

Neuwirth, Markus / Rohrmeier, Martin (2016): Wie wissenschaftlich muss Musiktheorie sein? Chancen und Herausforderungen musikalischer Korpusforschung. ZGMTH 13/2, 171–193. <https://doi.org/10.31751/915>

© 2016 Markus Neuwirth, Martin Rohrmeier



Dieser Text erscheint im Open Access und ist lizenziert unter einer Creative Commons Namensnennung 4.0 International Lizenz.

This is an open access article licensed under a Creative Commons Attribution 4.0 International License.

eingereicht / submitted: 01/04/2017

angenommen / accepted: 29/04/2017

veröffentlicht / first published: 30/06/2017

zuletzt geändert / last updated: 10/05/2018

# Wie wissenschaftlich muss Musiktheorie sein? Chancen und Herausforderungen musikalischer Korpusforschung<sup>1</sup>

Markus Neuwirth, Martin Rohrmeier

**ABSTRACT:** Korpusbasierte Forschung nimmt in der Sprach- und Literaturwissenschaft schon seit Langem einen wichtigen Platz ein. In der Musikforschung dagegen gewann sie erst vor Kurzem an Bedeutung. Die Gründe für diese verspätete Akzeptanz sind vielfältig und mitunter einer tiefgreifenden Skepsis gegenüber der Anwendung statistisch-quantitativer Methoden auf Musik als Kunstobjekt geschuldet. Der vorliegende Beitrag motiviert musikalische Korpusforschung, indem er grundsätzliche Probleme herkömmlicher Repertoireforschung (intuitive Statistik, methodische Intransparenz, Urteilsheuristiken) und gegenwärtiger Korpusforschung (z. B. Stichprobenerhebung, mangelnde Korpora und Annotationsstandards) aufzeigt und anhand repräsentativer Studien in den Bereichen Harmonik, Kontrapunktik, Melodiebildung und Rhythmik/Metrik exemplarisch diskutiert. Der Beitrag schließt mit einem Plädoyer für die Einbeziehung quantitativer Ansätze in der Musiktheorie im Rahmen eines übergeordneten ›Mixed Methods‹-Paradigmas.

Corpus-based research has long been occupying a prominent position in literary studies and linguistics. In musicology, by contrast, it is about to gain in importance only fairly recently. The reasons for this delayed acceptance are manifold. Among other things, they are rooted in a deep skepticism toward applying statistical-quantitative methods to music as an object of art. This article supports musicological corpus research by pointing out general problems inherent to traditional repertoire research (intuitive statistics, methodological non-transparency, and heuristics in judgment) as well as current corpus research (e.g., biased sampling, paucity of corpora, and lack of annotation standards). These problems are discussed in reference to prominent studies in the domains of harmony, counterpoint, melody, and rhythm/meter. The article concludes by making a case for the integration of quantitative approaches in music theory into the overarching framework of a ›mixed methods‹ paradigm.

## 1. Musiktheorie und Korpusforschung: Ein Zustandsbericht

Wie häufig und in welchen Situationen folgt im ›klassischen Stil‹ die IV. auf die V. Stufe? Wann begann die Folge ›Romanesca→Prinner‹ im Verlauf des 18. Jahrhunderts an Bedeutung zu verlieren? Gibt es einen korrelativen Zusammenhang zwischen der monothematischen Anlage der Exposition und der Rekombination des Seitensatzes in der

<sup>1</sup> Die Arbeit an vorliegendem Beitrag wurde u. a. durch das »Zukunftskonzept« der *Technischen Universität Dresden* im Rahmen der Exzellenzinitiative der *Deutschen Forschungsgemeinschaft* finanziell unterstützt. Wir danken den anonymen Gutachtern sowie Fabian Moss für wertvolle Hinweise und Anregungen zu früheren Textfassungen. Wie im Verlauf des Texts deutlich wird, ist der Titel einem ähnlich lautenden Aufsatz von Michael Polth nachempfunden (Polth 2000b).

