

## Ästhetische Phonetik als Wissenschaft von der ästhetischen Kommunikation in Sprache und Sprachmusik

Georg Heike

### 1. Vorläufige Begriffsbestimmung

Unter ästhetischer Kommunikation verstehen wir eine Übermittlung von ästhetischer Information, d. h. von Information, bei der es nicht in erster Linie um die Darstellung von Sachverhalten und um sprachliche Inhalte geht. Wenn wir das Organon-Modell von Bühler bzw. in der weiterentwickelten Form als Ebenen-Modell (Heike 1969) zur weiteren Erklärung des Gemeinten heranziehen, so können wir z. B. sagen, daß es in der ästhetischen Kommunikation nicht um die Übermittlung von Signalen geht, die durch Konventionalisierung einen eindeutigen Bezug zur Außenwelt oder zu vereinbarten geistigen Inhalten besitzen. Worum geht es also in der ästhetischen Kommunikation, worüber informiert die ästhetische Information? In einem ersten Versuch können wir folgendes sagen: Ästhetische Kommunikation findet erstens nicht in alltäglichen Kommunikationssituationen statt, sondern in künstlich herbeigeführten Situationen; zweitens muß im allgemeinen eine ästhetische kommunikative Absicht des Expedienten (Künstlers) vorausgesetzt werden. Einem Kommunikationsakt dieser Art liegt also eine Intention des Künstlers zugrunde, die in ihren ästhetischen Bezügen prinzipiell unbekannt und voraussetzungslos ist. Unter der Voraussetzung, die im Übrigen für alle Ausführungen im Zusammenhang mit dieser Darlegung gilt, daß ästhetische Kommunikation grundsätzlich kulturspezifisch ist und wir hier nur für den abendländisch-europäischen Bereich sprechen, kann man dennoch wenigstens zwei Aspekte aufführen, durch die sich ästhetische Information charakterisieren läßt: Das ist zum einen der emotionale Aspekt und zum zweiten der formale. Bei der Diskussion des ersten Aspekts wäre bis zu einem gewissen Grade die Apellfunktion (Bühler) heranzuziehen, indem man sie auf konventionalisierte und gar stilisierte Formen expressiver Signale erweitert. Festzustellen ist also, daß bei ästhetischen Signalen wie allgemein bei expressiven Signalen in Sprache und auch wohl Musik, ein Zusammenhang besteht zwischen Signaleigenschaften, d. h. dem äußeren Erscheinungsbild, der äußeren Gestalt von Signalen auf der einen und expressiven oder emotionellen Gegebenheiten auf der anderen Seite. Dieser Zusammenhang hat sicherlich eine allgemein menschliche Basis, hat also universelle Geltung, diese kann jedoch und wird in der Regel kulturspezifisch und sprachspezifisch konventionalisiert, überformt und verändert. Es können sicherlich auch neue kulturspezifische Konventionen entstehen, die also im Prinzip arbiträr wären wie die linguistischen Zeichen. Die Konventionalisierung und Stilisierung kann so weit gehen, daß sie über Jahrhunderte durch eine strenge Tradition weitgehend unverändert übermittelt wird (Peking-Oper) und sich dadurch so weit von der biologischen oder psychischen Basis entfernt hat, daß sie jeweils neu erlernt werden muß.

