

# **Von Kolonialpolitik zu rechnergestützter Musikforschung Die Wilhelm-Heinitz-Sammlung außereuropäischer Musik an der Universität Hamburg: Ein historischer und methodologischer Überblick**

## **Einleitung**

Erst durch die technische Errungenschaft der Schallaufzeichnung wurde es möglich, nicht-materielle Kulturleistungen wie insbesondere Musik, gesprochenes Wort und somit die oral tradierten Wissensbestände schriftloser Gesellschaften zu bewahren und in Form von Schallarchiven der wissenschaftlichen Untersuchung zugänglich zu machen. Schallaufnahmen gehören naturgemäß zu den wichtigsten Quellen der Systematischen, Vergleichenden und Historischen Musikwissenschaft, der Musikethnologie sowie der Ethnologie, Volkskunde, Phonetik und Linguistik.

Das systematische Sammeln ethnographischer Tonträger an der Universität Hamburg lässt sich bis ins Jahr 1910 zurückverfolgen. In jenem Jahr begründete der Theologe und Afrikanist Carl Meinhof (1857–1944) das *Phonetische Laboratorium*, die weltweit erste Einrichtung zur experimentellen Untersuchung der menschlichen Sprachlautformung, angegliedert an die Abteilung für afrikanische Sprachen des Hamburgischen Kolonialinstituts.

Die historische Entwicklung der Sammlung bis hin zu ihrer derzeitigen Instantiierung in Form einer digitalen Datenbank sowie die hierbei angewandten Methoden und Technologien werden im Folgenden überblicksartig zusammengefasst.

## **Das Hamburgische Kolonialinstitut 1908–1914**

Das Hamburgische Kolonialinstitut wurde 1908 als erste staatliche Hochschule Hamburgs gegründet. Obgleich die frühen Förderer des Kolonialinstituts, insbesondere der spätere erste Bürgermeister Werner v. Melle und der damalige Direktor des Hamburgischen Museums für Völkerkunde, Georg Thilenius, die Vision einer allgemeinen Universität hegten, bestand seine vornehmliche Aufgabe zunächst darin, die aus den deutschen Kolonien und Schutzgebieten erbrachten Informationen und Objekte wissenschaftlich zu untersuchen und die hieraus gewonnenen Erkenntnisse wiederum für die Verwaltung der vom Deutschen Reich vereinnahmten Gebiete in Afrika, Asien und Ozeanien nutzbar zu machen. Während sich die institutionelle Forschungstätigkeit bald diversifi-

zierte (u.a. Völkerkunde, Geographie, Zoologie, Botanik, zahlreiche verschiedene Regional- und Sprachwissenschaften) beschränkte sich das Lehrangebot weitgehend auf ein zweisemestriges Kurrikulum zur Vorbereitung auf den gehobenen kolonialen Beamtendienst (Paul 2008).

Für Verwaltungsaufgaben von zentralem Interesse war naturgemäß die Beschäftigung mit den jeweiligen einheimischen Sprachen. Mit Institutsgründung wurde daher der weltweit erste akademische Lehrstuhl für Afrikanistik (Meinhof) geschaffen, dem im Jahr 1910 das Phonetische Laboratorium unter Giulio Panconcelli-Calzia (1878–1966) angegliedert wurde. Früh wurde hier der Wert des noch relativ neuen Mediums der mechanischen Schallaufzeichnung für die Sprachforschung erkannt und genutzt. Zu der Sammlung reiner Sprachaufnahmen kamen bald auch zahlreiche Tonträger mit Aufnahmen von Musik und Gesang hinzu, häufig auch von kommerziellen Schallplattenverlagen wie etwa Odeon und Parlophon-Lindström herausgegeben, die die Kolonialgebiete als Absatzmärkte für ihre Erzeugnisse entdeckt und dort bereits während der 1910er Jahre ein Händlernetz etabliert hatten.