Reprise? Wie erfolgt in tonaler Musik üblicherweise die Platzierung von Kadenzen im Taktschema? Gibt es bevorzugte Klangfolgen in Lassos Duetten? Wie verbreitet sind absteigende Basstetrachorde in den Songs der Beatles?

Diese und viele weitere Fragen sind empirischer Natur und lassen sich nur nach Durchsicht umfangreicher Repertoires zuverlässig beantworten. Die streng-quantitative Untersuchung großer Datenmengen in maschinenlesbarer Form – kurz ›Korpusforschung‹ – stellt eine zentrale Strategie zur Beantwortung generalisierender Fragestellungen<sup>2</sup> dar und ist innerhalb der von Eric Clarke und Nicholas Cook 2004 wirkmächtig proklamierten ›Empirical Musicology‹ (bzw. ›Digital Musicology‹) zu verorten.<sup>3</sup> Besagte empirische Fragestellungen können die Charakterisierung eines musikalischen Systems (z. B. Dur-Moll-Tonalität, hindustanischer Raga, türkischer Maqam), eines Stils (z. B. Mozarts Stil, Verdis Melodiebehandlung, Jazz-Harmonik) oder musikalischer Phänomene (z. B. Formen von Kadenzen) betreffen. Darüber hinaus eröffnet die Analyse von Metadaten, Textmaterial und anderen Formen von Sekundärquellen ein enormes Potenzial für die Musikforschung.

Anders als in der Musiktheorie ist die systematische, computergestützte Analyse von Korpora in der Sprach- und Literaturwissenschaft bereits seit Langem ein etabliertes Forschungsparadigma, das die theoretische Forschung produktiv ergänzt.<sup>4</sup> Die Gründe für die verspätete Akzeptanz von Korpusforschung in der Musiktheorie sind vielfältig. Sie liegen wohl u. a. darin, dass das Selbstverständnis der Musiktheorie als einer wissenschaftlichen Disziplin (mit dem Ziel einer formal-wissenschaftlichen Beschreibung musikalischer Systeme) mit dem einer ›bloßen‹ Handwerkslehre bzw. generell einer pädagogischen Disziplin ebenso konkurriert wie mit dem eines historisch-hermeneutischen Fachs.<sup>5</sup> Darüber hinaus gehören computationale und statistische Verfahren nicht zum selbstverständlichen Rüstzeug Musikforschender<sup>6</sup>, und es herrscht unter vielen Musiktheoretiker\*innen und historischen Musikwissenschaftler\*innen noch stets eine tiefgreifende Skepsis gegenüber der Anwendung statistisch-quantitativer Methoden auf Musik: Diese würden kaum neue Erkenntnisse liefern und bestehende musiktheoretische Intuitionen lediglich

- 2 Entsprechend hierzu ist in der kognitiven Musikforschung die Annahme verbreitet, dass mentale Repräsentation des impliziten Wissens eines musikalischen Systems oder Stils auf statistischen Merkmalen beruht, wie sie in der Korpusforschung charakterisiert werden (vgl. Huron 2006; Rohrmeier/Rebuschat 2012).
- 3 Vgl. Clarke/Cook 2004; die Themenhefte 31/1 und 31/3 der Zeitschrift *Music Perception* (darin: Huron 2013; Temperley/VanHandel 2013; Gjerdingen 2014; Ito 2014; White 2014) sowie 11/1 der Zeitschrift *Empirical Musicology Review* (darin: White/Quinn 2016); außerdem Quinn 2014.
- 4 Vgl. z. B. Manning/Schütze 1999 und Jannidis/Kohle/Rehbein 2017.
- 5 Vgl. Brown/Dempster 1989; Holtmeier 1997; Temperley 2001a und 2001b; Sayrs/Proctor 2008. In vielen Traktaten des 18. Jahrhunderts etwa mischen sich empirisch-deskriptive Äußerungen mit präskriptiv-deontologischen. Diese Traktate hatten aber in der Regel nicht primär den Anspruch, die kompositorische ›Realität‹ passgenau zu beschreiben, sondern zielten vielmehr darauf ab, dem angehenden Komponisten (oder Musiker) brauchbare Rezepte an die Hand zu geben, ästhetisch zufriedenstellende Produkte anzufertigen.
- 6 So gibt es bedauerlicherweise kaum Statistik- und Methoden-Lehrbücher, die auf die spezifischen Bedürfnisse von empirischer Musikforschung zugeschnitten wären – ein Umstand, der die Integration quantitativer Methoden in den Forschungsalltag von Musikwissenschaftlern deutlich erschwert. Zu den wenigen Ausnahmen zählen Beran 2004 sowie Weihs/Jannach/Vatolkin/Rudolph 2016.

