taktigen »Akzentpulsen«. Gestrichelte Linien weisen auf Tonverbindungen hin, die für die Interpretation der gestischen Muster zusätzlich herangezogen werden können. In einem Fall (bei 1) relativieren sie die Deutung. .

Ein Beispiel soll nun betrachtet werden, ohne den Anspruch einer voll-ständigen Erörterung, das Thema der c-Moll-Fuge aus dem ersten Teil des WK, BWV 847. An seinem in Abbildung 2 wiedergegebenen Notentext wurden die vorzugsweise intonatorisch-gestisch interpretierbaren Stellen kenntlich gemacht – und die dort lokalisierten »Tongesten«. Die »Anziehungspunkte« für gestische Elemente, für gestische Markierungen, sind metrischer Art, nämlich die schweren Zählzeiten im 4/4-Takt. Auf eine weiterweisende Geste in Takt 1 folgt eine terminale an der Grenze zu Takt 2, die nächste ist getilgt, so daß die übernächste (in Takt 3) ein um so höhere Gewicht erhalten kann.

Dazu ist (1) zu anzumerken, daß die für ein musikalisches »Intonationsmuster« kritische Tonbewegung bereits zur Akzentstelle hin auftreten kann, da der musikalische Ton als Einzelton keine Gleitbewegungen kennt und diese daher dort streicht, wo sie im Sinne gesprochener Sprache vorkommen würden, nämlich auf dem sie auslösenden Akzent. Hiermit ist ein Gegenstand angesprochen, der in der Terminologie des Kieler Intonationsmodells (KÖHLER 1991b, 1992) unter dem Stichwort »Gipfel- und Talpositionen« behandelt wird und in der englischsprachigen Literatur unter »peak alignment« (LADD 1983): Für die exakte Platzierung von gestischen Melodiebewegungen an den dafür empfindlichen Stellen sind in Musik andere Regeln anzunehmen als in Sprache. Erste Hinweise für solche Regeln bietet die phonetische Beschreibung der »stilisierten Intonation« (stylized intonation): Hier geht LADD (1983) von einer phonologischen Kategorie (feature) der ausgehaltenen Tonhöhe aus (sustained pitch). Das Aushalten des Tones führt musikalisch zu im Vergleich zur Sprache bemerkenswerten Konsequenzen: Beispielsweise kann die besondere Akzentuierung eines Schlußtons (die sprachlich der Emphase auf der Schlußsilbe gleichkäme) entgegen dem Stereotyp der terminalen Intonation einen steigenden Tonsprung zur Folge haben, der Schlußfall wird gekappt. Dies wäre mit der prinzipiellen Gültigkeit der terminalen Intonationsgeste in Musik dennoch vereinbar: Die terminale Formel (etwa eine V-I-Fortschreitung im Baß) würde hier nur von einer »sustained-pitch-Konvention« überlagert (so daß es zum Beispiel zu einer Quartsprungklausel kommen kann, statt zu einem Quintfall). Sprachlich hätte eine stark akzentuierte Schlußsilbe nach einem ausgeprägten Hochpunkt einen steilen Fall. In dem Spezialfall der Textvertonung ergibt sich aus solchen Musik-Sprache-Kontrasten das Problem einer angemessenen Transformation zwischen sprachlicher und musikalischer Melodieebene. Generell kann in solchen Fällen die Harmonik kompensatorisch eintreten.

Das ausgewählte Beispiel macht aber (2) vor allem deutlich, daß Melodiegebilde, selbst schon kurze, in ein Pulsieren »freier« und syntaktisch-gestisch gebundener Tonfolgen aufgelöst werden können. Zugleich wird ersichtlich, daß man die behauptete pulsierende intonatorisch-gestische Struktur auf unterschiedlichen hierarchischen Niveaus ansetzen kann: Der Sequenz elementarer Gesten ist eine zweite Schicht unterlegt in Form einer gestuften fallenden Bewegung, die in eine ausführliche, als terminal zu verstehende Schlußformel hineinführt. Die einzelnen auf diese Weise integrierten linearen Elemente, »weiterweisende« und »schließende«, stehen in