Mit Kriegsbeginn 1914 zerfiel das kurzlebige deutsche Kolonialreich und das Kolonialinstitut verlor seine zentrale Aufgabe; die Einrichtung und zahlreiche der wissenschaftlichen Institute blieben zunächst jedoch bestehen.

### **Die Anfänge der Vergleichenden Musikwissenschaft in Hamburg**

Im Jahr 1915 begann der gelernte Fagottist Wilhelm Heinitz (1883–1963) eine Tätigkeit als wissenschaftlicher Mitarbeiter am phonetischen Laboratorium, wandte sich jedoch schnell der Untersuchung der musikalischen Äußerungen, die ihm in Form der vorhandenen Schallaufnahmen zugänglich waren, sowie der Untersuchung der in den Sammlungen des Instituts und des Museums für Völkerkunde (insbesondere auch durch die Erträge der Hamburger Südsee-Expedition von 1908–1910) vorhandenen Musikinstrumente zu. Heinitz' Forschungen stießen bei seinem Mentor Meinhof auf Wohlwollen: Noch im selben Jahr wurde die *Forschungsabteilung für vergleichende Musikwissenschaft* geschaffen und Heinitz als deren Leiter (und einziger fester Mitarbeiter) eingesetzt. Die Sammlungen wurden in den folgenden Jahrzehnten stetig erweitert.

Im Jahr 1919 erfolgte die Gründung der Universität Hamburg und viele der Einrichtungen des ehemaligen Kolonialinstituts gingen in noch heute bestehenden Asien-Afrika-Institut auf. Der 1911 eingeweihte Bau, der einst das Kolonialinstitut beherbergte, ist bis heute das Hauptgebäude der Universität Hamburg.

Wilhelm Heinitz wurde 1949 zum ersten Professor für Vergleichende Musikwissenschaft an das neugegründete Musikwissenschaftliche Institut berufen. Er überführte diejenigen Bestände der im Kolonialinstitut begonnenen Tonträgersammlung, die die Zerstörungen des Zweiten Weltkriegs überstanden hatten, an das Musikwissenschaftliche Institut. Heute sind in der Sammlung 654 Tonträger

– Schellackplatten, Acetat- und Metallfolienschnitte – mit insgesamt etwa 2.600 Einzelaufnahmen aus allen fünf Kontinenten erhalten.

### **Das E.S.R.A. -Projekt**

Im Rahmen des Projektes *E.S.R.A.*<sup>1</sup> (UHH Ethnographic Sound Recordings Archive) werden die erhaltenen Bestände seit Januar 2013 digitalisiert und die Digitalisate sowie zugehörige Metadaten in eine bedarfsoptimierte Datenbank eingepflegt. Die Erstellung der Digitalisate erfolgt hierbei nach international etablierten Richtlinien um ein Höchstmaß an Interoperabilität im Rahmen möglicher zukünftiger Standardisierungen zu gewährleisten (vgl. Elschek 2000).<sup>2</sup>

Der Zugriff auf die weltweit bestehenden Repositorien ethnographischer Audiodaten ist bisher, trotz zahlreicher entsprechender Bemühungen, durch einen geringen Grad an Vereinheitlichung und einen Mangel an nationaler wie internationaler Vernetzung erschwert; die wissenschaftliche Nutzbarkeit der vorhandenen Ressourcen ist durch das Fehlen einheitlicher Identifikatoren und bedarfsgerechter Parametersätze zur Referenzierung vorhandener Primärdatenbestände stark eingeschränkt.