bekräftigen. Außerdem würden generalisierende Korpusansätze dem Kunstcharakter von Musik und der ästhetischen Individualität einzelner Werke zuwiderlaufen – ein Argument, das genau genommen nicht auf den Wert von Computerverfahren an sich zielt, sondern auf den epistemischen Gehalt generalisierender Beobachtungen.

Durchforstet man die einschlägige musiktheoretische Literatur nach kritischen Äußerungen zum Stellenwert von Statistik, wird man mühelos fündig. Bei Charles Rosen etwa heißt es, die Untersuchung von Auftretenshäufigkeiten sei hinsichtlich der Bestimmung von musikalischer Form nichts anderes als die »postulation of a meaningless statistical abstraction«, die nichts zu einem adäquaten Verständnis dessen beitragen würde, wie Musik zu ihrer Zeit (und bis heute) funktioniere.<sup>7</sup> Auch Paul Wingfield verweist im Kontext der musikalischen Formenlehre auf die »limitations of the statistical evidence« und argumentiert, dass »drawing up hierarchies of ›defaults‹ on the basis of such basic statistical analysis is a questionable enterprise, because it tells us nothing about the unique network of connections surrounding a particular work.«<sup>8</sup>

Schon Carl Dahlhaus erkannte, dass statistische Häufigkeit nicht mit einer Norm (etwa im Bereich der Formenlehre) oder mit dem Weber'schen Idealtyp gleichgesetzt werden kann:

Bloße Statistik, die sich neutral dünkt, würde sich einerseits ins Unendliche verlieren und andererseits niemals zu einem Formbegriff gelangen, da bereits die scheinbar unverfängliche Voraussetzung, das Häufigste sei das Wesentliche, eine ästhetische Vorentscheidung wäre, die an den statistisch erfaßbaren Tatbestand von außen herangetragen werden müßte.<sup>9</sup>

Etwas differenzierter bringt es Leonard Meyer zum Ausdruck:

One danger of employing statistics is the temptation merely to count [...]. The problem with mere counts is that they produce averages or means. What we really need to know, however, is the relation of the average to some norm or ideal type. Only in this way can the characteristic features of the specific style be recognized.<sup>10</sup>

Zwei Aspekte sind hier hervorzuheben: Erstens, so bedeutsam der Hinweis auf die (mögliche) Divergenz von kompositorischer Norm und statistischem Durchschnitt auch ist, er macht die Kenntnis von Letzterem nicht entbehrlich; im Gegenteil: Wie Meyer zu Recht nahelegt, ist die Kon- bzw. Divergenz der beiden Komponenten an sich bereits informativ. Zum zweiten liegt den Aussagen von Dahlhaus und Meyer ein überaus simples Verständnis von statistischer Modellierung zu Grunde, das auf Methoden des 21. Jahrhunderts, wie sie aktuell etwa in den Bereichen des ›Data Mining‹ und ›Machine Learning‹ zur Anwendung kommen, in keiner Weise zutrifft.<sup>11</sup>

7 Vgl. Rosen 1988, 7 (Zitat: ebd.).

8 Wingfield 2008, 171.

9 Dahlhaus 1978, 156. Zum Problem des Weber'schen Idealtyps und seiner Anwendung im Bereich der Musikforschung vgl. Gossett 1989; Neuwirth 2011.

