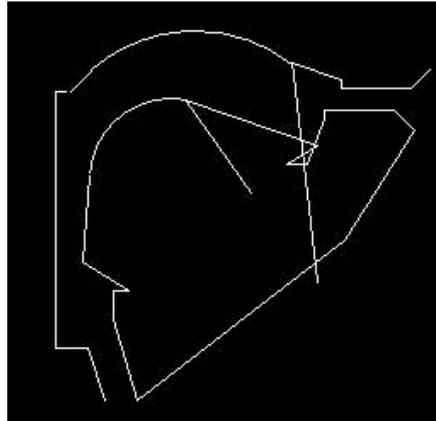


Roman Pfeifer



Studie zu einem Lautgedicht

für 4-Kanal Tonband

2005/06

9'

Studie zu einem Lautgedicht

Studie zu einem Lautgedicht ist eine Musik für 4-Kanal-Tonband, entstanden in den Studios des ICEM der Folkwang-Hochschule Essen in den Jahren 2005/2006. Diese Partitur versucht die konzeptuellen Hintergründe dieses Werks zu beleuchten und die verschiedenen Bau- und Organisationsweisen der Komposition verständlich zu machen.

Der Hörer als Sprechautomat:

Ausgangspunkt für das Tonbandstück *Studie zu einem Lautgedicht* ist ein Modell der Sprachwahrnehmung, nach dem die Lautwahrnehmung nur mit Bezug auf die Lautproduktion möglich ist. Das heißt, dass der Hörer Laute als Bewegungsabläufe von Artikulationsorganen wahrnimmt, indem er anhand der klingenden Laute die Bewegungen des Sprechapparates stumm mitvollzieht. Diese Theorie der Sprachwahrnehmung wurde am Bell Institut von einer Forschungsgruppe um den Phonetiker Alvin M. Libermann seit 1965 entwickelt. Sie ist unter dem Namen Motortheory of Speech bekannt geworden.

„The objects of speech perception are the intended phonetic gestures of the speaker, represented in the brain as invariant motor commands that call for movements of the articulators through certain linguistically significant configurations“
(Liberman, Alvin M., *Speech: A Special Code*, London 1996, S. 237.)

Der Hörer fühlt bei sich selbst, wie sich die Artikulatoren des Sprechers bewegen. Die Theorie verschiebt das Primat der Wahrnehmung von Klanggestalten auf Bewegungsgestalten und bietet eine aktive Hörtheorie, bei der der Hörer körperlich involviert ist, um Sprache zu dekodieren. Diese Erkenntnis ist Ausgangspunkt der Komposition *Studie zu einem Lautgedicht*, und zwar sowohl für die Organisationen der formalen Strukturen als auch für die Klangsynthese.

Das Klangmaterial von *Studie zu einem Lautgedicht* wurde durch ein Computermodell des Sprechapparats im Programm Praat generiert, in welchem die Sprechbewegungen durch Aktionspotentiale von 29 „Muskeln“ beschrieben werden. Diese Laute wurden mit elementaren elektronischen Klängen, also dem Knackser, dem Sinuston und dem Rauschen, kombiniert.

Ziel war nicht, den Computer zum Sprechen zu bringen und die verschiedenen menschlichen Phoneme nachzuahmen; vielmehr werden die verschiedenen Arten des Artikulierens – also der Konsonant als Widerstand, Reibung, Enge, Vibration, Schlag, Verschluss und Sprengung und der

Vokal als Überlagerung von Verformung verschiedener Resonanzräume – verwendet, um charakteristische Bewegungsformen für die einzelnen Organe zu erfinden, die dann durch Kombination zu komplexen Mundraumchoreografien führen.