Unter dem formalen Aspekt verstehen wir, daß die äußere Form, d. h. die Geordnetheit eines Kunstwerks, eine Information in sich darstellt. Selbst auch für den Fall, daß jegliche Geordnetheit vermieden wird, liegt offensichtlich der ästhetische Reiz gerade in der Negation von Ordnung und von konventionalisierten Ordnungsschemata (in der Musik z. B. bei John Cage). Zeichen können nicht nur wie in der Sprache und in der Linguistik definiert einen eidetischen Sinn haben, sondern, wie schon mehrmals bemerkt (Meyer-Eppler), auch einen operativen. Dazu gehören z. B. die Regeln, nach denen Zeichen als Elemente einer Nachricht, welchen Inhalts auch immer, angeordnet werden. Bei ästhetischen Nachrichten wird offenbar mit den Regeln spielerisch umgegangen, bzw. es werden neue Regeln geschaffen, die durch Wiederholungen gefestigt und eben dadurch als solche erkennbar gemacht werden. Der ästhetische Reiz formaler Beziehungen zwischen Zeichen und Zeichenfolgen besteht zum einen im Erkennen dieser Beziehungen und zum anderen darin, daß ein Spiel zwischen Erwartung und Überraschung stattfindet. Eingeschränkt auf den speziellen Fall sprachlicher Information kann man mit dem Begriffsinventar der Informationstheorie den Unterschied zwischen Haupt- und Trägerinformation machen, um das zuletzt Gesagte zu verdeutlichen. Unter Hauptinformation versteht man das, worum es in der Informationsübermittlung in erster Linie geht, nämlich um die Übermittlung geistiger Inhalte und Bezüge zur Außenwelt, also um die semantische Information, und unter Trägerinformation versteht man die strukturellen Voraussetzungen dafür, d. h. die Ordnung und Struktur zwischen den Elementen der Zeichen, die notwendig ist, damit die Zeichen unterschieden, identifiziert und als Träger der Hauptinformation nutzbar gemacht werden (Struktur des Codes). Die Trägerinformation ist also strukturiert und geordnet und folgt definierten Regeln, deren Ermittlung eine Hauptaufgabe der Linguistik ist. Beispiele zur Verdeutlichung des Gesagten wären unter anderem Wortspiele, Lautspiele - überhaupt der spielerische Umgang mit der Struktur von Sprache. Es ist deshalb auch nicht verwunderlich, daß z. B. im Dadaismus der konsequente Schritt, nämlich die Beziehung zwischen Träger- und Hauptinformation aufzugeben und allein innerhalb der Strukturiertheit der Trägerinformation zu bleiben, gemacht worden ist.

Es ist klar, daß wir uns in diesem Zusammenhang in erster Linie mit poetischer Sprache, d. h. also mit Sprache für bestimmte Zwecke und Situationen, für Sprache also als künstlerisches Ausdrucksmittel, befassen.

Unter Sprachmusik verstehen wir, daß über das Aufgeben der semantischen Funktion von Sprache hinausgegangen wird, d. h. der Stand des Dadaismus und der bekannten sogenannten Lautgedichte wesentlich weiter vorangetrieben wird. Verallgemeinernd kann man sagen, daß sämtliche phonetischen Parameter, die zur Artikulation von Sprache benötigt werden, auch davon losgelöst als musikalische Parameter nutzbar sind. Wie dies im einzelnen geschehen kann, soll im dritten Abschnitt dieses Beitrags ausführliche behandelt werden. An dieser Stelle wollen wir nur noch einmal festhalten, daß bei der Aufgabe der sprachlichen Funktion gesprochener Sprache in dem eben skizzierten extensiven Ausmaß bald die Frage auftaucht, ob die klanglichen Resultate noch als sprechsprachlich erzeugt wahrgenommen werden. Mit anderen Worten: Man muß sich die Frage stellen, ob es eine Merkmalskategorie "sprachlich" gibt, oder ob vielmehr auch damit zu rechnen ist, daß zwischen den Extremen "sprachlich erzeugt" und auf der anderen Seite "nicht sprachlich erzeugt" ein Kontinuum besteht, das auch als solches erlebt werden kann.