einem komplizierten Verhältnis eines stets wieder gebrochenen Gleichgewichts – das sich auch in den aus dem Gefüge des Themas ableitbaren möglichen harmonischen Strukturen spiegelt. So führt der »terminale« Quintfall in Takt 2 auf die dominantische V. Stufe »hinauf«, aber auch die Schlußwirkung am Ende des Themas ist geschwächt (Dort steht nicht der sonst stetig wiederholte Grundton c, sondern die Mollterz auf einer Sechzehntelnote). – Es kann damit ein neuer gestischer Zyklus beginnen (hier als nächstes der zweite Themeneinsatz), so daß sich das Pulsieren schließlich über das ganze Stück hinweg fortsetzt. An dieser Stelle soll nicht geklärt werden, welche gestische Einheit die eigentlich relevante ist, die halbtaktige oder die zweitaktige: Hier soll vielmehr das allgemeine Prinzip exponiert werden.

### **Schlußfolgerungen**

Das hier vorgeschlagene sprachlich-musikalische Melodiekonzept kann folgendermaßen resümiert werden: Sprechmelodie und musikalische Melodie organisieren in sich je auf ihre Weise, aber teils auch mit gemeinsamen Mitteln, ein (hierarchisch geordnetes) Pulsieren von Zeitabschnitten mit verringerter und erhöhter gestischer Bedeutsamkeit. Dies wird möglich durch syntaktische Muster, charakteristische Harmoniefolgen und Klauselbildungen, die Bereitstellung von real auditiv wahrnehmbaren Akzentstellen oder nur mental präsenten metrischen Schemata. Ein weiteres Mittel sind die Variation der Ereignisdichte und (damit) Pausen. In den entsprechend gekennzeichneten Zeitabschnitten werden bei der Produktion musikalische beziehungsweise sprachliche Elemente mit intonatorisch-gestischer Funktion plazierte. Umgekehrt wird, was dort plazierte ist, gestisch gehört. Gestisch relevant ist stets die Zeitstrecke vom Einsetzen der Kennzeichnung bis zum Ende der melodischen Einheit.

Offen bleibt zunächst, wie die intonatorisch-gestische Analyse oder Strukturierung von Musik mit anderen Aspekten der Musikwahrnehmung und -produktion interagiert; hier könnten wiederum Parallelen zur Sprache aufschlußreich sein. Offen bleibt ferner, ob es eine bestimmte musikalische Aggregatebene gibt, die die Interpretation als gestische Einheit besonders anzieht (etwa die Periode, der Halbsatz etc.) und wenn, warum dies so ist.

Die Grundposition dieses Aufsatzes läßt sich in zwei Komponenten zerlegen, eine allgemeine und eine spezifische:

Die allgemeine besagt, daß melodische Phänomene zwei Ebenen haben,

die der Tonhöhe und die der zeitlichen Strukturierung. Diesen Aspekt der zwei Ebenen findet man an vielen Stellen wieder: Er spiegelt sich in der musikalischen Alltagstheorie, wenn beispielsweise von Musik als »Melodie und Rhythmus« gesprochen wird (eine Formulierung, die ihre Wurzeln schon in der antiken Musiktheorie hat, vgl. DAHLHAUS 1961). Er ist in den oben zitierten phonetischen Intonationsmodellen enthalten, die sich so interpretieren ließen. Und er erscheint in musiktheoretischen und musikpsychologischen Ansätzen: Zum Beispiel operiert LERDAHL & JACKENDOFFs Modell (1983) mit einer derartigen Zwei-Ebenen-Struktur: Unterschieden werden dort time-span- und grouping-Analyse, time span reduction und prolongational reduction. In dem gleich nochmals kurz zu erwähnenden Melodie-Modell von NARMOUR (1990) werden die acht möglichen 3-tönigen »melodischen Basisstrukturen« (basic structures) mit Hilfe von nicht in ihnen selbst liegenden Parametern abgegrenzt (other-parametric closure), wobei Faktoren wie Dauer, Metrum, Dissonanz-Konsonanz eine Rolle spielen (NARMOUR 1991). Auch hier werden also (mindestens) zwei Ebenen benötigt, eine im Tonhöhenbereich, eine im Zeitbereich.

