

# Vokale in Klängen – Eine LPC-basierte Extraktion der Vokalqualität zur Abbildung von Violinklängen im Vokaldiagramm

Stella Müller<sup>1</sup>, Robert Mores<sup>2</sup>

<sup>1</sup>Hochschule für Angewandte Wissenschaften Hamburg, 20099 Hamburg, Deutschland, Email: stellamue@web.de

<sup>2</sup>Hochschule für Angewandte Wissenschaften Hamburg, 20099 Hamburg, Deutschland

## Einleitung

Sprache umgibt uns seit unserer Geburt und prägt unsere auditive Wahrnehmung grundlegend. Prägt Sprache auch unser musikalisches Klangempfinden? Weisen Musikinstrumente ähnliche Eigenschaften wie Vokale auf? Möglicherweise leistet eine extrahierbare Vokalqualität die Verbindung zwischen subjektiv wahrgenommener Klangqualität und technischen Gegebenheiten eines Instrumentenklangs.

In der vorliegenden Arbeit wird eine automatisierte Extraktion der Vokalqualität aus Instrumentenklingen vorgestellt. Der Ansatz wird exemplarisch anhand von Violinklängen reflektiert.

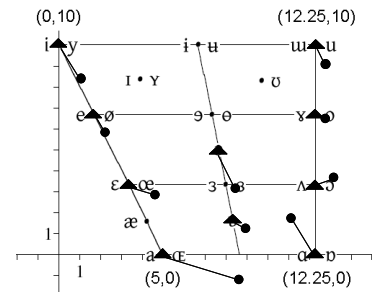
## Vokale

**Phonetische Einordnung:** Mit Hilfe des Vokaldiagramms nach Jones, welches von der International Phonetic Association (IPA) übernommen wurde, können Vokalqualitäten systematisiert, kontinuierlich und unabhängig von bestimmten Sprachen veranschaulicht werden [1]. Die Position eines Vokals im Diagramm korrespondiert mit der Zungenposition. Die Vokalqualität wird je nach Rundung der Lippen in primäre und sekundäre Kardinalvokale weiter differenziert und üblicherweise nach Gehör in das Vokaldiagramm eingeordnet [2].

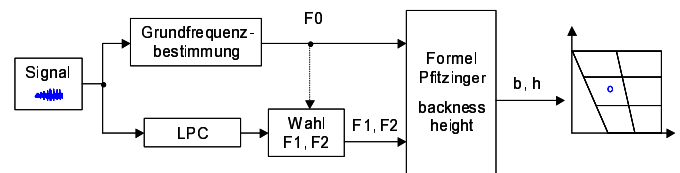
**Akustische Merkmale:** Durch Formung des Mundraums und Positionierung der Zunge entstehen für einen Vokal typische Resonanzen. Dabei charakterisieren im Wesentlichen zwei Formanten (im Folgenden F1 und F2) im Bereich zwischen ca. 300Hz und 3000Hz die Vokalqualität. Daher werden Vokale in phonetischen Untersuchungen häufig in der F1-F2-Ebene dargestellt, wobei die Formantfrequenzen ungefähr mit Zungenlage (F2) und -höhe (F1) korrelieren. Die genaue Vokalqualität ergibt sich nun aus der Position in der F1-F2-Ebene in Verbindung mit der noch zu ermittelnden Grundfrequenz F0 [3], [4].