## 2. Die ästhetische Kommunikationskette

Die ästhetische Kommunikationskette weist im Gegensatz zur lautsprachlichen Kommunikationskette einige Besonderheiten auf. Während bei der lautsprachlichen Kommunikation die Nachrichten im Normalfall in beiden Richtungen abwechselnd übermittelt werden (Dialog), ist dies bei der ästhetischen Informationsübermittlung nicht der Fall. In der Regel kann man daher davon ausgehen, daß der Empfänger ästhetischer Nachrichten nur über eine passive Kompetenz verfügt und deshalb im Gegensatz zur sprachlichen Kommunikation gar nicht in der Lage ist, selbst Signale auszusenden. Der Expedient erwartet zwar Reaktionen und in irgend einer Form eine Rückmeldung vom Perzipienten, vorausgesetzt, daß eine kommunikative Absicht vorliegt, jedoch erfolgt diese Rückmeldung oder Reaktion nicht in der gleichen kommunikativen Situation. Ein weiterer Unterschied besteht darin, daß die Nachricht des Senders einer Interpretation bedarf, jedenfalls im Normalfall, d. h. außer dem eigentlichen Sender muß es einen Interpreten geben, der die Aufgabe hat, die Nachricht zu realisieren und auf diese Weise zum Perzipienten zu übertragen. In solchen Fällen besteht die Nachricht also in einer symbolischen, weitgehend abstrakten Repräsentation (Partitur bzw. Text in der allgemeinsten Bedeutung), die der Interpretation und der Realisation bedarf. Das besondere an dieser Zweiteilung in Autor und Interpret ist, daß bei der Interpretation in diesem Fall Information hinzukommt und dadurch beliebige Varianten der Realisation entstehen können. Obwohl in der Regel Information hinzutritt, kann natürlich auch intendierte Information des Autors fehlen, bzw. in einer Weise realisiert sein, die nicht seinen Absichten entspricht. Es ist aber auch mit dem Fall zu rechnen, und das ist sicherlich nicht selten, daß die hinzutretende Information für den Autor eine positive Überraschung darstellt. Die Diskrepanz im positiven oder negativen Sinne zwischen intendierter Nachricht und Realisation mag selbst dann gegeben sein, wenn Autor und Interpret identisch sind. Diese Diskrepanz wird dann aufgehoben, wenn die elektronisch oder in sonstiger Art und Weise erzeugte Musik oder Sprache auf Tonträgern abgespeichert vorliegt und/oder öffentlich vorgeführt wird. Interessanterweise gibt es kaum Beispiele elektronisch erzeugter Sprachkunstwerke, das mag aber daran liegen, daß die ästhetische und natürliche Qualität synthetischer Sprache bzw. elektronisch erzeugter Sprache bis dato unbefriedigend ist. Die Erfahrungen mit elektronisch bzw. maschinell erzeugten synthetischen ästhetischen Signalen haben jedoch gezeigt, daß das Ausbleiben der interpretatorischen Varianz (jede Realisation bzw. genauer gesagt jede Wiedergabe ist physikalisch im Prinzip mit jeder anderen völlig identisch) in ästhetischer Hinsicht unbefriedigend ist. Die Folge dieser Erfahrung war, daß in der Regel elektronisch erzeugte Signale mit natürlichen, durch Interpreten erzeugten Signalen gemischt verwendet werden: Eine Auflösung dieses Dilemmas, das darin besteht, daß der Autor auf der einen Seite das hörbare Resultat so weit wie möglich definieren möchte, auf der anderen Seite die Lebendigkeit der Interpretation, die durch Abweichungen und Variationen zustande kommt, nicht missen möchte, wäre folgendermaßen zu bewerkstelligen: Würde es gelingen, bzw. würde die Technologie es ermöglichen, Computerprogramme zu entwickeln, die in der Lage wären, auch den Interpretationsvorgang in ästhetisch ausreichender Weise zu simulieren, so könnte im Kontext der Möglichkeiten dieses Programmes auch der Autor die Interpretation einschließlich ihrer Variationsmöglichkeiten bestimmen. Im Falle von Sprachkunstwerken würde das bedeuten, daß der Sprachproduktionsprozeß, in einem Computerprogramm simuliert, sämtliche Parameter gesprochener Sprache erfassen müßte, d. h. also nicht nur verständliche und natürlich klingende Sprache, sondern auch sämtliche diagnostischen, expressiven und zufallsbestimmten Eigenschaften. Mit den bereits entwickelten Möglichkeiten der Sprachsynthese durch ein Simulationsmodell des natürlichen Sprachproduktionsprozesses (Kölner Modell) wäre dieses Ziel im Prinzip zu erreichen, vorausgesetzt allerdings, daß ausreichende Information, d.h. also wissenschaftliche Forschungsergebnisse für diese Bereiche gesprochener Sprache vorliegen, was zur Zeit nicht der Fall ist.

