

Die Bedeutung der Phonetik in der Vokalkomposition von Stockhausen

Georg Heike, Köln

Ein Beitrag zum Thema „Bedeutung der Phonetik in der Vokalkomposition von Stockhausen“ muß meiner Meinung nach zunächst die Terminologie klären. Es mag vielleicht spitzfindig und lästig erscheinen, Begriffe, unter denen sich wohl jeder etwas vorstellen kann, noch einmal einer Klärung zuzuführen. Gestatten Sie mir aber dennoch, dies zu tun, da meiner Meinung nach die Behandlung des gestellten Themas auch danach geraten kann, wie klar oder unklar und wie genau festgelegt die Begriffe sind, die man benutzt.

Ad 1 Vokalkomposition:

Meine Anmerkung zu dem Begriff, der wohl von dem traditionellen der Vokalmusik ausgeht, soll nur darauf verweisen, daß das, was man in Musiklexika findet, genaugenommen entweder teilweise falsch oder zumindest unklar ist. So findet man beispielsweise unter „Vokal“ in der Erläuterung zur vokalen Musik, daß es sich um stimmhafte Laute handelt, mit denen also offenbar Vokalmusik gemacht wird, wie der fachliche Laie annehmen müßte. Weder sind Vokale durch ihre ausschließliche Stimmhaftigkeit definiert – es gibt, wie jeder weiß, auch die Möglichkeit, Vokale zu flüstern – noch verzichtet der Komponist von Vokalmusik auf die Konsonanten eines Textes. Im Gegenteil sind es gerade in der neuen Musik die Konsonanten, die das besondere Interesse der Komponisten finden. Anstelle von Vokalmusik oder Vokalkomposition wäre es sicherlich – und das geschieht gelegentlich ja auch – adäquater, von Sprachlautkomposition oder von Sprachkomposition zu reden. Gemeint ist auf jeden Fall das Komponieren mit phonetischem – nicht unbedingt mit sprachlichem – Material. Welche Konsequenzen letztlich daraus gezogen werden können, möchte ich zum Schluß meines Beitrags behandeln.

Ad 2: Phonetik

Unter Phonetik kann sich offensichtlich auch jeder etwas vorstellen. Die Phonetik ist eine sehr traditionsreiche Wissenschaft, ihre Methoden und ihre Inhalte haben sich aber nach dem Kriege grundsätzlich geändert. Diese Veränderungen haben insbesondere in der Bonner Phonetik durch Werner Meyer-Eppler stattgefunden, und dies ist auch der Grund dafür, warum Stockhausen die Phonetik und übrigens auch die Kommunikationsforschung für sich und damit auch für die Neue Musik entdeckt hat. Meyer-Eppler war ab Mitte der 50er Jahre durch einige Vorträge während der Darmstädter Ferienkurse für Neue Musik bekannt geworden als Phonetiker, aber vor allem auch als Kommunikationswissenschaftler, der sich nicht nur mit Sprache, sondern auch mit ihrem Bezug zur Neuen Musik beschäftigte. Stockhausen hat nicht Phonetik an sich studiert, sondern so, wie sie eben durch Meyer-Eppler im Zusammenhang mit der Informationstheorie und Kommunikationsforschung betrieben und in einer besonderen Weise vermittelt wurde. Vorlesungen von Meyer-Eppler, die ich zusammen mit Stockhausen eine Zeitlang besuchte, gaben ganz deutliche Hinweise darauf, daß Meyer-Eppler während der Vorlesungen, bei denen er immer frei sprach, in seinen Beispielen und in der Darstellungsweise auf ästhetische Belange vor allem in der Musik einging, sobald Stockhausen unter den Zuhörern war. Ebenfalls war zu bemerken, daß Stockhausen von dem Dargebotenen in der Regel sehr zu eigenen Gedanken angeregt wurde und sofort Umsetzungsmöglichkeiten im ästhetischen Zusammenhang gesehen hat. Laut Stockhausen war Meyer-Eppler der beste Lehrer, den er im Leben getroffen hat.¹

¹ Zitiert nach Werner Klüppelholz, „Sprache als Musik. Studien zur Vokalkomposition bei Karlheinz Stockhausen, Hans G. Helms, Mauricio Kagel, Dieter Schnebel und György Ligeti.“. Pfau Verlag, Saarbrücken. 2. Aufl. 1995. S. 40.