### **Dateninfrastruktur – Vernetzung und Interoperabilität**

Die im Rahmen des Projekts entwickelte Metadatenstruktur berücksichtigt im Bereich der verwendeten Sprachen und Protokolle bereits bestehende Standardisierungsbestrebungen (DublinCore, OAI-PMH, XML) und ist bestrebt, eine einheitliche Semantik als Grundlage einer optimierten Programmierschnittstelle (*application programming interface*, API) zu etablieren, um das Ziel einer überregionalen Vernetzung gleichartiger Datenbanken und die Vereinheitlichung von Such- und Zugangsmodalitäten zu befördern.<sup>3</sup> Mittelfristiges Ziel des E.S.R.A.-Projekts ist daher, neben der Neuerschließung und einer adäquaten archivarischen Betreuung der Bestände, die nachhaltige Verbesserung der allgemeinen Verfügbarkeit musikwissenschaftlicher Forschungsdaten durch die Etablierung einer domänenübergreifenden digitalen Infrastruktur für die Kategorisierung und Klassifizierung von Digitalisaten historischer ethnographischer Schallaufnahmen und deren effizienter, dezentraler Verfügbarmachung durch flexible, datenbankübergreifende Suchstrategien (*audio-data-mining*) (vgl. Proutskova 2007). Derzeit wird ein exemplarisches Pilotsystem prototypisch implementiert.

---

<sup>1</sup> Siehe hierzu URL: <http://www.esra.uni-hamburg.de> [Zugriff vom 26.05.2015].

<sup>2</sup> International Association of Sound and Audiovisual Archives Technical Committee (2005); in URL: [http://www.iasa-web.org/IASA\\_TC03/TC03\\_English.pdf](http://www.iasa-web.org/IASA_TC03/TC03_English.pdf). [Zugriff vom 26.05.2015].

<sup>3</sup> Vgl. The National Recording Preservation Board of the Library of Congress (2010) sowie The National Recording Preservation Board of the Library of Congress (2012).

Eine grundsätzliche Schwierigkeit bei der Erschließung historischer Tonträgerbestände stellt die häufige Unvollständigkeit, nicht selten auch das vollständige Fehlen, einer zugehörigen Dokumentation dar. Eine manuelle Annotation historischer ethnographischer Schallaufnahmen würde ein hohes Maß an fachlicher Kenntnis der Musik der betreffenden Region und Epoche voraussetzen; der Zeit- und Arbeitsaufwand wäre enorm und schließt sich daher wohl selbst für langfristig und stabil finanzierte Institutionen weitgehend aus.

Die jüngsten Entwicklungen im Bereich der rechnergestützten Signalverarbeitung und Klanganalyse (*Music Information Retrieval*, MIR) ermöglichen es nunmehr, auch komplexe Schallvorgänge wie vollständige Musikaufnahmen in Form eines spezifischen „akustischen Fingerabdrucks“ zu repräsentieren. Solche direkt aus dem digitalen Audiosignal abgeleiteten Deskriptoren erlauben eine eindeutige Referenzierung und Kontextualisierung innerhalb selbst umfangreichster Korpora, unabhängig von Vollständigkeit und Sprache der Dokumentation und der Systematik der jeweils verwendeten Metadatenstruktur. Ein solch komplexer Deskriptor stellt zudem per se bereits ein Datenobjekt von großer analytischer Tiefe dar.

Innerhalb der derzeit entwickelten Implementation soll daher eine bedarfsrechte Metadatenstruktur aus (a) direkt aus dem digitalen Audiosignal abgeleiteten musikalisch-akustischen Deskriptoren und (b) standardisierten, semantisch integrierten deskriptiven Kategorien realisiert werden. Die E.S.R.A.-Datenbank stellt hierzu Primärdaten (in Form hochauflösender Digitalisate) zur Verfügung und bietet, im Sinne eines umfassenden, konstruktiven Dialogs mit der Fachgemeinschaft bereits im Entwicklungsprozess (detaillierte Bedarfsanalyse anhand unterschiedlichster Forschungsfragen, Weiterentwicklung von Best-Practices-Richtlinien), ein fachöffentliches Forum zur Evaluierung des Projektfortschritts. Die Formulierung objektspezifischer Parameter und deren Verknüpfung mit einem einheitlichen Repräsentationsformat strebt mithin eine weitere Professionalisierung von Informationsstrukturen, die Erschließung bisher latenter Ressourcen, sowie ein hohes Maß an Interoperabilität, langfristiger Verfügbarkeit und Nachnutzbarkeit an.