Wie bei jeder Kommunikationskette muß auch im Falle der ästhetischen Kommunikation Bilateralität gegeben sein, d. h. zwischen Expedienten und Perzipienten muß zusätzlich zu der physikalischen Verbindung durch das übertragene Signal auch eine nichtmaterielle Verbindung durch einen gemeinsamen Code, durch ein gemeinsames System, z. B. Sprachsystem, vorliegen. Wie umfangreich der gemeinsame Sprachbesitz (in der allgemeinsten Bedeutung des Wortes) sein muß, ist im Falle der ästhetischen Kommunikation bis dato weitgehend unbekannt. Man kann jedoch davon ausgehen, daß ein Minimum an Übereinstimmung gegeben sein muß, damit eine Kommunikation, und sei sie auch noch so reduziert, stattfinden kann. Zu den Besonderheiten ästhetischer Nachrichten gehört es jedoch, daß der gemeinsame Sprachcode sich ständig ändert bzw. sogar von Nachricht zu Nachricht verändert wird (als Folge des Innovationsanspruchs), so daß vom Perzipienten nicht nur ein Lernvorgang, sondern auch ein Interpretationsvorgang auf der Basis des ihm bis dahin gegebenen Wissens abverlangt wird. Dieser Lernvorgang, der auch Entdeckungsprozeduren beinhaltet, macht jedoch das Besondere der ästhetischen Kommunikation heutzutage aus. Es liegt in dem ungewöhnlich hohen Anspruch, der an die Lernfähigkeit und Lernbereitschaft und Interpretationsfähigkeit des Perzipienten gestellt wird. Man muß jedoch auch die Relativität dieses hohen Anspruchs in Betracht ziehen, d. h. man muß davon ausgehen, daß die Bereitschaft zum Entdecken und Aufnehmen neuer und ungewöhnlicher ästhetischer Produkte heute sehr gering ausgeprägt ist, weil im Gegensatz zu früheren kulturellen Epochen die Rezeption bzw. Konsumtion ästhetischer Produkte aus der Vergangenheit im Vordergrund stehen. Der bekannte Dirigent Harnoncourt

hat mit Recht das Augenmerk darauf gerichtet, daß es beispielsweise zu Lebzeiten Mozarts unvorstellbar gewesen wäre, ein Konzert mit Werken ausschließlich barocker Komponisten zu geben. Für Barockmusik wären weder die Voraussetzungen zum Verständnis noch ein Interesse bei den Zuhörern vorhanden gewesen. Diesen Sachverhalt kann man vielleicht damit erklären, daß in früheren Jahrhunderten die Musik einen sehr viel größeren Stellenwert als heutzutage hatte. Sie diente nicht dem Konsum, sie war nicht jederzeit und beliebig verfügbar, sondern jeweils ein besonderes Ereignis, daß nicht nur einen Neuigkeitswert hatte, sondern auch haben mußte.

In der zeitgenössischen Kunst dieses Jahrhunderts ist es offenbar erstmals so, daß in Sonderfällen bei dem Exponenten keine kommunikative Absicht im eigentlichen Sinne vorzuliegen braucht. Dies ist gegeben, wenn ästhetische Signale ausschließlich aufgrund von Zufallsmanipulationen oder anderen Manipulationen komplexer Art (beispielsweise durch hochkomplexe Computerprogramme) erzeugt werden, so daß die Frage nach der Intention solcher ästhetischer Signale außerhalb eines kommunikativen Ansatzes im engeren Sinne beantwortet werden muß. Im Rahmen der Terminologie der Informationstheorie liegt in einem solchen Fall eben auch keine Kommunikationskette vor, sondern eigentlich nur eine Beobachtungskette. Das bedeutet, daß der Rezipient oder Perzipient davon ausgehen muß, daß die ästhetischen Signale keinerlei Bezug zu irgendeinem gemeinsamen sprachähnlichen Code aufweisen und deshalb nur als physikalische Signale ohne ästhetische Intention eines intelligenten Lebewesens angesehen werden müssen. Dies ist der Grund, weshalb hier keine Kommunikation im strengen Sinne, sondern höchstens eine Beobachtung möglich ist. Signale dieser Art, wie sie auch in der Umwelt vorkommen, können nur beobachtet und eventuell in Beziehung zu ihrer Hervorbringung betrachtet und erklärt werden. Wenn solche Signale dennoch als ästhetische Signale in Betracht kommen, so aus folgenden Überlegungen:

