

ihrer Zeit. Von großem Interesse – vor allem vielsagend hinsichtlich der unterschiedlichen Persönlichkeiten – ist dabei ihr Umgang etwa mit gemeinsamen Schüler:innen.

Aufgrund der Quellenlage erscheint die ganze Arbeit auf den ersten Blick eher wie eine Arbeit über Hans von Bülow und weniger über Joachim Raff – was unter anderem an der nennenswert hohen Zahl an Briefen von Bülows liegt, die erhalten und hier publiziert sind. Dabei finden allerdings einschneidende Ereignisse in den Leben der beiden Musiker nur begrenzt Berücksichtigung, weil sie nur teilweise oder gar nicht im Fokus ihres Briefwechsels standen. So ist die Zeit von Bülows als Leiter der neugegründeten Königlichen Musikschule München nur begrenzt ein Thema des Buchs – beide dürften sich in dieser Phase mehrheitlich persönlich und nicht über Briefe ausgetauscht haben.

Kannenberg geraten zeitgemäße Formate insbesondere hinsichtlich einer digitalisierten und damit vielfältig nutzbaren Edition des Briefcorpus nicht aus dem Blick: Parallel zu der vorliegenden Printedition des Briefwechsels zwischen Raff und von Bülow ist – in der Art einer niedrigschwelligen Hybridausgabe – eine eBook-Version erschienen, in der nicht nur sämtliche URLs direkt, sondern die Quellensignaturen der Briefedition mit den Datenbankeinträgen des Online-Portals des Joachim-Raff-Archivs unmittelbar miteinander verknüpft sind. Diese Version ist die für eine wissenschaftliche Auseinandersetzung mit dem Briefcorpus adäquatere, wenngleich die Publikation in Buchform überaus wertvoll und lesefreundlich ist. Immerhin befinden sich in dieser Briefedition auch eine ganze Reihe von Briefen der Ehefrauen Doris Raff und Marie von Bülow, Hans von Bülows zweiter Frau, die sich unmittelbar auf die Musiker und ihr Binnenverhältnis beziehen. Im Ganzen besteht der editorische Teil der Veröffentlichung aus 176 Einzelbriefen und -konvoluten, wobei Kannenberg als Herausgeber vielfach aufeinander bezogene Briefe miteinander verbunden hat und in seinen

vielschichtigen Anmerkungen, aber auch in einer instruktiven Einleitung aufdeckt, wie weitreichend bisweilen die Eingriffe der Ehefrauen und / oder der Tochter Ruffs, Helene Raff, als Nachlassverwalterin in älteren, teilweisen Veröffentlichungen von Briefen gewesen sind.

Beiläufig erwähnt Kannenberg, dass die Freunde Raff und von Bülow auch ein nicht abgeschlossenes Promotionsverfahren verbindet – ein Umstand, der es nach Auskunft einer Fußnote dem Verfasser nicht gerade erleichtert hat, „die Anfertigung der eigenen Arbeit durchzuhalten“ (S. 46). Das schließlich obsiegende Durchhaltevermögen bereitete Simon Kannenberg eine beachtliche Dissertationsleistung, die vollkommen zu Recht mit dem Dissertationspreis der Gesellschaft für Musikforschung 2020 ausgezeichnet wurde – und allen Interessierten, die sich mit Künstler:innen-Netzwerken des 19. Jahrhunderts in Europa befassen, bereitet die Publikation nicht nur ein Lesevergnügen, sondern auch einen bemerkenswerten Fundus an Quellenmaterial, das weit über die engere Musikerfreundschaft hinauszuweisen vermag.

(Februar 2022)

Birger Petersen

FELIX CHRISTIAN THIESEN: Mikroklänge – Plinks. Zur Erkennbarkeit kürzester musikalischer Klangobjekte. Baden-Baden: Tectum Verlag 2021. 298 S. (Wissenschaftliche Beiträge aus dem Tectum-Verlag: Musikwissenschaft. Band 15.)