Die spezifische Komponente besagt, daß Musik und Sprache den notwendigen formalen Rahmen zweier einander ergänzender Ebenen der Melodie auf ähnliche Weise ausfüllen und daß den grundsätzlichen Gemeinsamkeiten zwischen sprachlichen und musikalischen Melodien damit eine weitere hinzugefügt wird: ihre gestische Struktur, beruhend auf dem intonatorischen Grundgegensatz steigender und fallender Muster, der sich in Sprache mit den als »syntaktische Gesten« operierenden distinktiven Tonhöhenmustern verbindet.

Damit wird eine unmittelbare funktionelle Parallele zwischen sprachlicher Intonation und musikalischer Melodie gezogen.

Ein Versuch, sprachliche Intonationsverläufe aus dem Blickwinkel der musikalischen Melodie nachzuzeichnen, stammt von NARMOUR (1991): NARMOUR unterlegt sprachlichen Intonationskurven (wie sie von GARDING (1981) generiert wurden) seine melodischen »Basisstrukturen«. Die musikalisch-sprachliche Analogiebildung wird auf der molekularen Ebene der Wortakzente und auf der Ebene des Satzes vollzogen. NARMOUR zeigt auf diese Weise die Generalisierungsmöglichkeiten seines Melodiemodells. Auch hier erscheinen Intonation und musikalische Melodie somit aus einem Stoff und werden von denselben Prinzipien determiniert. NARMOUR, der von der Seite der musikalischen Melodie ausgeht und nicht von der sprachlichen, vernachlässigt aber die Dynamik der für die Satzintonation konstitu-

tiven melodischen Gesten und nutzt damit nicht deren möglichen Erklärungswert für musikalische Strukturen. Im vorliegenden Text steht gerade dieser Aspekt im Vordergrund.

Zweck der Darstellung war, zu zeigen, daß es musikpsychologisch Sinn macht, musikalische Melodien im Lichte der Sprechmelodie zu betrachten. Nach der hier vertretenen Auffassung sind die Gestalten der Intonation in Musik als strukturelles und expressives Element inbegriffen.

Ob diese Sichtweise Bestand hat und was daraus für die Verarbeitung, Repräsentation und Produktion musikalischer Melodien folgt, müssen empirische Untersuchungen erweisen. Es liegen bereits einige Befunde vor, die im Rahmen dieses Konzepts gewonnen wurden (DOMBROWSKI 1987, 1992, in Vorbereitung). Sie konzentrieren sich auf die Auswirkung der Satzintonation auf die musikalische Deklamation von Texten: Überprüft wurde, (1) ob sich die erwartete Intonation sprachlicher Sätze in den Melodien niederschlägt, die zu ihnen improvisiert wurden. Untersucht wurde dies an Kindern im Alter von 5 bis 12 Jahren und an Erwachsenen. (2) Analysiert wurde ferner, welche musikalischen Effekte sich ergeben, wenn bei textgebundenen vokalen Improvisationsaufgaben die Aufmerksamkeit unterschiedlich stark auf die Sprechmelodie der Texte gelenkt wird. Die gewonnenen Befunde bestätigten im ganzen die Vorhersagen. Ein interessantes Teilergebnis war, daß charakteristische stilistische Varianten der Leseprosodie in den zum Lesetext erfundenen Melodien reflektiert wurden (DOMBROWSKI, in Vorbereitung). Dieser Befund konnte so interpretiert werden, daß einige der Untersuchungsteilnehmer den zu singenden und zuvor zu lesenden Text als gestische Einheit aufgefaßt haben. Andere dagegen haben, einem rhythmisch-metrischen Schema folgend, die darin enthaltenen Akzentpositionen »abgearbeitet«. Die Teilnehmer haben sich also in unterschiedlichem Maße auf die zeitlich strukturelle und die gestische Ebene der sprachlichen und musikalischen Melodik konzentriert. Die Annahme einer funktionellen Verbindung von gestischen Elementen und zeitlichen Ordnungsmustern in einem Zwei-Ebenen-Modell eröffnet also den Zugang zu kognitiven Stilen und operativen Strategien bei der Wahrnehmung und Produktion von Musik<sup>3</sup>.

Die Betrachtung der musikalisch melodischen Phänomene im Licht sprachlicher Intonation ist mehr als eine bloße äußerliche Umdeutung oder sogar Simplifizierung der musikalischen Sachverhalte: Es ergeben sich daraus - wie aus den kurz erwähnten Beispielen ersichtlich - interessante empirische Fragestellungen.