10 Meyer 1989, 60.

Verlässt man den Bereich der musikalischen Formenlehre und wendet sich etwa der Untersuchung von Tonalität zu, so findet man auch hier kritische Stimmen zur Angemessenheit quantitativer Ansätze. In dem lesenswerten Aufsatz »Wie wissenschaftlich kann Analyse sein?« argumentiert Michael Polth, Tonalität (verstanden als System musikalischen Zusammenhangs) ließe sich nicht ohne Berücksichtigung von Kontext und idealerweise aller Tonsatzparameter erfassen.<sup>12</sup> Dies schließe einen wissenschaftlichen Ansatz, der sich naturgemäß um Abstraktion und Reduktion bemüht, a priori aus, denn »[w]er von jeglichem Vorverständnis und jeglichem Kontext abstrahiert, abstrahiert von der Tonalität selbst«<sup>13</sup> und verliert, so könnte man ergänzen, eben den Gegenstand aus dem Blick, den man zu untersuchen vorgibt. Polth exemplifiziert seine Kritik an Roland Eberleins Studie *Die Entstehung der tonalen Klangsyntax* (1994), deren Autor er nicht nur einen »überspitzten Begriff von Wissenschaftlichkeit«<sup>14</sup> attestiert, sondern ein Unterfangen zuschreibt, das von vornherein zum Scheitern verurteilt gewesen sei. Sofern es in der tonalen Zusammenhangsbildung um die Artikulation von klanglichen Qualitäten gehe, erweise sich ein mit Quantitäten operierender statistischer Zugang per se als unzureichend. Obwohl Polth in seiner Diagnose, in der Handhabung eines anspruchsvolleren Tonalitätsbegriffs sowie in Bezug auf seine Mahnung zur Reflexion durchaus zuzustimmen ist, ist die daraus gezogene Schlussfolgerung (Verzicht auf quantitative Verfahren) möglicherweise keine zwingende Konsequenz.

Der vielfach vorgebrachte Positivismusvorwurf (in seinen verschiedenen Spielarten), generalisierende quantitative Forschung sei nicht in der Lage, die im jeweiligen historischen Einzelfall realisierten Qualitäten zu erfassen, ist häufig eine beinahe reflexartige Reaktion auf korpusgestützte Untersuchungen. Die in der historisch oder kulturwissenschaftlich geprägten Musikforschung weitverbreitete Skepsis gegenüber quantitativen Ansätzen beruht im Kern auf der Befürchtung, diese würden qualitativ-hermeneutische Ansätze ersetzen und menschliche Analytiker\*innen auf lange Sicht entbehrlich machen. Wie wir in diesem Beitrag zeigen möchten, schließen historisch differenzierendes, kontextsensitives ›close reading‹ und tendenziell vergrößerndes ›distant reading‹<sup>15</sup> sich nicht aus, sondern ergänzen einander auf produktive Weise: ›close reading‹ liefert Hypothesen, die sich auf einer breiteren Datenbasis überprüfen lassen; ›distant reading‹ dagegen liefert wichtige Erkenntnisse, auf die sich detaillierte Einzelfallanalyse zusätzlich stützen kann. Qualitativ-hermeneutische Fragestellungen können also durch Korpusforschung auf eine neue Ebene von Wissenschaftlichkeit gehoben werden, ohne an Legitimität einzubüßen: Qualitativ-hermeneutische Methoden operieren in der Regel mit einer größeren Anzahl an Kategorien, die sich darüber hinaus im Sinne einer »dichten Beschreibung« (Clifford Geertz) durch ein höheres Maß an Granularität auszeichnen als die idealerweise wenigen, großkörnigeren Kategorien der Korpusforschung. Genau aus

11 Vgl. zu »Data Mining« und maschinellem Lernen etwa mit Blick auf die digitalen Geisteswissenschaften Jannidis/Kohle/Rehbein 2017.

12 Polth 2000b, 64.

13 Ebd., 66.

14 Ebd.

15 Vgl. Moretti 2013.

diesem Grund lässt sich eine Korpusstudie auch nicht aus der Kumulation vieler kleiner hermeneutischer Analysen gewinnen. Die vielfältigen Möglichkeiten des Zusammenwirkens quantitativer und qualitativer Verfahren in einem integrativen ›Mixed Methods‹-Rahmen auszuloten, bleibt eine der dringlichen Aufgaben zukünftiger Forschung.