Wie lange brauchen wir, um Musik einordnen zu können? Erfahrungsgemäß müssen wir nicht lange hinhören, wenn wir gerade im Stau von einer Radiostation zur nächsten wechseln, um zu entscheiden, ob wir bei der angewählten Radiostation bleiben oder weitersuchen. Woran machen wir diese Entscheidungen fest? Welche Aspekte der Musik können wir überhaupt wahrnehmen

in diesen kurzen Musikausschnitten, den sogenannten „Plinks“?

In der Monographie „Mikroklänge – Plinks. Zur Erkennbarkeit kürzester musikalischer Klangobjekte“ formalisiert Felix Christian Thiesen diese Fragen im Rahmen von umfangreicher, empirischer Arbeit. Die Kontextualisierung dieser Fragen in der historischen Wahrnehmungsforschung aus dem 19. und 20. Jahrhundert, sowie in künstlerischen Perspektiven des 20. Jahrhunderts, die Thiesen in dem Einleitungsteil vornimmt, ist äußerst lesenswert und wunderschön bebildert. Bei der Erörterung der bisher getätigten Forschung wird dabei deutlich, wie viel noch zu tun ist, um diese Fragen beantworten zu können, und wo genau die Baustellen sind, die Thiesen mit seiner empirischen Arbeit angehen möchte.

Thiesen hebt dabei hervor, dass die existierende Forschung vor allem deswegen nicht unter einen Hut gebracht werden kann, weil verschiedene Forscher:innen die Fragen verschieden formalisierten. Es wurde zum Beispiel gefragt, wie lange Plinks sein müssen, dass Hörer:innen den Titel des Stücks erkennen oder bis Hörer:innen eine emotionale Reaktion festmachen. Umso beeindruckender ist Thiesens Versuch, die unterschiedlichen Schwellenwerte aus der Forschung in einer Graphik zu vereinen, die für weiterführende Forschung eine Ressource von großem Wert sein wird (Abb. 5, S. 45).

Nur eine noch um Ergebnisse aus Forschung mit physiologischen Methoden erweiterte Abbildung wäre wertvoller gewesen. Zum Beispiel sind elektrophysiologisch abbildbare Erkennungsleistungen im Bereich von 150 ms bis 250 ms in der Arbeit von Stefan Koelsch (Music-syntactic processing and auditory memory: Similarities and differences between ERAN and MMN, in: *Psychophysiology* 46 [2009], S. 179–190) bekannt, und neuere mit Hilfe von Pupillometrie gewonnene Erkenntnisse von Robert Jagiello und Kollegen (Rapid Brain Responses to familiar vs. unfamiliar Music – an EEG and

pupillometry study, in: *Scientific Reports* 9 [2019], S. 1–13) deuten auf Erkennungsleistungen nach 100 ms hin. Diese hätten den als relevant identifizierten Zeitbereich weiter unterstützt und zu einer ausführlicheren Zusammenstellung beigetragen.

Psychoakustische Klangdeskriptoren werden in einem eigenen, einleuchtenden Kapitel in der Einleitung behandelt. Dieses unterstreicht, dass diese Klangdeskriptoren als potentielle Prädiktoren unserer rapiden Musikererkennung dienen können. Hier wäre eventuell eine Einordnung der Forschung zu crossmodalen Variablen von Elke Lange und Karl Frieler (Challenges and opportunities of predicting musical emotions with perceptual and automatized features, in: *Music Perception* 36 [2018], S. 217–242) nützlich gewesen, die in dem folgenden Teil über die getätigten empirischen Untersuchungen Einfluss nimmt.

Diese Untersuchungen erstrecken sich über drei Pilotstudien und zwei großangelegte Onlinestudien mit jeweils mehr als 500 Proband:innen. Zunächst wird der Frage nachgegangen, wie lange man braucht, um ein Stück wiederzuerkennen; danach, welche Aspekte der Musik zu einer möglichst genauen Einordnung der Musik beitragen. Dafür stellt Thiesen ein von ihm entwickeltes Prinzip vor, das sogenannte Matrjoschka-Prinzip, mit Hilfe dessen er möglichst im Timbre homogene Ausschnitte aus Songs extrahieren möchte.