## Summary

In speech prosody the contrast between rising and falling contours is an important device for the formation of structure. Such contours unfold their functions in a succession of accents and intonational units. The present contribution shows that both statements can be applied analogically to the context of music and that the melodic organization of language is relevant to a psychological comprehension of music.

A starting point for this analogy is that the characteristic use of rising und falling contours is a language universal, originating in elementary expressive behaviour, i.e. in vocal-communicative gestures, which have now been included in linguistic and – as proposed here – in musical processes, as well (cf. OHALA 1983, 1984, BOLINGER 1986).

The melodic continua of language and music are, therefore, viewed as chains of tone gestures, or rather as

- (1) gesturally marked, and
- (2) temporally structured courses (Two-level view).

In this connection, central importance is attributed to the dialectics of progredient-open-rising vs. terminal-closed-falling patterns.

The line of thought, developed here, is illustrated with linguistic and musical examples. It is further related to several theoretical approaches to the psychology of music (e. g. by CLYNES, CLARKE, LERDAHL & JACKENDOFF, SCHERER).

## Anmerkungen

- 1 HANSLICKs Bezeichnung »tönend bewegte Formen« trifft ebenso gut Musik wie die Sprechmelodik: HANSLICK (1854/1989, S. 26) begrenzt die Expressivität der Musik auf ihre Fähigkeit, das Dynamische, das heißt die »Bewegung« der Gefühle, aufzunehmen. Diese lasse sich »in den Momenten schnell, langsam, stark, schwach, steigend, fallend« nachbilden: Gerade das sind aber Gegensatzpaare, die auch für die Beschreibung prosodischer Verläufe zentral sind. Sie betreffen Tonhöhenverlauf, Intensität und Tempo beziehungsweise Dauer.
- 2 Gemeint ist hier die Bezugnahme auf die linguistische Intonationsforschung. Eine gestische Auffassung von Musik legt aber auch ASSAFJEWs (1976/1930/1942) Intonationentheorie nahe, die jedoch in die vorgetragenen Überlegungen nicht einbezogen wurde.
- 3 Die Annahme zweier derartiger kognitiver Stile im Umgang mit musikalischen und sprachlichen melodischen Phänomenen erlaubt interessante musiktheoretisch-systematische und musikhistorische Interpretationen. So kann die von DAHLHAUS (1978) vorgenommene Unterscheidung zweier verschiedener Begriffe der musikalischen Periode den beiden Alternativen zugeordnet werden: Ein am Vers orientierter Periodenbegriff unterstreicht die zeitliche Ordnung, ein rhetorischer Periodenbegriff die gestische Struktur eines melodischen Gefüges.