## 2. Probleme und Herausforderungen für die Korpusforschung

Im folgenden Abschnitt behandeln wir drei Problemfelder der Korpusforschung. Zunächst (1) die Probleme der intuitiven Statistik und der methodischen Intransparenz; dann (2) Heuristiken, wie die Verfügbarkeitsheuristik und den Bestätigungsfehler, die sich verzerrend auf Urteilsbildungen auch im Kontext wissenschaftlicher Forschung auswirken können; und schließlich (3) das Problem verzerrter und zu geringer Stichproben.

### 2.1 Intuitive Statistik und methodische Intransparenz

De facto ist die exemplarische oder exhaustive Untersuchung von Korpora (im Sinne von ›Repertoires‹) und eine darauf basierende Typenbildung keinesfalls eine Seltenheit in der musikwissenschaftlichen und -theoretischen Forschung, sondern vielmehr die Regel. Einige einschlägige Beispiele aus dem Bereich des ›klassischen Stils‹:

1. Ethan Haimo schlägt eine Rekonstruktion kompositorischer Normen im sinfonischen Schaffen Joseph Haydns vor. Zwar zeigt er diese Normen nur an einigen wenigen Beispielen in extenso auf, ordnet aber die jeweiligen kompositorischen Entscheidungen Haydns vor dem Hintergrund einer Personalnorm ein, die offenbar von einem umfangreicheren (wenn auch nicht näher bestimmten) Repertoire abgeleitet wurde.<sup>16</sup>
2. In *Classical Form* unternimmt William Caplin den Versuch einer idealtypischen Rekonstruktion formaler Optionen in der Musik von Haydn, Mozart und Beethoven zwischen ca. 1780 und 1810.<sup>17</sup>
3. Michael Polth untersucht in seiner Studie zu *Sinfonieexpositionen des 18. Jahrhunderts* einen breiten Korpus vorklassischer Sinfonien und schlägt auf dieser Basis distinkte Formtypen vor.<sup>18</sup>
4. In ihrer voluminösen Abhandlung *Elements of Sonata Theory* von 2006 gehen James Hepokoski und Warren Darcy dahingehend über Caplin hinaus, dass sie ein deutlich umfangreicheres Repertoire einbeziehen (auch im Hinblick auf die Entstehungszeit der berücksichtigten Werke), um eine detailliertere Typologie von Sonatenformen entwickeln zu können.<sup>19</sup>

Die Liste ließe sich beliebig fortsetzen. Trotz der jeweiligen Unterschiede in der Wahl des konkreten Gegenstands, der analytischen Methode oder des Repertoireumfangs

16 Vgl. Haimo 1995.

17 Vgl. Caplin 1998.

18 Vgl. Polth 2000a.

19 Vgl. Hepokoski/Darcy 2006. Kritisch zu Umfang und Zusammensetzung des Korpus vgl. Wingfield 2008, 141 f.

bleibt das Kernproblem unverändert, dass jede der genannten Studien lediglich auf qualitative Begrifflichkeiten (›häufig‹, ›oft‹, ›zumeist‹, ›selten‹ etc.) zurückgreift, um Häufigkeiten von Formoptionen zu erfassen. Ihnen allen ist der Gebrauch von ›intuitiver Statistik‹<sup>20</sup> gemeinsam; es scheint, als ließen sich die musikalischen Eigenschaften eines Korpus hinreichend gut mit quasi-statistischen Angaben charakterisieren. Während solche intuitiv-statistischen Charakterisierungen ohne Zweifel einen heuristischen Nutzen haben, wirft die beschriebene Praxis eine Reihe von Problemen auf, die im vorliegenden Zusammenhang reflektiert werden müssen.