Zum einen genügt es vielleicht bereits, wenn sie als ästhetische Signale deklariert und in einer Situation einem Publikum dargeboten werden, in der normalerweise ästhetische Kommunikation stattfindet. Aufgrund dieser Voraussetzung kann ästhetische Information in verschiedenster Hinsicht zum Tragen kommen: Durch die Negation ästhetischer Konventionen oder auch durch die Auswahl der Erzeugungsprozeduren, wobei vor allem gegebenenfalls der ausdrückliche Hinweis auf die Erzeugungsprozedur im Titel oder Kommentar Assoziationen und ästhetische Interpretationen der zufällig erzeugten Signale beim Perzipienten bewirken können. Zwei hier willkürlich ausgewählte Beispiele wären die folgenden: Ein Stück mit dem Titel "Wassermusik", bei dem die zufällige Reihenfolge von Regentropfen für die klangliche Struktur maßgeblich wäre, oder eine Komposition mit dem Titel "Klang der Sterne", die sich aus der Übertragung einer Photographie des Sternenhimmels auf Notenpapier ergäbe. Auch ein Text aus Buchstaben oder Lauten ließe sich herstellen durch eine einmal gewählte Codierung des Sternenhimmels in die Buchstaben eines Alphabets bzw. in die Laute eines Phoneminventars. Zur Kategorie ästhetischer Signale dieser Art gehören auch solche Erzeugungsprozeduren, bei denen der Erzeugungsmechanismus zwar nicht zufällig, sondern in irgendeiner Weise geordnet erfolgt, die Auswirkungen dieser Ordnung auf das ästhetische Resultat jedoch unvorhersehbar sind. Beispiele für Kunstprodukte dieser Art sind solche, bei denen die Relationen zwischen den Buchstaben eines Textes, die sich ja aus der Geordntheit der Sprache ergeben, durch eine willkürlich festgesetzte Codiervorschrift in musikalische Töne umsetzen lassen (Klarenz Barlow, Gerhard Rühm). Die Tatsache jedoch, daß in solchen Fällen der inhaltliche oder sonstige Kontext von Texten beabsichtigt ist, schafft unter Umständen ästhetisch reizvolle Assoziationen bei der Rezeption (beispielsweise in 'Das Leben Chopins' von Rühm, eine Umsetzung der Biographie Chopins in Klaviermusik).

Ein weiterer Grund für die ästhetische Wirkung von Zufallsprodukten bzw. Produkten unbekannter Ordnung liegt darin, daß dem Perzipienten dieser Sachverhalt nicht in voller Bedeutung bewußt ist und er deshalb bei der künstlerischen Präsentation solcher Signale geneigt ist, eine potentielle ästhetische Absicht vorzusetzen. Entscheidend tritt ein Charakteristikum der menschlichen Perzeption hinzu, nämlich der sog. Zwang zur Gestalt, der darin besteht, daß dargebotene Signale jedweder Art grundsätzlich auf dem Hintergrund und im Kontext der Erfahrung interpretiert und mit Bekanntem identifiziert werden. Hinzu kommt, daß zufällig erzeugte Signale eine statistische Gleichverteilung nur bei einer hinreichend großen Zahl, also über einen längeren Zeitraum hinweg aufweisen. Im Einzelfall können durchaus, wenn auch wiederum zufallsbedingt, Häufungen auftreten, die den Eindruck einer Gestaltung und Ordnung erwecken.

Für die menschliche Perzeption gilt anscheinend allgemein der Sachverhalt, daß ein ständiger Bezug zum Erinnerungspotential vorgenommen wird, wobei vermutlich die ursprüngliche biologische Funktion der auditiven Wahrnehmung im Vordergrund steht. Das Hören ist offensichtlich nicht primär für die interhumane Kommunikation genetisch entwickelt, sondern für die Beobachtung der Umwelt im Hinblick auf das Überleben. Für das Überleben in einer gefährlichen Umwelt ist es ungeheuer wichtig, die akustischen Signale daraufhin zu interpretieren, ob ihre Verursacher eine Gefährdung darstellen können oder nicht. Die Erfahrung hörender Lebewesen lehrt, daß Schallsignale in einem ursächlichen Zusammenhang mit ihrer Hervorbringung stehen. Sie lehrt auch, daß manche Signale eine Bedrohung durch die Signalquelle bedeuten können und andere nicht. Diese Zusammenhänge zwischen Signaleigenschaften und lebenswichtiger Information sind offenbar bisher kaum erforscht. Wir geben einige drastische Beispiele, die der Deutlichkeit halber so gewählt sind, daß Konnotation und Signaltypus sich erfahrungsgemäß ausschließen und in dieser Kombination also nie auftreten können: "bedrohliches" Schnurren, "liebvolles" Donnerkrachen, "zärtliches" Klappern (einer Schlange), "aufschreckendes" Säuseln im Blätterwald, "beruhigendes" Sturmgeheul etc.