## **Das Music Information Retrieval System**

Die Suche in Katalogen und Archiven wird gewöhnlich mit Hilfe von Daten über das betreffende Objekt, so genannten Metadaten, realisiert. Im Fall von Musikarchiven sind dies zum einen Daten über das Stück selbst, wie z.B. Titel, KomponistIn, beteiligte MusikerInnen, Instrumentation, usw. Zum anderen fallen unter den Begriff „Metadaten“ aber auch alle weiteren außermusikalischen Spezifikationen, wie beispielsweise Medienart, Standort, Zeitpunkt der Erfassung oder Anschaffungsdatum. Um ein bestimmtes Stück in einer Sammlung ausfindig zu machen, vergleicht ein Algorithmus die vom Benutzer eingegebenen Metadaten mit denen aller im Archiv vorhandenen Musikstücke. Als

Ergebnis wird ein anderes Metadatum, beispielsweise der Standort des Tonträgers, an den Benutzer zurückgemeldet. Diese Methode des *text-based retrieval* ist sowohl das bekannteste als auch weit verbreitetste Modell der Kommunikation von Suchaufträgen. Neben den offensichtlichen Vorteilen birgt diese Herangehensweise zwei entscheidende Nachteile. Die Qualität der textbasierten Suche in einer Datenbank hängt von der Qualität der eingepflegten Metadaten ab. Sind in der Datenstruktur entscheidende Merkmale nicht repräsentiert, wie z.B. der Titel, kann eine Suche kein zufrieden stellendes Ergebnis liefern. Dies ist vor allem im Hinblick auf Transliterationen von entscheidender Bedeutung. So kann man nach „Pjotr Iljitsch Tschaikowski“ oder „Peter Tschaikowsky“ suchen, aber auch nach „Pëtr Il’ič Čajkovskij“. Jede Version wird unterschiedliche Ergebnisse liefern. Dieselbe Argumentation gilt auch auf Seiten der Benutzer. Eine Suchanfrage muss weitestgehend mit den gespeicherten Daten übereinstimmen. Im einfachsten Fall kann die Verwechslung von Majuskeln und Minuskeln oder von „ß“ und „ss“ zu ungewollten Ergebnissen führen.

Des Weiteren wird im text-based retrieval nur *über* Metadaten gesprochen. Das Musikstück an sich, bleibt völlig außer Acht. Deshalb kann eine text-basierte Suche Ähnlichkeiten zwischen Musikstücken nur in dem Maße berücksichtigen, in dem sie von Metadaten impliziert werden. Im Allgemeinen ist die Annahme berechtigt, dass unter einem Merkmal Objekte mit ähnlichen Eigenschaften klassifiziert sind. Speziell im musikalischen Kontext ist eine solche Annahme in der Regel nicht zielführend. Dies wird bei der Vergegenwärtigung musikalischer Genres sofort deutlich. Die Menge der unter „Populäre Musik“ subsummierten Stücke ist heterogen und zwar auch dann, wenn sie auf einen Zeitbereich eingegrenzt wird. *7 years* von Lukas Graham, die aktuelle Nummer Eins der deutschen Single-Charts, klingt wesentlich anders als *Lush Life* von Zara Larsson.<sup>4</sup> Es ist deshalb generell nicht zu erwarten, dass ein Suchauftrag nach Populärer Musik auch Stücke liefert, die musikalische Ähnlichkeiten aufweisen.

Die in E.S.R.A. enthaltenen Sammlungen weisen teilweise geringe Mengen oder keine Metadaten auf. Die Anwendung eines auf text-based retrieval basierten Systems ist daher nicht effizient möglich.