Nun zur näheren Definition der Phonetik, so wie sie in Bonn erstmals betrieben wurde und dann auch Vorbild für eine moderne Phonetik im deutschsprachigen Raum wurde.² Bezeichnend für die Phonetik, wie sie Meyer-Eppler vermittelte, war die enorme Erweiterung der alten Phonetik als Wissenschaft von der Beschreibung und Vermittlung der Aussprache, und zwar durch das Einbeziehen der Kommunikationsforschung und der Informationstheorie, aber auch im Zusammenhang damit der Psychologie der Wahrnehmung zu einer Phonetik als Wissenschaft von der lautsprachlichen Kommunikation, die sich mit allen Bereichen der Kommunikationskette zu befassen hat. Dabei sind vor allem zwei Gesichtspunkte zu erwähnen, die die Bonner Phonetik charakterisierten und die einen Einfluß auf Stockhausen gehabt haben:

1. Die Berücksichtigung der außerlinguistischen Bereiche bei der Produktion und Funktion phonetischer Signale, als da sind expressive, diagnostische Information und die Phänomene bei gestörter Sprache;
2. Die besondere Berücksichtigung der Wahrnehmungsseite der Kommunikationskette, d. h. die Bewertung des akustischen Signals bei der menschlichen Perzeption als Voraussetzung für seine kommunikative Funktionsfähigkeit.

In apparativ-technischer Hinsicht war in Deutschland im und kurz nach dem Krieg ein großer Nachholbedarf entstanden. In dieser Zeit konnte man nur in Bonn moderne Signalanalyse und auch bereits Signalsynthese betreiben und studieren, und dies, vor allem durch die Zusammenarbeit und die Beziehungen von Meyer-Eppler zum Rundfunk in Köln mit sehr guter Technologie. Hier wird deutlich, warum und in welchem Ausmaß die Phonetik für Stockhausen so bestimmend wurde: Meyer-Eppler war auch Initiator für die Entstehung der elektronischen Musik in Köln und damit in der Welt überhaupt. Man kann zusammenfassend feststellen, daß die Bedeutung der Phonetik, Kommunikationsforschung und Informationstheorie, wie Stockhausen sie in Bonn kennengelernt hat, weit über die Vokalkomposition Stockhausens hinausgeht und sich in vielfältigen Bezügen auch in seinen Publikationen nachweisen läßt. Obwohl ich an einigen wenigen Beispielen diesen Nachweis bringen werde, ist eine systematische Darstellung sämtlicher Einflüsse auf Stockhausens Werk in diesem Zusammenhang nicht möglich. Es ist sicherlich auch viel sinnvoller, das Prinzipielle herauszuarbeiten und eventuell eine Systematik der möglichen und tatsächlich erfolgten Umsetzungen und sogar Transformationen von Wissenschaft in musikalische Produktionen und Konzepte zu versuchen.

Im Falle der Vokalkomposition, um bei diesem Terminus zu bleiben, bietet es sich an, die drei konstituierenden Stadien der lautsprachlichen Kommunikationskette (Sender – Signal – Empfänger) zu benutzen, um die wichtigsten Umsetzungen beispielhaft zu beschreiben.

Ad 1:

Die wissenschaftliche Beschreibung und Erklärung der Lautproduktionsprozesse bei der Artikulation gibt Möglichkeiten an die Hand, die weit über den Buchstabentext hinausgehen, auf den sich traditionelle Vertonung beschränkt. Die eigentliche Lautung zu definieren und damit auch Texte differenzierter zu verarbeiten ist mit den Mitteln der phonetischen Notation möglich. Der Komponist kann damit dem musikalischen Interpreten vermitteln, was dieser über das, was er in der Musikhochschule gelernt hat, hinaus noch tun kann. Interessanterweise, nebenbeibemerkt, wird meines Wissens Phonetik in diesem Zusammenhang und in dieser Anwendung an keiner Musikhochschule gelehrt, nicht einmal in zusätzlichen Kursen. Ein von mir vor Jahren auf Empfehlung und Rat von Stockhausen gemachter

² Schüler im engeren und weiteren Sinne wurden Lehrstuhlinhaber in Köln, München, Kiel, Frankfurt.

daß musikalisch-akustische Parameter ein Äquivalent in einer Erlebniskategorie beim Hörer haben sollten. Was die Verarbeitung von Sprache im *Gesang der Jünglinge* betrifft, so wäre in diesem Zusammenhang vor allem zu nennen die beabsichtigte und durchgeführte Komposition der Verständlichkeit als Parameter. Aus der Informationstheorie ist ein sehr wichtiges Begriffspaar zu erwähnen, das Stockhausen nicht nur in der Verarbeitung von Sprache bestimmt hat, sondern auch ganz allgemein in seiner Komposition. Es sind die Begriffe der Redundanz und der Entropie und ihre Bedeutung in einer ästhetischen Kommunikationsforschung. Ähnlich wie in der sprachlichen Kommunikation kann man auch in der ästhetischen Kommunikation eine Optimierung annehmen, die sich irgendwo zwischen maximaler Redundanz und maximaler Entropie befindet. Die Übertragung wissenschaftlicher Erkenntnisse und Untersuchungsmethoden aus dem Bereich der sprachlichen auf die musikalische Kommunikation ist ein Hauptanliegen der Bonner Schule Meyer-Epplers und hat neue und andere Aspekte der Behandlung beider Bereiche im Gegensatz zum traditionellen Verhältnis von Sprache und Musik gebracht.