## Methode

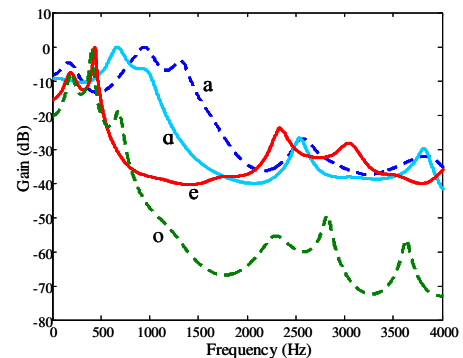
**Automatisierung:** Mittels LPC-Analyse wird die Formantstruktur berechnet, die beiden relevanten Formanten werden gewählt und die Grundfrequenz bestimmt. Die Übersetzung der akustischen Parameter in eine Vokalqualität ist anhand der Formeln von Pfitzinger möglich, welche F0, F1 und F2 auf die Ebene des Vokaldiagramms für primäre Kardinalvokale nach Jones abbilden. Vokalqualitäten werden als Punkte im Vokaldiagramm dargestellt, Abbildung 1. Es werden ausschließlich primäre Vokale berücksichtigt, da sonst Uneindeutigkeiten bezüglich der Zuordnung der Formanten auf eine Vokalqualität entstehen [4], [5].



**Abbildung 1:** Nach IPA definierte Zielpunkte der verifizierten Vokale (Dreieck) im Vergleich zu den automatisch berechneten Vokalqualitäten (Punkt), Angaben in cm.



**Abbildung 2:** Verarbeitungsschritte der Analyse von vokalhaltigen Klängen.



**Abbildung 3:** LPC-Spektren einiger Vokalqualitäten bei etwa gleichbleibender Grundfrequenz.

**Wahl der Formanten:** Für die qualifizierte Auswahl wird der vokalrelevante Frequenzbereich in drei Abschnitte unterteilt und anhand eines eigenen Ansatzes auf Energiedichteverhältnisse untersucht. Die Methode, Vokale allein durch F0, F1 und F2 zu beschreiben, entspricht nicht der gegebenen Komplexität der Vokale und muss erweitert werden: Zum einen weisen Vokale teilweise drei bis vier, hintere Vokale oft nur einen Formanten auf [6]. Zum anderen treten mit zunehmender Grundfrequenz deren Harmonische in den Vordergrund und die Formantstruktur geht verloren [3]. Während die Sprachtonhöhe zwischen 110Hz und 330Hz (m/w) variiert und der Effekt noch nicht markant stört, erreichen Violinen Tonhöhen über 1200Hz. D.h. im Extremfall liegen ein oder zwei Harmonische im

betrachteten Bereich. Hinzu kommt, dass bei hoher Grundfrequenz die Grundfrequenzresonanz in den Bereich von  $F_1$  kommt und nicht immer zweifelsfrei entschieden werden kann, ob  $F_0 = F_1$  oder nicht. Anhand eines systematischen Vergleichs vieler Vokalspektren unterschiedlichster Grundfrequenzen gelingt die Entwicklung der Regeln für eine qualifizierte Auswahl.

## Verifikation

Die hier entwickelte automatisierte Extraktion ist anhand von Vokalaufnahmen der IPA verifiziert. Für die Kardinalvokale [i], [e], [ɛ], [a], [ɑ], [ɔ], [o], [u] und die zentralen Laute [ə] und [ɐ] ergibt das Verfahren nur geringe Abweichungen, siehe die nach IPA definierten Zielpunkte und die automatisiert erzielten Positionen im Vokaldiagramm in Abbildung 1 [1]. Die Abweichung beim Vokal [a] liegt an überprüfbar Abweichungen der Klangfarbe von der erwarteten Vokalqualität in der Aufnahme. Für alle anderen Vokale treffen die berechneten Punkte in einen plausiblen Bereich.

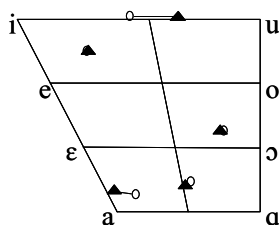
## Analyse der Violinenklänge

Enthalten nun vokalähnliche Klänge die entsprechenden Komponenten im Spektrum? Lassen sich aus Violinenklängen Vokalqualitäten extrahieren?

Ausgewählte Violinenaufnahmen werden angehört und die wahrgenommene Klangfarbe nachgesungen. Die Vokalqualitäten beider Signale werden berechnet und im Vokaldiagramm dargestellt, um zwischen Wahrnehmung und Berechnung zu vergleichen. Zur Kontrolle der Automatisierung werden die relevanten Formanten zusätzlich aus dem Spektrum abgelesen und mit den automatisch berechneten Ergebnissen verglichen.