E.S.R.A. ist als eine über das Internet erreichbare Web-Applikation konzipiert und steht damit einem breiten Nutzerkreis aus WissenschaftlerInnen, MusikerInnen und Laien zur Verfügung. Der variable Grad an Expertise des genannten Personenkreises, im Umgang mit Archiven einerseits und musikwissenschaftlichen Fragestellungen andererseits, stellt höchste Ansprüche an die Flexibilität der Schnittstelle zwischen Anwender und Datenbank. Somit ist eine möglichst einfach zu erlernende und intuitiv zu bedienende Schnittstelle ein entscheidendes Qualitätskriterium eines Online-Archives. Die Gestaltung der Schnittstelle wirkt sich allerdings nicht nur einseitig auf die Bedienung des Archives

---

<sup>4</sup> Angaben nach URL: <http://www.officialcharts.com/> [Zugriff vom 17.03.2016].

und damit auf die Benutzer aus. Darüber hinaus hat die Schnittstelle umfassenden Einfluss auf die internen Abläufe der Datenbank. Dieser reicht von der Datenspeicherung bis hin zu Performanz der zur Suche verwendeten Algorithmen und bedingt somit auch die Qualität des Ergebnisses einer jeden Suchanfrage.

Aus diesen Gründen wird im E.S.R.A., neben einem textbasierten Suchsystem, eine *content-based retrieval* implementiert. Hierbei wird die Datenbank anhand von Merkmalen durchsucht, die direkt den Audiosignalen entnommen werden. Das System basiert daher nicht auf möglicherweise subjektiven und inkonsistenten Metadaten. Stattdessen wird ein empirischer Abgleich von Eingabe- und Objektdaten realisiert.

## **Systemarchitektur**

Das geplante music information retrieval system kombiniert ein stochastisches Mustererkennungssystem zur formalen Darstellung musikalisch-struktureller Aspekte mit einem künstlichen neuronalen Netzwerk. Dieses ordnet die zuvor gewonnenen Strukturen gemäß ihrer Ähnlichkeit auf einer zweidimensionalen Karte.

In der momentanen Implementation werden die Möglichkeiten einer auf Klangfarbe basierenden, stochastischen Rhythmustheorie zur Darstellung rhythmischer Ähnlichkeit erprobt. Der Grundgedanke dieser Theorie ist, dass die Wahrnehmung von Rhythmus durch die Klangfarbe der eingesetzten Musikinstrumente beeinflusst wird. Diese Annahme bietet den Vorteil, dass eine rhythmische Analyse allein auf Grundlage der Audiosignale möglich wird. Dadurch ist die Analyse frei von musiktheoretischen Voraussetzungen und kann kulturübergreifend eingesetzt werden.

Zur Realisierung dieses Vorhabens muss geklärt werden, welche Merkmale in einem Audiosignal zur Darstellung der Änderung der Klangfarbe im zeitlichen Verlauf geeignet sind. Außerdem müssen diese Merkmale an rhythmisch relevanten Zeitpunkten gemessen werden. Weiterhin muss die Zeitreihe der gemessenen Merkmale durch das Mustererkennungssystem analysiert werden, um eine empirische Repräsentation eines jeden Stückes im Archiv zu generieren. Schließlich müssen die abstrakt-mathematischen Repräsentationen hinsichtlich ihrer Ähnlichkeit geordnet und auf eine intuitiv zu verstehende Art dargestellt werden.

## ***Feature Extraction***

Durch Untersuchungen konnte festgestellt werden, dass die Unterscheidbarkeit von Klangfarben hauptsächlich auf der wahrgenommenen Helligkeit basiert (vgl. Grey 1977). Daher wird zur Analyse im Falle von E.S.R.A. nur der *spektrale Zentroid* gemessen. Dabei handelt es sich um ein Maß der Energiever-