So wie sich in einer Matrjoschka Puppen ineinander verschachteln, so verschachteln sich die gewonnenen Stimuli ineinander: Zuerst wird ein 800 ms langer Ausschnitt aus einem strukturell bedingten Teil des Songs extrahiert, danach ein 400 ms langer Ausschnitt aus dem längeren Ausschnitt gewonnen, und so weiter. In der ersten Studie wird mit Hilfe von Music Information Retrieval (MIR)-Ansätzen validiert, dass so gewonnene Stimuli in sich konsistenter sind als Stimuli, die nicht mit Hilfe des Prinzips gewonnen wurden.

Was dabei unklar bleibt: Warum nicht ebendiese MIR-Ansätze genutzt worden sind, um die maximal konsistenten oder prototypischen Stellen eines Songs zu bestimmen. Dazu hätte man einen Song in 50 ms lange Segmente trennen, die psychoakustischen Klangdeskriptoren dieser Segmente mit Hilfe von MIR errechnen und schließlich die für einen 800 ms langen Stimulus benötigten 16 aufeinanderfolgenden Segmente auswählen können, für welche diese Klangdeskriptoren am meisten miteinander korrelieren. Wie stark der so determinierte Ausschnitt mit dem Deskriptorenmittelwert korreliert, gäbe dann Auskunft darüber, wie prototypisch der Ausschnitt für den jeweiligen Song ist. Sollte, wie Thiesen vermutet, der Entnahmezeitpunkt für die Erkennung maßgeblich sein, so müsste man hier sehen, dass die konsistentesten Ausschnitte vermehrt im Refrain anstatt in Strophen auftreten.

Ein anderer Ansatz für die zweite Vorstudie zur Bestimmung von möglichen Zielkategorien, also, ob Genreerkennung eine bessere Formalisierung darstellt als Titelerkennung oder die Erkennung eines vorhandenen Instruments, wäre ebenfalls möglich gewesen. Da hier das Ziel ist, anhand von Expertenübereinstimmung sich so gut wie möglich den wahren, aber leider unbekanntem Instrumental- oder Vokalkomponenten, den sogenannten Arrangementteilen, von Songs anzunähern, wäre hier der Einsatz der längeren Stimuli allein angemessen gewesen. Die Auswahl der Songs und Zielkategorien zum Zweck einer Teilreplikation einer existierenden Studie von Carol Krumhansl (Plink: „Thin slices“ of music, in: *Music Perception* 27 [2010], S. 337–354) ist hier zu erwähnen und zu loben.

Allerdings bleibt eine Diskussion darüber aus, ob Übereinstimmungen über das Nichtvorhandensein eines Arrangementteils gleichwertig mit Übereinstimmungen über das Vorhandensein eines Arrangementteils sind. Ein so möglicher Bias könnte dann letztendlich zu einer Verminderung der Aus-

sagekraft darüber führen, welche Arrangementteile besonders die Musikererkennung unterstützen, und zu Verzerrungen zugunsten des Einflusses von einzelnen Stücken führen, die ein seltenes Arrangementteil beinhalten, so zum Beispiel das auch von Thiesen erwähnte Stück „Uptown Funk“. So sind die Ergebnisse aus der sonst anderweitig sehr umsichtig geplanten ersten Onlinestudie mit leichten Zweifeln behaftet.

Besonders vorausschauend hier ist das Miteinbeziehen von crossmodalen Variablen, die es Laienproband:innen vereinfachen, akustische Eigenschaften zu beschreiben. Allerdings werden diese Daten nicht vollständig ausgeschöpft – so hätte man wie vorher angedeutet, die Relationen zwischen diesen Beschreibungen mit Hilfe der von MIR errechneten Klangdeskriptoren analysieren können. Ebenso fehlen Cronbach's α -Werte zur Bewertung der Homogenität der ermittelten crossmodalen Faktoren.