Zunächst ist in der herkömmlichen musikanalytischen Repertoireforschung zumeist ein Mangel an Transparenz zu diagnostizieren, denn nur selten legen Musikforscher die Kriterien offen, nach denen sie bei ihrer intuitiv-statistischen Analyse verfahren sind. Robert Gjerdingen bringt das Problem wie folgt auf den Punkt:

Music scholars bandy about words like ›typical‹, ›characteristic‹, or ›standard‹ with the open confidence of embezzlers who, knowing that they alone keep the books, cannot be imagined being called to account.<sup>21</sup>

In *Music in the Galant Style* (2007) reiht sich Gjerdingen selbst unter die von ihm kritisierten Forscher ein, da er zwar eine – an sich sehr informative – Übergangstabelle für galante Stimmführungsmodelle anbietet, die dabei zu Grunde liegende Stichprobe, von der behauptet wird, sie sei klein aber repräsentativ, jedoch nicht näher charakterisiert.<sup>22</sup>

## 2.2 Verfügbarkeitsheuristik und Bestätigungsfehler

Die beschriebene Art der intuitiven Urteilsbildung leidet nicht nur an einem Mangel an Präzision und geht mit methodischer Intransparenz einher; sie ist auch für eine Reihe von Verzerrungen empfänglich, wie sie in der kognitiven Psychologie seit Langem bekannt sind.<sup>23</sup> Dazu zählen vor allem Phänomene wie die ›Verfügbarkeitsheuristik‹ und der ›Bestätigungsfehler‹.

Die Verfügbarkeitsheuristik zeichnet sich dadurch aus, dass Menschen bevorzugt auf Informationen zurückgreifen, die ihnen schnell ›in den Sinn‹ kommen, um Urteile über Häufigkeiten oder Wahrscheinlichkeiten zu treffen (sog. ›ease of retrieval‹<sup>24</sup>). Diese den Zeitaufwand reduzierende Heuristik gelangt bei Urteilsbildungen im Alltag ebenso zur Anwendung wie bei solchen im wissenschaftlichen Kontext, sofern nicht Maßnahmen getroffen werden, um dem entgegenzuwirken. Wer über den Ratz'schen »Satz«<sup>25</sup> nachdenkt, dem wird nolens volens als erstes der Anfang von Beethovens Sonate f-Moll

20 Vgl. Fischer 1970, 104; Kahneman 2011, 5.

21 Gjerdingen 1991, 127. Ein gutes Beispiel ist etwa die Schichtung von Mittelzäsur-Optionen, die nicht explizit statistisch untermauert wird, in Hepokoski/Darcy 2006, 23–50.

22 Vgl. Gjerdingen 2007, 372. Ähnliches gilt für die harmonische Übergangstabelle anhand einer nicht näher spezifizierten Stichprobe von Barockmusik in Huron 2006, 250f.

23 Vgl. Kahneman 2011.

24 Vgl. ebd.

25 Ratz 1973, 21 ff.

op. 2/1 in den Sinn kommen; sie oder er wird dann möglicherweise bestimmte Eigenschaften wie die motivische Fragmentierung in der Fortsetzungsphrase, den dynamisch-vorwärtstreibenden Charakter, ein offenes halbschlüssiges Ende und ähnliches für typischer halten, als dies möglicherweise in der kompositorischen Praxis der Zeit der Fall war. Ob diese Konklusion tatsächlich wahr oder falsch ist, spielt im vorliegenden Zusammenhang keine Rolle; entscheidend ist hier die Problematik des methodischen Wegs, der zu besagter Schlussfolgerung führt.

Der sogenannte Bestätigungsfehler (›confirmation bias‹) bezeichnet eine ›Positive-Test-Strategie‹, bei der – bewusst oder unbewusst – ausschließlich auf Beobachtungen fokussiert wird, die eine Theorie bestätigen, ohne dass systematisch nach falsifizierenden Befunden gesucht würde.<sup>26</sup> Auch hier lassen sich aus der Musikforschung problemlos Beispiele anführen. Hier ein fast willkürlich gewähltes: Die Hypothese, dass sich Haydn bevorzugt in Sätzen mit monothematischer Exposition für eine substantielle Rekomposition der Reprise entscheidet, da ein Wiederaufgreifen des Seitensatzes in der Grundtonart eine unerwünscht redundante Struktur zur Folge hätte, belegt Ethan Haimo durch den Verweis auf wenige positive Fälle.<sup>27</sup> Die Vielzahl derjenigen Fälle, die dieser Hypothese zuwiderlaufen und alternative Erklärungsmuster für Haydns rekomponierte Reprisen nahe legen, wird völlig ausgeblendet.<sup>28</sup>