### 3. Sprachmusik: Komponieren mit phonetischem Material

Der Begriff Sprachkomposition datiert aus den fünfziger Jahren. Wir meinen damit den Sachverhalt der Musikalisierung phonetischer und linguistischer Parameter. Diese Verfügbarmachung der Sprachparameter für kompositorische Zwecke ist in Konsequenz durch das serielle Prinzip erst möglich geworden: Serielles Denken heißt, Abstufungen in allen freien Parametern so kompositorisch zu verwenden wie das traditionellerweise beim Tonhöhenparameter der Fall ist.

Zu den linguistisch-phonetischen Parametern gehören, um nur die wichtigsten zu nennen: die Verständlichkeit (also alle Abstufungen der Verständlichkeit zwischen den Extremen "maximal verständlich" und "unverständlich"), die Sprachzugehörigkeit oder Sprach- bzw. Dialektfärbung (gradueller Abstufungen der Ausprägung einer Dialektfärbung), schließlich auch die Art der Signalkategorie als Abstufung zwischen den Extremen "sprachlich erzeugtes Signal" und "nicht sprachlich erzeugtes Signal". Dieser Parameter, wobei nicht-sprachlich erzeugtes Signal identisch mit elektronischem Signal ist, wurde erstmals in dem "Gesang der Jünglinge" von Stockhausen komponiert, ebenso wie der Parameter "Verständlichkeit".

Wenn wir zunächst bei der Kategorie der sprachlich erzeugten Signale bleiben, so bietet sich als nächste Differenzierung die nach den bekannten ausschließlich phonetischen Parametern an, als da sind Stimmtonhöhe, Signalamplitude, Grad der Periodizität zwischen den Extremen "quasiperiodisch" und "stochastisch" (turbulenter Luftstrom), die Kategorie Dauerlaut zwischen den Extremen "kontinuierlich dauernd" und "abrupt", der Grad der Sonorität bzw. Vokalität auf der Skala zwischen vokalisch und konsonantisch, Merkmale der Vokalfarbe (hell - dunkel, kompakt - diffus etc.). Im Kontext dieser Überlegungen der Parametrisierung phonetischer Merkmale zeigt es sich, daß die in der Linguistik geläufigen polaren Merkmale bzw. Binärmerkmale der Phonologie eigentlich ein phonologisches Konstrukt sind und letztlich keine phonetische Basis besitzen. Anstelle der soeben verwendeten Terminologie vorwiegend aus dem akustischen und auditiven Bereich wäre es auch angebracht, die entsprechende Terminologie aus dem Bereich der phonetischen Genese (Phonation und Artikulation) zu benutzen. Diese Terminologie erscheint anschaulicher, weil sie den direkten Zusammenhang zwischen phonetischen Eigenschaften und deren Hervorbringung bzw. Aktionen zur Hervorbringung beschreibt. Die entsprechende Terminologie würde folgende Begriffe aufweisen: stimmhaft - stimmlos, phonatorischer Energieaufwand (Stimmbandspannung und subglottaler Druck), Tonhöhe (als empfundene Kehlkopfhöhe), Verengung, Hindernisbildung, Verschlussbildung, Nasalität, Labialität, hohe bzw. tiefe Zungenposition, vordere bzw. hintere Zungenposition etc.