Die Wirkung dieser Klänge schwankt zwischen starrem Automatismus und dem Ausdruck emotionaler Körperlichkeit. Im Versuch, dem menschlichen Laut zu gleichen, offenbart das Computermodell seine Künstlichkeit. Andererseits hat es die Möglichkeit, Laute so nachzuahmen, dass der Hörer die vermeintliche Körperlichkeit und Emotionalität des Klangs wahrnimmt. Das Hörerlebnis changiert auf diese Weise zwischen distanzierter Wahrnehmung einer künstlichen Klangwelt und direkter Bewegtheit durch die Klänge im stummen Mitartikulieren. Indem der Hörer die computererzeugten Laute innerlich nachahmt, werden seine Atem-, Stimm-, Rachen-, Zungen- und Lippenbewegungen umgeformt und verwandelt. Der Hörer erfährt sich als Sprechautomat; er wird gleichsam durch den Computer artikuliert.

Im Folgenden sollen die einzelnen Schichten der Komposition betrachtet werden.

Artikulationssynthese

zur Erzeugung künstlicher Laute ist zentraler Bestandteil der Komposition; hierfür wurde ein Modell des Phonetikprogramms Praat verwendet. In diesem Modell des Sprachlauts werden Aktionspotentiale von 28 „Muskeln“ beschrieben, um mittels eines physikalischen Modells Sprachlaute zu synthetisieren.

Die einzelnen Kurven dieser Aktionspotentiale wurden durch die in Praat integrierte Skriptsprache beschrieben (siehe Anhang).

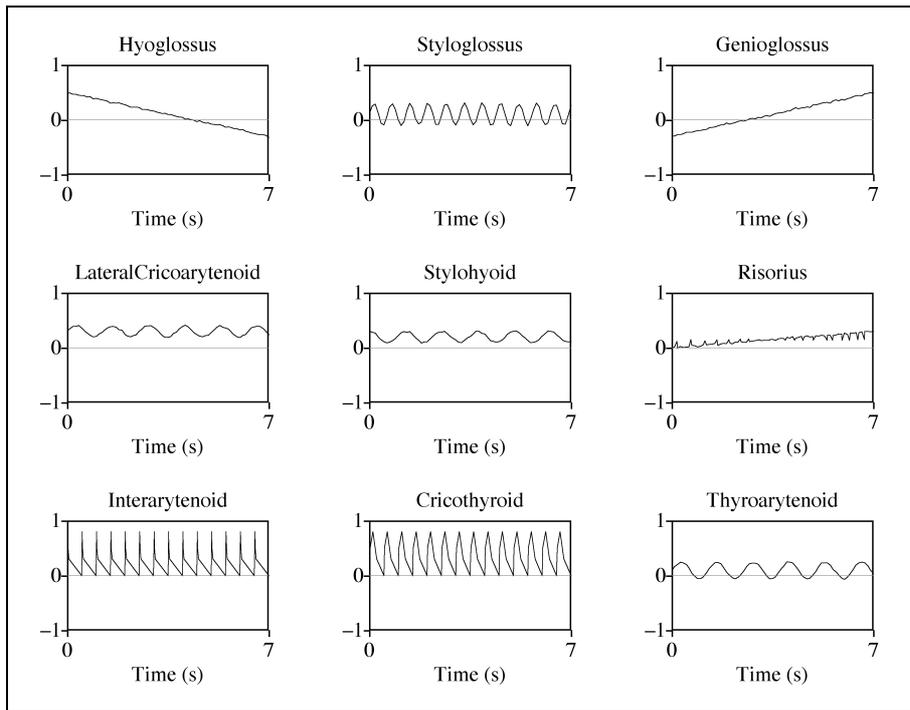
Diese elementaren Kurventypen sind:

- a) Statische Werte
- b) Lineare Interpolationen
- c) Sinus- und Cosinuswellen.