Schließlich sei den Statistikskeptiker\*innen mit dem Argument begegnet, dass jeder Fall von Korpusanalyse, sei sie intuitiv, exemplarisch oder statistisch, implizit auf Annahmen über ein zu Grunde liegendes Modell oder einen statistischen Prozess beruht.<sup>29</sup> Diese Prämisse wird jedoch in intuitiven Verfahren meist weder benannt, noch reflektiert, und mag in manchen Fällen aufgrund der mangelnden expliziten Charakterisierung auf inkonsistente Weise in den Theoriebildungsprozess eingehen. Pistons vielzitierte »Table of usual root progressions« (siehe unten, Tab. 1) ist in diesem Zusammenhang nicht nur ein Beispiel für intuitive Statistik, sondern gleichzeitig sogar ein explizit gemachtes einfaches Modell, nämlich ein handgefertigtes Markov-Modell, welches Übergangshäufigkeiten von harmonischen Stufen beschreibt.<sup>30</sup>

### 2.3 Das Problem verzerrter und zu geringer Stichproben

Ein weiteres gravierendes und in der Regel unterschätztes Problem betrifft die Frage der Repertoirebildung in der Musikforschung, denn der aus den empirischen Wissenschaften bekannte Begriff der ›Stichprobe‹ lässt sich hier selten umstandslos anwenden.<sup>31</sup> Unter

26 Vgl. Kahneman 2011.

27 Vgl. Haimo 1988 und 1995.

28 Vgl. Neuwirth 2010.

29 Eine statistische Charakterisierung, ob intuitiv oder quantitativ, ist ohne Annahme eines zu Grunde liegenden Strukturmodells nicht denkbar. Statistik fällt keineswegs mit dem bloßen Zählen von Objekten zusammen – und selbst dieses stellt ein einfaches Modell dar; vgl. Rohrmeier 2013.

30 Vgl. ebd.

31 Nach Klärung der Forschungsfrage und Operationalisierung der zentralen Begrifflichkeiten (d. h. Überführung theoretischer Konzepte in empirisch beobachtbare Variablen) ist ein nächster Schritt die Erhebung der Daten als Stichprobe (*sampling*).



einer Stichprobe versteht man dort bekanntlich eine Teilmenge derjenigen Population, die man untersuchen möchte. Der Begriff ›Stichprobe‹ impliziert also den Anspruch auf Generalisierung, der wiederum die Anwendbarkeit von Inferenzstatistik als Methode des Hypothesentests garantiert. Allerdings ist dieser Anspruch im musikwissenschaftlichen Fall häufig gar nicht gegeben, denn untersucht man etwa Mozarts Klaviersonaten, so tut man dies in der Regel nicht mit der Absicht, generalisierende Aussagen über Mozarts allgemeinen Kompositionsstil treffen zu wollen. Für eine Untersuchung dieser Fragestellung wäre es ratsam, ein breiteres Spektrum an Gattungen zu Grunde zu legen, nicht einen Ausschnitt aus dem Œuvre, der möglicherweise aufgrund von Vorselektion zu systematischen Verzerrungen führen könnte (z. B. durch Beobachtung von Eigenschaften, die qua Gattungszugehörigkeit nur auf die Sonaten zutreffen).