Über die Werke Stockhausens, die als Exempel für phonetische Textbehandlungen dienen, wie vor allem der *Gesang der Jünglinge*, *Stimmungen*, *Momente*, um nur die wohl wichtigsten zu nennen, möchte ich hier nicht ausführlich handeln. Das würde den Rahmen sprengen und ist ja auch andererseits und auch meinerseits schon mehrfach gemacht worden. Versuchen möchte ich vielmehr, das Typische und auch Neuartige herauszustellen und zu verallgemeinern. Dies kann zu einer vorläufigen Theorie der Sprachkomposition bzw. besser sogar zu einer Theorie einer allgemeinen ästhetischen Klangerzeugung führen, die im Hinblick auf Sprache nicht spezifiziert ist. Eine solche Theorie müßte den Gegensatz Sprachsignal↔Nichtsprachsignal als Parameter enthalten.

Ein Studium der Phonetik garantiert noch keineswegs einen sinnvollen und konsistenten, d. h. den kompositorischen Möglichkeiten adäquaten Gebrauch des phonetischen Wissens, wie viele Beispiele zeigen. Charakteristisch für Stockhausen war, daß er das Konzept der seriellen Kompositionsweise konsequent und stimmig auf die Behandlung des phonetischen Materials übertrug. Am deutlichsten läßt sich am Beispiel vom *Gesang der Jünglinge* aufzeigen, daß ein adäquates Komponieren mit phonetischem Material viel mehr bedeutet als Effekthascherei mittels bis dato ungewöhnlicher phonetischer Anweisungen.