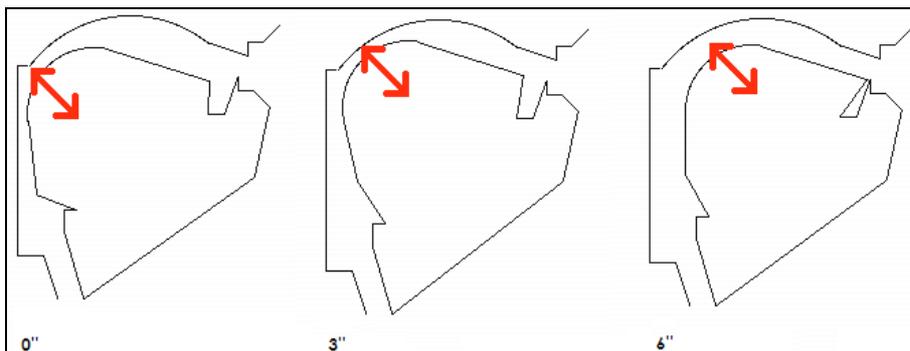
Diese drei Funktionen können durch Zufallswerte verzerrt werden. Weiterhin gibt es:

- d) Muster mit 4 Stützwerten, die periodisch wiederholt werden
- e) Zufällig verteilte Akzente.

Hiermit wurden die Parameterbewegungen für die einzelnen Muskeln erzeugt. Im folgenden Beispiel sind die Parameter in der oberen Zeile für die Bewegungen der Zunge verantwortlich, die mittlere Zeile steht für die Bewegungen des Rachens und der Lippen und die letzte Zeile beschreibt die Bewegungen der Stimmbänder.