Hinsichtlich der akustischen Signalkategorien kann man feststellen, daß es keinen grundsätzlichen Unterschied zwischen Sprachklängen und musikalischen Klängen gibt. Den geräuschhaften phonetischen Signalen entsprechen beispielsweise in der Musik geräuschhafte Instrumente ohne definierte Tonhöhe, der Plosivlautartikulation entsprechen impulsartige Tonerzeugungen, den Klangfarbenänderungen durch artikulatorische Veränderung der Resonanzbedingungen der Artikulation entsprechen in der Musik Klangfarbenänderungen durch Verwendung unterschiedlicher Instrumente, unterschiedlicher Spielweisen oder unterschiedlicher Register. Ein grundsätzlicher Unterschied zwischen Sprache und Musik besteht jedoch hinsichtlich der Strukturierung der Parameter zur Erzeugung von Signalen in Abhängigkeit von der Zeit. Zur Unterstützung dieser These können wir folgende Fakten aufführen:

1. Phonetische Parameter nehmen nur selten extreme Positionen ein. Sie schwanken in der Regel um Mittelwerte. Dies gilt allerdings jeweils nur innerhalb eines Parameters. Man kann also sagen, daß phonetische Parameter in gesprochener Sprache nicht optimal ausgenutzt sind hinsichtlich der Differenzierungsmöglichkeiten. Damit wird eine hohe Redundanz garantiert, die für die Verständlichkeit auch erforderlich ist.
2. Phonetische Parameter weisen in Abhängigkeit von der Zeit keine abrupten Sprünge auf.
3. Zwischen unterschiedlichen phonetischen Kategorien gibt es in der Regel keine kontinuierlichen Übergänge, d. h. die kategorialen Unterschiede sind signalphonetisch abgesichert.
4. Zwischen phonetischen Parametern liegt in der Regel eine Abhängigkeit vor, die allerdings unterschiedliches Ausmaß haben kann. Das bedeutet im Gegensatz zur musikalischen Parameterstrukturierung, daß nicht in jedem Parameter frei in den Parameterwerten variiert werden kann.
5. Was unter 1. die Strukturierung innerhalb eines Signalparameters angeht, trifft auch für die Strukturierung der Zeit zu. In gesprochener Sprache sind die Veränderungen in Abhängigkeit von der Zeit prinzipiell nicht so extrem strukturierbar wie in der Musik, d. h. es gibt nicht extrem lange Zeitdauern für bestimmte Parameterwerte und auch nicht extrem kurze. Das gleiche gilt auch für Signalpausen. Treten in gesprochener Sprache durch Zögerungsphänomene Pausen auf, so besteht die Tendenz, diese Pausen durch nicht-sinnvolle Äußerungen zu füllen (durch Verlegenheitslaute: 'gefüllte' Pausen).
6. Phonetische Signale sind typischerweise durch Klangfarbenmodulation charakterisiert, d. h. gerade die Veränderungen der Klangfarbe als Folge artikulatorischer Bewegungen sind es, die die Codierung sprachlicher Informationen ermöglichen. Man kann diesen Sachverhalt physikalisch so ausdrücken, daß es sich bei gesprochener Sprache um eine etwa im Silbenrhythmus (Öffnungs- und Schließbewegung) ablaufende gleichzeitige Modulation in Frequenz und Amplitude handelt.

Aufgrund der genannten Charakteristika phonetischer Signale erscheint es reizvoll, die von der musikalischen Signalproduktion sich unterscheidenden Charakteristika für innovative musikalische Strukturen zu nutzen, was in der Musikentwicklung der letzten Zeit auch schon gelegentlich zu beobachten ist (Steve Reich). Anstelle, wie das häufig geschieht, abgespeicherte phonetische Signale zu transformieren bzw. zu verarbeiten und als Elemente eines musikalischen Zusammenhangs zu benutzen, würde es sich sinnvollerweise anbieten, die sprechsprachlichen Veränderungen von Parametern als Funktion der Zeit per Analyse zu extrahieren und zur Strukturierung musikalischer Parameter zu verwenden. Das Ziel wäre mit anderen Worten die Musikalisierung phonetischer Parameter auf der Basis einer parametrischen Analyse und einer Darstellung der Parameterverläufe als abstrakte Funktion der Zeit. Um nur ein Beispiel zu nennen: Eine genaue Tonhöhenanalyse sprachlicher Äußerungen (also die Frequenz der Stimmlippen-schwingungen als Funktion der Zeit dargestellt, in der Phonetik auch gelegentlich Sprachmelodie-Verlauf genannt) ließe sich als Muster für weitere Verarbeitungen in musikalischen Zusammenhängen benutzen, da Frequenzwerte als Funktion der Zeit beliebig veränderbar, transponierbar und manipulierbar sind. Man könnte auf diese Weise getreu dem