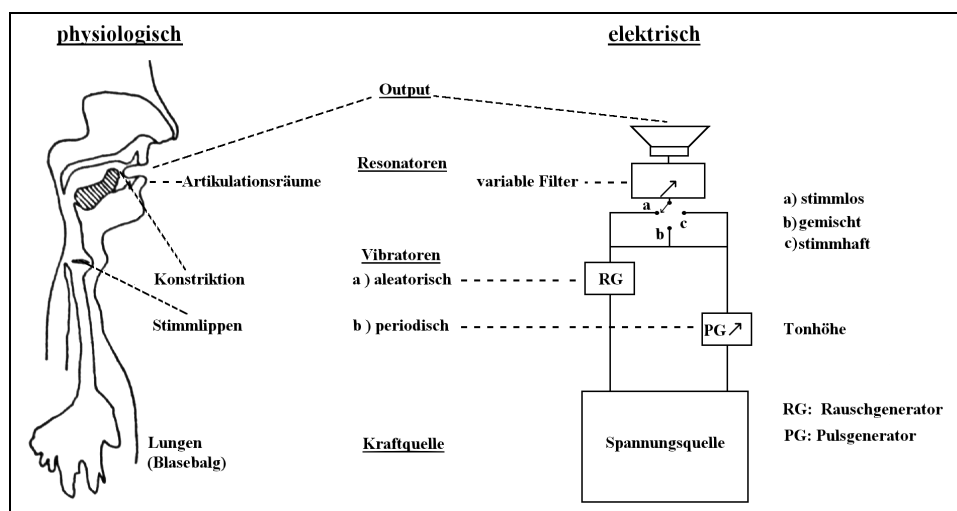


Abb. 2: Entsprechungen physiologischer und elektrischer Klangerzeugung

Der physikalische Begriff des Parameters hat vermutlich seit Meyer-Eppler bei Komponisten das Bewußtsein darüber geschärft, womit sie eigentlich komponieren. Nachdem klar geworden war, was traditionellerweise an Parametern schon immer komponiert wurde (Tonhöhe, Lautstärke, Klangfarbe, Tempo), stellte sich nun in jeder neuen Komposition mit neuem Klangmaterial die Frage nach den zu komponierenden Parametern, und zwar nicht nur auf der Materialebene, sondern auch auf den höheren Ebenen des Umgangs mit dem Material, also der Form und vor allem schließlich der Perzeption. Abbildung 2 veranschaulicht die phonetischen Parameter, die sich auf den Erzeugungsvorgang beziehen (z. B. Stimmhaftigkeit, Geräuschhaftigkeit, Abruptheit, diverse Vokalparameter etc.). Parameter höherer Ordnung ergaben sich konsequent aus dem musikalisch-kompositorischen Denken: Vokal/Konsonant als kontinuierlicher Parameter, die Verständlichkeit bei Verwendung von Texten, der Gegensatz Sprachsignal / elektronisch bzw. instrumental erzeugtes Signal als Pole eines Kontinuums etc. Die folgenden Beispiele zeigen, mit welchen Mitteln die beiden wichtigsten Parameter, nämlich der Parameter elektronischer Klang/Sprachklang und der Parameter Verständlichkeit im *Gesang der Jünglinge* komponiert wurden.

- Beispiel 1: Verschiedene Versionen einer Silbe bzw. eines Wortes werden miteinander überlagert, zum Teil verhallt. Im Extremfall resultieren Rauschblöcke chorischen Charakters, jedoch ohne sprachliche Erkennbarkeit: Ab 5'33⁴ das Wort „Feuer“ (noch erkennbar) und ab 4'42 eine sprachlich nicht identifizierbare Überlagerung.
- Beispiel 2: Einzellaute werden kontrapunktisch mit chorisch-akkordischen und syllabischen Partien überlagert. In 2'52 wird das Wort „alle“ mehrfach mit sich selbst überlagert, daraus bleibt das l als Dauerlaut stehen und wird gleichzeitig mit elektronischen Klangmustern und dem Einzelwort „preist“ überlagert.
- Beispiel 3: Durch Manipulation des Bandlaufs werden Transformationen in andere Tonlagen und auch die Umkehrung verwendet. Bei ca. 9'20 Transformation in sehr hohe Lage mit Beschleunigung; bei 5'46 das Wort „Sommersglut“ durch Umkehrung der Wortteile überlagert und kontrapunktiert, gefolgt von der Inversion des Textes „preiset den Herrn“;
- Beispiel 4: Übergänge und Mischungen von gesprochenen Lauten und elektronischen Klängen: Etwa 6'12: aus dem gesprochenen Wort „Eis“ bleibt der letzte Konsonant, das s, als Dauerlaut stehen und transformiert sich in ein elektronisch erzeugtes Rauschsignal.
- Beispiel 5: Komposition der Sprachverständlichkeit als Parameter. Stockhausen definiert 7 Verständlichkeitsgrade. Abbildung 3 zeigt sie in der seriellen Anordnung 5-1-2-6-7-3-4, die quasi als Exposition am Anfang des Werkes steht (20,5“ bis nach 52,3“).

1: nicht verständlich	5: nicht genau verständlich
2: kaum verständlich	6: verständlicher
3: ganz wenig verständlich	7: verständlich
4: fast verständlich	

⁴ Die Zeitangaben beziehen sich auf die CD 3 (Stockhausenverlag) und gelten einschließlich 10“Vorlauf.

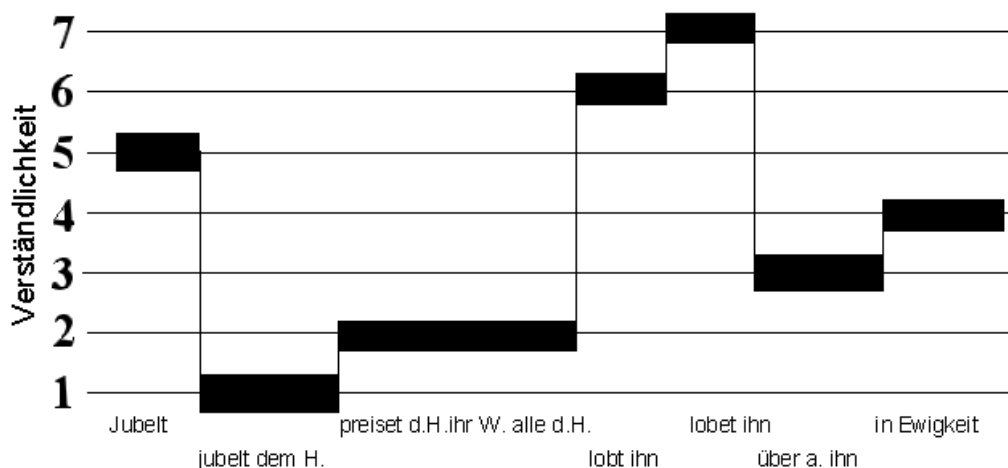


Abb. 3: Die 7 komponierten Verständlichkeitsgrade im *Gesang der Jünglinge*, aus dem Anfang (10,5“ bis 42,3).