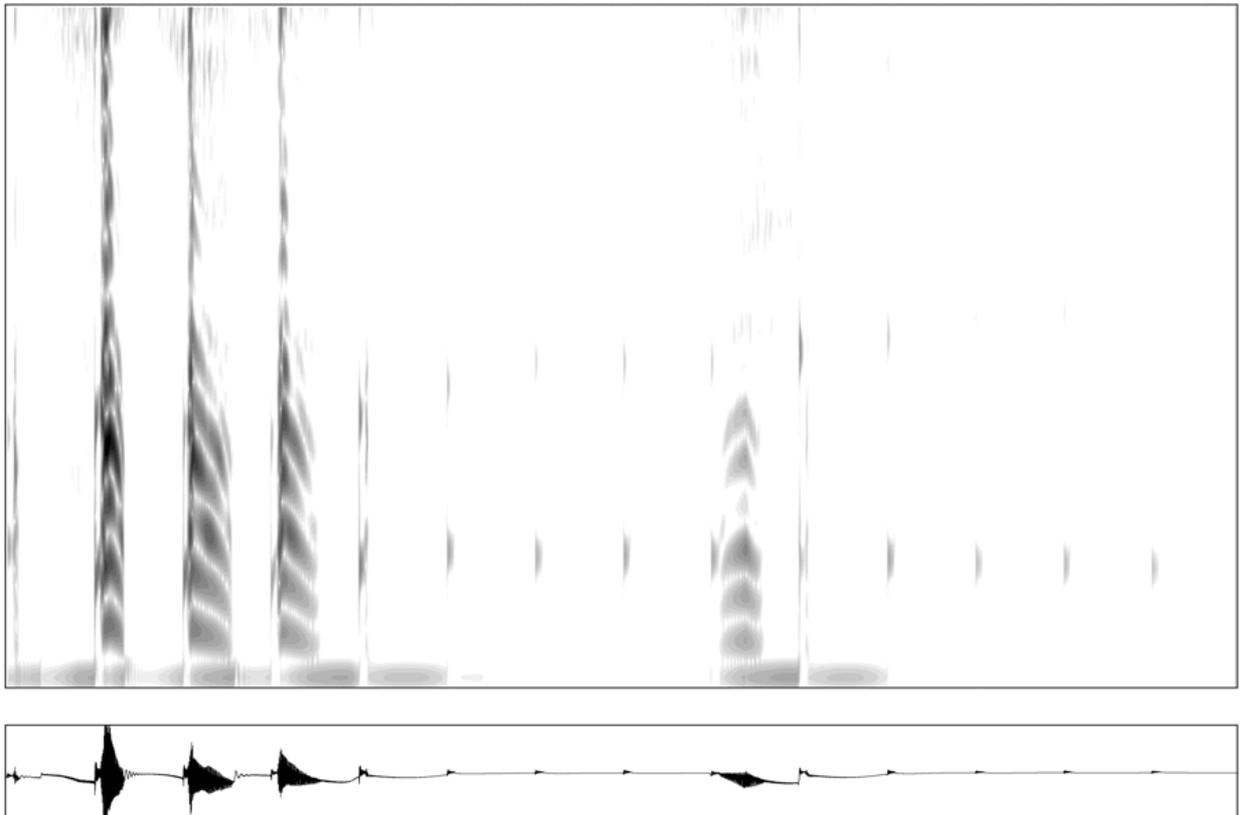


Anhand der folgenden Abbildung kann man erkennen, wie sich der Mundraum durch diese Funktionen über die Zeit von sieben Sekunden verändert. Diese Veränderung wird durch eine periodische Bewegung der Zunge vor und zurück überformt, die durch die Pfeile angedeutet wurde, und durch eine regelmäßige Folge von Impulsen der Stimmbänder artikuliert.



Das Ergebnis sind nicht klare Phoneme sondern Artikulationsbewegungen, in denen die einzelnen Artikulatoren ihren spezifischen Geschwindigkeiten folgen. Es ergeben sich Folgen, in denen

Varianten eines Bewegungssets zum Klingen gebracht werden. Im Spektrogramm dieser Artikulation sind diese verschiedenen Aspekte in Hüllkurven und spektralen Veränderungen erkennbar. Es entstehen Gruppen von ähnlichen Klängen, die in der Komposition in verschiedenen Schichten überlagert wurden. Die Architektur des Klangraums wurde demnach so gestaltet, dass diese Schichten sich gut voneinander abheben. Jeder Klang hat seinen festgelegten Platz, punktuell wird Hall, durch Convolution mit verschiedenen Impulsantworten erzeugt, verwendet, und einzelne Schichten werden in unterschiedliche Hallräume, meist mit kleinen Volumen und kurzen Hallzeiten, projiziert.



Neben den künstlichen Lauten gibt es zwei zusätzliche Materialschichten. Sinustöne und Rauschbänder werden an die Laute angelagert und verlängern so die verschiedenen Tonhöhen und Spektren. Sie betonen die Künstlichkeit der synthetischen Laute. Eine zweite Materialschicht bilden Impulsketten, die in verschiedensten Ausprägungen, Dichten und Pulsationen die Form des Stückes maßgeblich gestalten.

Impulsketten

unterschiedlicher Länge sind durch einfache rhythmische Proportionen 4:5:6:7 geprägt. Sie bestehen zumeist aus kleinen formelhaften Mustern, die geringfügig variiert werden.

$\bullet = 128$ ($\bullet = 96, \bullet = 200$)

22

mp ~ mf *p* *mp ~ mf* *p*

22

mp ~ mf *p* *mp ~ mf* *p*

Diese Ausgangsmaterialien werden in verschiedenen Tempi, die ebenfalls in solchen elementaren Proportionen zu einander stehen, produziert. Dies macht es möglich, verschiedene Abfolgen dieser Ketten zu konstruieren, indem gleiche Tempi genutzt werden, um metrisch zu modulieren. In diesem Beispiel wird das Muster in den Tempi 128:96:160 (also 4:3:5) angelegt und bietet so eine Vielzahl von Möglichkeiten, über gemeinsame Geschwindigkeiten das Material in verschiedenen Kontexten zu platzieren. Die Tabelle gibt einen Überblick über die entstehenden Pulsgeschwindigkeiten:

Tempo	Rhythmische Proportionen			
1/4	1/20	1/16	1/12	1/10
128	640	512	384	320
96	480	384	288	240
160	800	640	480	400

Eine resultierende Kette, die aus diesen Formeln komponiert ist, sieht demnach folgendermaßen aus.