seriellen Strukturierungsprinzip zwischen eindeutig phonetischen Tonhöhenverläufen und solchen, die im Zusammenhang der musikalischen Komposition bevorzugt werden, vermitteln und damit einen neuen strukturierbaren kompositorischen Parameter gewinnen. Was für den Tonhöhenparameter gilt, trifft im Prinzip in der gleichen Weise auch für alle anderen Parameter zu.

Die bisher erwähnten Verfahren bestehen darin, gesprochene Sprache zu analysieren, zu manipulieren und eventuell in der von der kompositorischen Idee verlangten Weise zu transformieren. Die moderne Sprachsynthese unter Verwendung des Computers müßte nun eigentlich ein ideales Mittel zur Verfügung stellen, um die phonetischen Parameter, die - wie wir bereits festgestellt haben - grundsätzlich sowohl in Sprache als auch in Musik Geltung haben, in einer beliebigen Weise zu strukturieren, so daß zwischen den Extremen eindeutig phonetischer Sprachsignale und eindeutig musikalischer Klangproduktionen ohne jeglichen Sprachbezug beliebig vermittelt werden könnte. Man muß leider feststellen, daß die technologische Entwicklung im Bereich der Sprachsynthese so ist, daß aufgrund der Komplexität der Aufgaben und Ziele bisher die Verständlichkeit und die Natürlichkeit der gesprochenen synthetischen Sprache im Vordergrund der Entwicklung standen. Die zu erwartende technologische Weiterentwicklung könnte diesen Sachverhalt ändern. Dabei bietet sich allerdings von den verfügbaren Verfahren zur Sprachsynthese für die genannten kompositorischen Zwecke eigentlich nur eines an, nämlich das der Simulation des Sprachproduktionsvorgangs durch ein computerimplementiertes Phonations- und Artikulationsmodell (Kölnener Artikulationsmodell, siehe z. B. Kröger 1998). Der Vorteil solcher Synthesysteme liegt darin, daß die phonetischen Parameter in ihrer anschaulichen und durch ihre organische Produktion bestimmten Art und Weise kontrollierbar sind. Es werden also nicht beispielsweise Formantfrequenzen von Vokalen kontrolliert, sondern die Artikulationsbewegungen der Zunge, des Unterkiefers und der Lippen. Dadurch wird es möglich, die Artikulations- und Phonationsorgane als von dem speziellen Zwecke der Sprachproduktion befreite Organe und quasi als Musikinstrumente zu benutzen. Durch eine weiterentwickelte Sprachsynthese dieser Art wäre der Komponist in der Lage, beliebig zwischen Sprachsignalen und nicht sprachlich verwendeten phonetischen Parametern zu wechseln. Die Komposition phonetischer Parameter mit Hilfe eines solchen computerimplementierten Synthesystems könnte durchaus im Rahmen eines musikalischen Kompositionsprogramms erfolgen. Da solche Computerartikulationsmodelle in der Regel auch eine visuelle Darstellung der modellhaften Artikulations- und Phonationsbewegungen ermöglichen, wäre der übliche Nachteil elektronischer Musikproduktion, nämlich die Unanschaulichkeit, bis zu einem gewissen Grade vermieden. Dem Komponisten bliebe es, wie im Falle der elektronischen Musik bis jetzt auch, natürlich nach wie vor überlassen, bei der Aufführung eines entsprechenden Sprachmusikwerks rein synthetisch erzeugte phonetische Signale zusammen mit agierenden menschlichen Interpreten zu verwenden.

#### **Literatur**

Heike, G. (1969): Sprachliche Kommunikation und linguistische Analyse. Heidelberg.  
Kröger, B.J. (im Druck): Ein phonetisches Modell der Sprachproduktion. Tübingen.

**Prof. Dr. Georg Heike**  
Im Bendgesgarten 26  
D-53902 Bad Münstereifel

Institut für Phonetik  
Greinstraße 2  
D-50939 Köln