## Ergebnisse

Wie Abbildung 3 zeigt, liegen die Vokalqualitäten der Violinenklänge und der gesungenen Vokale überwiegend nahe beieinander. Offensichtlich ist die Abbildung von Violinenklängen im Vokaldiagramm auf die beschriebene Art und Weise möglich und stimmt mit dem Höreindruck überein.



**Abbildung 4:** Vergleich der berechneten Vokalqualitäten von ausgewählten Violinenklängen (Dreieck) und deren nachgesungenen Vokale (Kreise).

Vokalqualitäten wie [i] und [e] sind meist nur in höheren Tonlagen zu finden. Das Spektrum dieser Tonlagen wird durch wenige Harmonische der Grundfrequenz beherrscht, sodass keine sinnvollen Formantfrequenzen bestimmt

werden können und die aus den Harmonischen zu berechnenden Vokalqualitäten nicht dem Höreindruck entsprechen.

Häufig sind in den Violinenklängen Anteile mehrerer Vokalqualitäten enthalten oder die Vokalqualität kann nicht eindeutig zugeordnet werden, da das mittelfrequenzreiche Resonanzprofil von Violinen meist mehr als die bei Vokalen üblichen bis zu 4 Formanten aufweist..

Prinzipiell lässt sich aber feststellen: Vokale, die in Klängen enthalten sind, spiegeln sich auf gleiche Weise in den Spektren wider wie mit Stimme erzeugte Vokale und können somit dargestellt werden. Dabei eignet sich die Violine aufgrund der sehr hohen Tonlage nur bedingt. Auf diese Weise ist eine Perspektive einer Klangbeschreibung reproduzierbar ausformuliert, die ansonsten vielleicht nur unbewusst in die Bewertung von Klangqualität mit eingeflossen ist.

## Ausblick

Durch die Vielzahl von Formanten der Violine entsteht häufig ein uneindeutiger Höreindruck. Um diesem gerecht zu werden und mögliche Vokalkomponenten herauszufiltern, bedarf es noch weiterer Untersuchungen.

Das Violoncello bietet sich als Alternative an, dessen typische Tonlage im Bereich der Sprechstimme liegt. Damit fällt nicht nur die Assoziation zur Stimme leichter, auch die Grundtöne liegen in einem Bereich, in welchem die Formantstruktur nicht von Harmonischen beeinträchtigt wird.

## Dank

Das BMBF fördert das Projekt „Wissenschaftliche Begleitung der klang-ästhetischen Weiterentwicklung der Violine“. Ergebnisse und Hörbeispiele unter [7].

## Literatur

- [1] homepage der Universität Victoria, URL: <http://web.uvic.ca/ling/resources/ipa/charts/IPAlab/IPAlab.htm> (Zugriff am 30.01.2008)
- [2] Wagner, K.H.: Phonetik und Phonologie. Universität Bremen, 63-70, URL: <http://www.fb10.uni-bremen.de/khwagner/phonetik/pdf/phonetik.pdf> (Zugriff am 30.01.2008)
- [3] Peterson, G. E., Barney, H. L.: Control Methods Used in a Study of the Vowels. In: The Journal of the Acoustical Society of America 24, Heft 2, 1952, 175-184.
- [4] Pfitzinger, H. R.: Acoustic correlates of the IPA vowel diagram. In: Proceedings of the XVth Int. Congress of Phonetic Sciences vol. 2, Barcelona, 2003, 1441-1444.
- [5] Pfitzinger, H. R.: Towards functional modelling of relationships between the acoustics and perception of vowels. In: ZAS Papers in Linguistics 40, 2005, 133-144.
- [6] Stumpf, Carl: Die Struktur der Vokale. In: Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin. Gesamtsitzung vom 4. April 1918, Berlin, 333-358.
- [7] URL: <http://www.mt.haw-hamburg.de/home/mores/> -> Veröffentlichungen