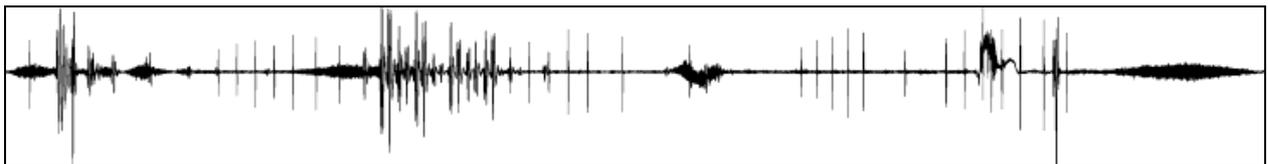
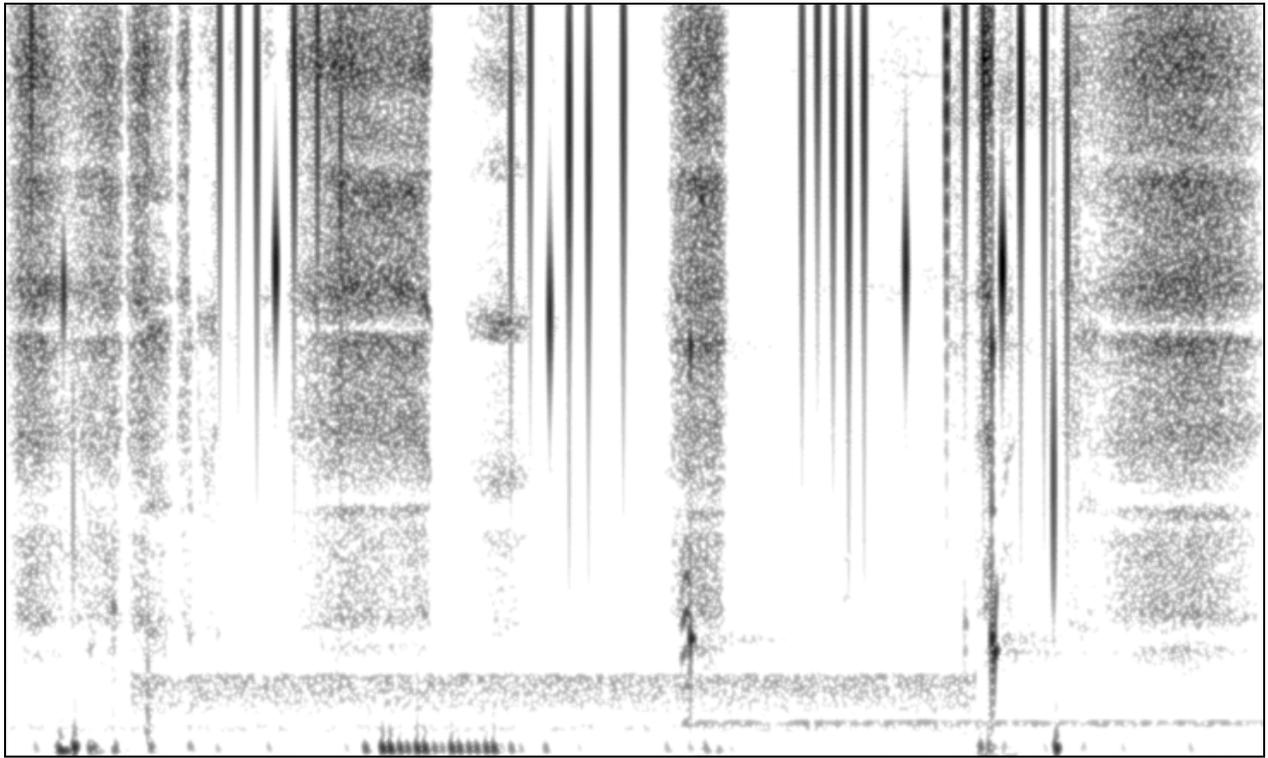
640 512 384 320, 640 512 384 640..
 384 288 240, 480
 480 400 800 640 480 400 800 640