## Literatur

- ASSAFJEW, B. V. (1976) – *Die musikalische Form als Prozeß*. Moskau Leningrad 1930-1947. Aus dem Russischen von E. Kuhn. Berlin: E. Lippold.
- BECKING, G. (1958/1928) – *Der musikalische Rhythmus als Erkenntnisquelle*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- BEHNE, K.-E. (1982) – *Musik – Kommunikation oder Geste?* In: K.-E. BEHNE (Hrsg.), *Musikpädagogische Forschung*, 3. Laaber: Laaber. S. 125–145.
- BIERWISCH, M. (1971) – *Regeln für die Intonation deutscher Sätze*. In: *Studia Grammatica VII: Untersuchungen über Akzent und Intonation im Deutschen*. 2. Aufl. Berlin: Akademie. S. 99–201.
- BOLINGER, D. (1964) – *Intonation as a universal*. In: H. G. LUNT (ed.): *Proceedings of the 9th International Congress of Linguists*. The Hague: Mouton. S. 833–848.
- BOLINGER, D. (1978) – *Intonation across languages*. In: J. P. GREENBERG, C. A. FERGUSON & E. A. MORAVCSIK (eds.): *Universals of human language*, vol. 2. Phonology. Stanford: Stanford University Press.
- BOLINGER, D. (1986) – *Intonation and its parts: Melody in spoken English*. London: Edward Arnold.
- BOLINGER, D. (1989) – *Intonation and its uses*. Stanford: Stanford University Press.
- BRUCE, G. (1977) – *Swedish word accents in sentence perspective*. (=Travaux de l'Institut de Linguistique de Lund XIII).
- CARLSON, R., FRIBERG, A., FRYDÉN, L., GRANSTRÖM, B. & SUNDBERG, J. (1989) – *Speech and music performance: Parallels and contrasts*. *Contemporary Music Review*, 4, 391–404.
- CARLSON, R., GRANSTRÖM, B. & HUNNICUTT, S. (1990) – *Multilingual text-to-speech development and applications*. In: W. A. Ainsworth (ed.): *Advances in Speech, Hearing, and Language Processing*, Vol. I. London: JAI Press, 269–296.
- CHOMSKY, N. & HALLE, M. (1968) – *The sound pattern of English*. New York: Harper & Row.
- CLARKE, E. F. (1987) – *Levels of structure in the organization of musical time*. *Contemporary Music Review*, 2, 211–238.
- CLYNES, M. (1977) – *Sentics: The touch of emotions*. New York: Anchor/Doubleday.
- CLYNES, M. (1986) – *Generative principles of musical thought: Integration of microstructure with structure*. *Communication and Cognition*, CCAI, 3, 185–223.
- CLYNES, M. & NETTHEIM, N. (1982) – *The living quality of music: Neurobiologic basis of communicating feeling*. In: M. CLYNES (ed.): *Music, mind, and brain: The neuropsychology of music*. New York: Plenum, pp. 47–82.
- CRUTTENDEN, A. (1981) – *Fals and rises: meanings and universals*. *Journal of Linguistics*, 17, 77–91.
- CRYSTAL, D. (1987) – *The Cambridge encyclopedia of language*. Cambridge: Cambridge University Press.
- DAHLHAUS, C. (1961) – *Melodie: Systematisch*. In: F. BLUME (Hrsg.), *Die Musik in Geschichte und Gegenwart*, 9, Kassel: Bärenreiter. S. 25–55.
- DAHLHAUS, C. (1978) – *Satz und Periode: Zur Theorie der musikalischen Syntax*. *Zeitschrift für Musiktheorie*, 9, 16–26.
- DEUTSCH, D. (1991) – *The tritone paradox: An influence of language on music perception*. *Music Perception*, 8, 335–347.
- DEUTSCH, D. (1994) – *The tritone paradox: Some further geographical correlates*. *Music Perception*, 12, 125–136.
- DOLSON, M. (1994) – *The pitch of speech as a function of linguistic community*. *Music Perception*, 11, 321–331.
- DOMBROWSKI, E. H. (1987) – *Musik und Sprache: Psychologische Untersuchungen zu ihrem Verhältnis am Beispiel der musikalischen Abbildung der sprachlichen Satzartkategorien*. Unveröffentlichte Diplomarbeit. Institut für Psychologie der Universität Kiel.
- DOMBROWSKI, E. H. (1992) – *Satzintonation und Melodie in textierten Improvisationen: Ein entwicklungspsychologischer Vergleich*. *Musikpädagogische Forschung*, 13, 224–234. (Hermann J. Kaiser (Hrsg.), *Musikalische Erfahrung: Wahrnehmen, Erkennen, Aneignen*).
- DOMBROWSKI, E. H. (In Vorbereitung). *Zum Einfluß der Satzintonation auf das musikalische Verhalten*.