Erwähnt werden muß (siehe Text zur CD 3), daß Stockhausen offenbar eine Korrelation von Verständlichkeit und Klangform auf der Skala Sprachklang / elektronischer Klang annimmt: verständliche Passagen sind auch erkennbar gesungen, unverständliche sind nahezu vollständig elektronisch erzeugt. Zitat (s. Beiheft, S. 60): „... der Begriff Verständlichkeit ist also semantisch und klanglich-phonetisch differenziert.“

Bei der Parameterisierung phonetischer Merkmale für kompositorische Zwecke wird es interessanterweise auch oder gerade dem linguistisch orientierten Phonetiker bewußt, daß die für die Beschreibung linguistischer Kategorien benutzten polaren bzw. binären Merkmale wie z. B. stimmhaft/stimmlos, abrupt/dauernd, hell/dunkel etc. eigentlich phonologische Konstrukte sind. Auf der Ebene der phonetischen Substanz liegt nahezu immer Kontinuirlichkeit vor. Phonetische Parameter weisen jedoch auch dann im Vergleich zu instrumentalen bzw. elektronisch steuerbaren Signalparametern einige Besonderheiten auf:

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 Punkt 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: sog. '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.

Das letzte besonders typische Merkmal eines phonetischen Parameters ist übrigens erst jüngst in der phonetischen Theorie in seiner konsequenten Bedeutung erkannt worden und hat zur Entwicklung des sog. gesturalen Ansatzes sowohl in der Deskription als auch in der Modellierung geführt. *Gesten* stehen in einem unmittelbaren und auch beim Perzipieren in einem ursächlichen und nachvollziehbaren und somit auch rekonstruierbaren Zusammenhang mit den verursachenden Organbewegungen. Eine zuverlässige und ungefährliche Registrierung von Artikulationsbewegungen ist übrigens erst seit ca. 10 Jahren mit der Entwicklung des Elektromagnetischen Articulographen möglich. Eine Umsetzung in musikalische Parameterstrukturierungen könnte interessante Erweiterungen im Bereich der Steuerung von Zeitabläufen ergeben.

Die Verwendung von Bewegungsgesten für die elektronische Klangerzeugung in Sprache und Musik wäre im Rahmen eines computerimplementierten Produktionsmodells nach dem Quelle-Filter – Prinzip am sinnvollsten. Werden Quelle und Filter wie bei der Sprachproduktion durch ein Phonationsmodell und einen Artikulationstrakt spezifiziert, dann wäre das Ergebnis gesprochene Sprache. Der Vorteil eines solchen Synthesystems, wie im Fall des Kölner Artikulationsmodells bereits weitgehend entwickelt, 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, die beispielsweise auch darin bestehen würde, daß anstelle eines Phonationsmodells eine beliebige Quelle simuliert werden könnte, wäre der Komponist in der Lage, zwischen Sprachsignalen und nichtsprachlich 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 solchen von agierenden menschlichen Interpreten zu verwenden.

Literatur:

- Heike, Georg: „Ästhetische Phonetik als Wissenschaft von der ästhetischen Kommunikation in Sprache und Musik“, in: Angela Biege und Ines Bode (Hrsg.): Theorie und Empirie in der Sprechwissenschaft. Festschrift zum 65. Geburtstag von Eberhard Stock am 17. Juni 1998. Verlag Werner Dausien, Hanau und Halle/S. 1998. S. 65-74.

Heike, Georg: „Musiksprache und Sprachmusik. Anmerkungen zu einer Theorie der ästhetischen Kommunikation“, in: Georg Heike: Musik und Sprache – Gesammelte Aufsätze. Pfau Verlag, Stuttgart (im Druck)

Kröger, Bernd J.: „Ein Phonetisches Modell der Sprachproduktion“, Niemeyer, Tübingen 1998.

Klüppelholz, Werner: „Sprache als Musik. Studien zur Vokalkomposition bei Karlheinz Stockhausen, Hans G. Helms, Mauricio Kagel, Dieter Schnebel und György Ligeti.“. Pfau Verlag, Saarbrücken. 2. Aufl. 1995. S. 40.

Schwerdtfeger, Dettloff: „Karlheinz Stockhausens Oper *Donnerstag aus Licht*. Ziel und Anfang einer kompositorischen Entwicklung“. Magisterarbeit, Köln, WS 97/98.

Prof. Dr. Georg Heike
Institut für Phonetik
Greinstraße 2
D-50939 Köln