Ein zweites Beispiel dieser Impulsketten ist melodischer aufgebaut. Es ist harmonisch aus einem Ausschnitt (49. - 60. Teilton) aus der Obertonreihe mit Grundton **D** aufgebaut. Das 11-tönige Muster lässt sich rhythmisch in 3:1:7 Töne gliedern, in Bezug auf die Bandbreite in 3:2:3:3 Töne. Es ist gegen Ende deutlich zu vernehmen.

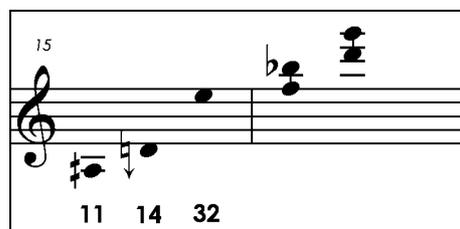
Das Muster wird in den Tempi 80 und 96 produziert und es entstehen erneut gemeinsame rhythmische Werte, die für verschiedene Verbindungen genutzt werden können

Rauschbänder und Sinustöne

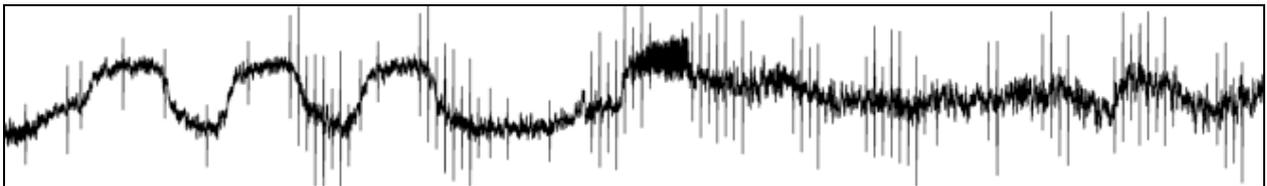
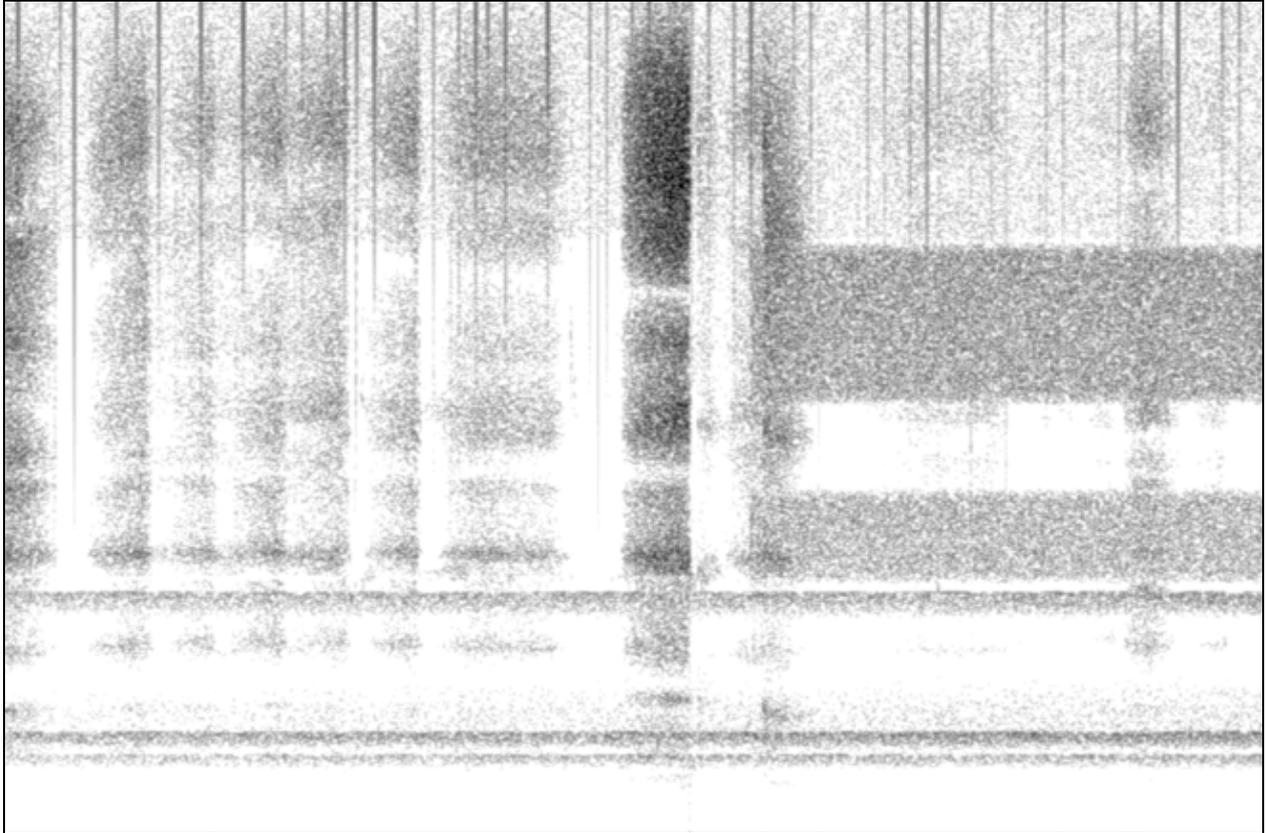
bilden die dritte Materialschicht in *Studie zu einem Lautgedicht*. Sie werden einerseits genutzt, um klingende Tonhöhen oder spektrale Maxima der synthetischen Laute zu verlängern. In der folgenden FFT-Darstellung kann man deutlich erkennen, wie im ersten Drittel ein Laut durch ein Rauschband verlängert wird. Etwas nach der Mitte dient dann ein Sinuston als Tonverlängerung, bevor im letzten Drittel der Einsatz eines synthetischen Plosivs dazu verwendet wird, das Rauschband wieder abzuschalten. Ebenso kann man die anderen Schichten, die breiten Frikativbänder der Artikulationssynthese-Laute und die Impulsketten, deutlich erkennen.