- DRAKE, C. & PALMER, C. (1993) – *Accent structures in music performance*. *Music Perception*, 10, 343–378.
- ESSEN, O. v. (1964) – *Grundzüge der hochdeutschen Satzintonation*. Ratingen: Henn-Verlag.
- GEORGIADES, T. (1954) – *Musik und Sprache: Das Werden der abendländischen Musik*. Berlin: Springer.
- GRUHN, W. (1978) – *Musiksprache, Sprachmusik, Textvertonung: Aspekte des Verhältnisses von Musik, Sprache und Text*. Frankfurt: Diesterweg.
- HALLIDAY, M. A. K. (1967) – *Intonation and grammar in British English*, The Hague: Mouton.
- HANSLICK, (1989/1854) – *Vom musikalisch Schönen*. 21. Aufl. Wiesbaden: Breitkopf & Härtel.
- HELFRICH, H. (1985) – *Satzmelodie und Sprachwahrnehmung: Psycholinguistische Untersuchungen zur Grundfrequenz*. Berlin: de Gruyter.
- HERDER, J. G. (o. J./1769) – *Viertes kritisches Wäldchen*. In: H. Düntzer (Hrsg.), *Werke*, Bd. 20. Berlin.
- JACKENDOFF, R. (1991) – *Musical parsing and musical affect*. *Music perception*, 9, 199–230.
- KATZ, J. J. & POSTAL, P. H. (1964) – *An integrated theory of linguistic descriptions*. Cambridge, Mass. MIT.
- KOHLER, K. J. (1977) – *Einführung in die Phonetik des Deutschen*. Berlin: Erich Schmidt.
- KOHLER, K. J. (1991a) – *A model of German intonation*. *Arbeitsberichte des Instituts für Phonetik der Universität Kiel (AIPUK)*, 25, 295–360.
- KOHLER, K. J. (1991b). *Form and function of intonation peaks in German: A research project*. *Arbeitsberichte des Instituts für Phonetik der Universität Kiel (AIPUK)*, 25, 11–28.
- KOHLER, K. J. (1992) – *Prosodisches Transkriptionssystem für die Etikettierung von Sprachsignalen*. *Arbeitsberichte des Instituts für Phonetik der Universität Kiel (AIPUK)*, 26, 239–250.
- LADD, D. R. (1983) – *Phonological features of intonational peaks*. *Language*, 59, 721–759.
- LERDAHL, F. (1991) – *Underlying musical schemata*. In: Peter HOWELL, Robert WEST, Ian CROSS (Eds.), *Representing musical structure*. London: Academic Press. S. 273–290.
- LERDAHL, F. & JACKENDOFF, R. (1983) – *A generative theory of tonal music*. Cambridge, Mass.: MIT.
- LIEBERMAN, P. (1974) – *A study of prosodic features*. In: T. A. SEBOEK (ed.), *Current trends in linguistics*, Vol. 12. Den Haag: Mouton. S. 2419–2449.
- MEINHOLD, G. & STOCK, E. (1982) – *Phonologie der deutschen Gegenwartssprache*.
- NARMOUR, E. (1990) – *The analysis and cognition of basic melodic structures: The implication-realization model*. Chicago: The University of Chicago Press.
- NARMOUR, E. (1991) – *The melodic structures of music and speech: Application and dimensions of the implication-realization model*. In: Johan SUNDBERG, Lennart NORD & Rolf CARLSON (Eds.), *Music, language, speech, and brain*, Houndsmills, Basingstoke UK: Macmillan. S. 48–56.
- OHALA, J. (1983) – *Cross language use of pitch: An ethological view*. *Phonetica*, 40, 1–18.
- OHALA, J. (1984) – *An ethological perspective on common cross-language utilization of  $f_0$  of voice*. *Phonetica*, 41, 1–16.
- PIERREHUMBERT, J. B. (1987) – *The phonetics of English intonation*. Bloomington: IULC.
- SCHENKER, H. (1935) – *Der freie Satz*. Wien: Universal-Edition.
- SCHERER, K. R. (1989) – *Vocal correlates of emotion*. In: H. L. WAGNER & A. S. R. Manstead (eds.): *Handbook of Psychophysiology: Emotion and social behavior*. Chichester: Wiley. S. 165–197.
- SCHERER, K. R. (1991) – *Emotion expression in speech and music*. In: J. SUNDBERG, L. NORD & R. CARLSON (Eds.), *Music, language, speech, and brain*. Houndsmills, Basingstoke UK: Macmillan. S. 146–156.
- SCHERER, K. R. & BERGMANN, G. (1984) – *Vocal communication*. *The German Journal of Psychology*, 8, 57–90.
- SCHERER, K. & OSHINSKY, J. (1982/1977) – *Zur emotionalen Eindruckswirkung akustischer Reizparameter*. In Klaus Scherer (Hrsg.), *Vokale Kommunikation* (S. 326–342), Weinheim: Beltz.
- TRÄGER, G. L. & SMITH, H. L. (1951) – *An outline of English structure* (*Studies in Linguistics: Occasional Papers 3*). Norman, Okla.
- WINKLER, C. (1979) – *Untersuchungen zur Kadenzbildung in deutscher Rede*. München: Hueber.