Trotzdem sind die Ergebnisse der ersten Onlinestudie äußerst interessant, da sie mit Hilfe von Conditional Inference Trees zeigen, wie sich musikalische Erkennung entfalten könnte. Insgesamt werden in Thiesens Studie Titel und Sänger:in weniger häufiger erkannt als in der zu replizierenden Studie. Allerdings kann dies dadurch verantwortet sein, dass in der vorherigen Studie von Krumhansl (s. o.) Stimuli aus weitaus mehr Songs oder Quellmusikstücken geschnitten wurden, und diese vielleicht den Proband:innen besonders bekannt waren. Dementsprechend wäre die zusätzliche Erhebung von Familiaritätswerten von den Titeln allgemein nützlich gewesen, um festzustellen, ob geringe Erkennungsleistungen sich dadurch erklären, dass die Songs nicht bekannt sind.

Wenn man zum Beispiel Proband:innen von einer Liste von Stücken auswählen ließe, welche ihnen bekannt sind, und dann per Zufall die Quellstücke auswählt, aus denen man vielleicht mit Hilfe des oben beschriebenen Ansatzes die konsistentesten Ausschnitte berechnet, dann müsste man auch nicht mit der

Annahme operieren, dass musikalisch vorgebildete Proband:innen gleichzeitig auch über ein größeres musikalisches Lexikon verfügen. Dieses Lexikon muss schließlich auch nicht unbedingt Populärmusik beinhalten. Letztendlich wäre die Frage danach, wie lange man zuhören muss, bis man die Musik erkennt, erst dann möglich, wenn man vorher sicherstellt, dass die Musik bekannt ist.

Die zweite Onlinestudie und deren Vorstudie stellen eine wichtige und äußerst notwendige Erweiterung der vorhandenen Literatur dar, denn hier erstellt Thiesen Stimuli, von denen die Arrangementteile bekannt sind. Aus den Daten der Vorstudie ist direkt ersichtlich, dass die Anzahl der Arrangementteile Einfluss darauf hat, wie differenziert Expert:innen darin sind, alle Teile zu erkennen. Obwohl Thiesen hier schlussfolgert, dass bei Stimuli mit fünf Arrangementteilen und einer Länge von 400 ms „mit 17% korrekten Zuordnungen [...] nicht von einer guten Erkennungsleistung gesprochen werden kann“ (S. 144), wäre ein Hinweis darauf, dass eine zufällige Erkennungsleistung bei 50%⁵ Arrangementteile \approx 3% läge, nützlich, um die erbrachten Leistungen angemessen zu bewerten.

Thiesens zweite Onlinestudie ist mithin das bisher aussagekräftigste Experiment über die chronologische Entfaltung der Musikwahrnehmung. Mit dieser Studie steht nun eine Fülle an Daten bereit, aus denen Thiesen zum Beispiel folgert, dass für das richtige

Einordnen aller Arrangementteile die Erkennung von Gitarre, Klavier und Schlagzeug besonders wichtig sind. Hier könnte man nun direkt weitermachen, um die chronologische Entfaltung der Musikererkennung weiter zu spezifizieren.

Zum Beispiel könnte man die Erkennungsleistung für spezielle Instrumente in einem Song als Zielvariable modellieren. Damit könnte man dann Aussagen darüber treffen, inwieweit Stimuluslänge und Präsenz von bestimmten anderen Instrumenten oder Stimmen dazu führt, dass ein Klavier erkannt wird. Entsprechende Modelle für andere Instrumente könnten dann miteinander verglichen werden, um eventuelle Unterschiede in der Erkennung verschiedener Instrumente auf psychoakustische Klangdeskriptoren zurückzuführen.

Auch ohne diese Ergebnisse bietet Thiesens Monographie einen umfangreichen, bereichernden und inspirierenden Einblick in die Plink-Forschung und hat wohl auch deswegen den Dissertationspreis der Gesellschaft für Musikforschung 2021 erhalten. Sie weist Wege auf, in denen sich die musikpsychologische mit musikinformatischer Forschung gewinnbringend vereinen kann. Darüber hinaus kann Thiesens Arbeit auch Hinweise für die künstlerische Praxis, vor allem der kompositionstechnischen Arbeit mit kurzen Klangfragmenten liefern.

(Dezember 2021)

Anja-Xiaoxing Cui