Formal lässt das Stück sich in verschiedene Zonen einteilen, in denen bestimmte Laute (Plosive, Frikative, Atem ...) vorherrschend sind. Diese Charakteristiken werden durch Impulse, Rauschen und Sinustöne ergänzt und verstärkt. Besonders am Ende des Stücks kann man hören, wie die künstlichen Atemgeräusche nach und nach in ein hohes Rauschen eingehüllt werden. Dieses besteht zuerst aus drei engbandigen Rauschen, zu denen dann zwei Bänder mit dem Ambitus einer Quart hinzukommen.



Auch hier lassen sich in der FFT die einzelnen Schichten, also Laute, sehr hohe Impulse und Rauschen visuell gut unterscheiden.



Studie zu einem Lautgedicht

hat diesen Namen, da es ein erster Versuch ist, die Verschränkung von Sprachklang und Artikulationsbewegung, wie sie die Motortheorie als Grundlage für Produktion und Perzeption erkennt, als Konstruktionsprinzip der Musik zu verwenden und hörbar zu machen. Der Sprachlaut und seine künstliche Erzeugung bergen so ein Modell, das es vermag, die Wahrnehmung des Hörers buchstäblich in Bewegung zu setzen. Der Erforschung dieser Mechanismen der menschlichen Sprache ist dieses Stück so schon ein bisschen näher gekommen.