Die Untersuchungen eines bestimmten Repertoires (z. B. Mozarts Klaviersonaten) zielen also auf die Charakterisierung genau dieses Repertoires. Stichprobe und Population fallen in eins. Doch selbst wenn man eine umfangreichere Auswahl aus dem Mozart'schen Œuvre treffen würde, bliebe die Unterscheidung zwischen Stichprobe und (unbekannter) Population fragwürdig, da das Mozart'sche Schaffen uns ja im Wesentlichen in seiner Gesamtheit vorliegt und lediglich ein größerer Aufwand vonnöten wäre, um es völlig zu erschließen. Um den Unterschied zwischen dem historischen und sozialwissenschaftlichen Datentypus deutlich zu machen, führt David Huron hier die hilfreiche Differenzierung zwischen »finite data« und »unbounded data«<sup>32</sup> ein, wobei Ersteres den Datentypus bezeichnet, mit dem (Musik-)Historiker\*innen operieren. Dass es dabei nicht lediglich um terminologische Spitzfindigkeiten geht, sondern ein sachlich bedeutsamer Kern berührt wird, tritt zutage, wenn man sich in Erinnerung ruft, dass etwa Hypothesen- bzw. Signifikanztests und die damit einhergehenden statistischen Verfahren wesentlich darauf beruhen, dass von einer Stichprobe auf die Beschaffenheit der Grundgesamtheit probabilistisch rückgeschlossen wird.<sup>33</sup>

Erhebt man allerdings den Anspruch auf Verallgemeinerbarkeit, so ist darauf zu achten, dass die Erhebung der Stichprobe randomisiert erfolgt. Mit anderen Worten: Jede relevante Beobachtungseinheit muss prinzipiell mit der gleichen Wahrscheinlichkeit Eingang in die Stichprobe finden können, um so die Repräsentativität der Stichprobe zu gewährleisten. Andernfalls liegt eine nicht-randomisierte Stichprobe vor, die potenziell voreingenommen sein (›biased sample‹) und u. a. auch den oben genannten Problemen (z. B. Verfügbarkeitsheuristik) unterliegen kann.

Neben der Beschaffenheit der Stichprobe spielt auch ihr Umfang eine wichtige Rolle. Grundsätzlich gilt, dass kleinere Stichproben anfälliger für systematische Fehler sind als größere; andererseits ist aber ab einem mathematisch bestimmbar Maximalumfang keine Fehlerminimierung mehr zu erwarten.

In diesem Zusammenhang ist auch das Problem der Überanpassung (›overfitting‹) von zentraler Bedeutung. Will man Aussagen (bzw. Regeln) überprüfen, die charakteristische Eigenschaften eines Korpus, z. B. von Mozarts Klaviersonaten oder Bachs Chorälen, beschreiben und eben nicht auf Zufallsartefakte zurückgehen, ist entscheidend, dass

32 Huron 2013, 4.

33 Vgl. Huron 2013.

der Satz von Stücken, die zur Überprüfung benutzt werden, nicht mit dem Satz von Stücken, aus dem die Aussagen gewonnen wurden, überlappt. Das Prinzip der Trennung von ›training set‹ und ›test set‹ ist ein fundamentaler Grundsatz in der Computermodellierung.<sup>34</sup> Wird dieses Prinzip nicht beachtet, laufen die beobachteten Resultate Gefahr, nicht generelle Eigenschaften des Korpus abzubilden, sondern reine Artefakte der gewählten Stichprobe.<sup>35</sup> Dies trifft insbesondere auf detaillierte Regeln mit wenig Anwendungsfällen und kleine Korpora zu. Das Problem der Überanpassung betrifft auch die gängige musikwissenschaftliche Praxis und erfordert eine grundlegende Methodenreflexion und -revision.<sup>36</sup>

Wie werden in der Musikforschung Korpora erstellt? Nach welchen Grundsätzen erfolgt die Erhebung der zu untersuchenden Datenmenge, und inwieweit orientiert sich Musikforschung an den in den empirischen Wissenschaften etablierten Maßstäben? Korpusforschung ist entschieden von der Verfügbarkeit von Daten abhängig. Dies betrifft auf der untersten Ebene das Vorhandensein von Musik in symbolisch kodierter Form. Wer etwa den allgemeinen musikalischen Stil von europäischer Kunstmusik in den Jahren 1770–1790 untersuchen möchte und sich dabei nur auf die Werke Haydns und Mozarts stützt (bzw. die vielen sie umgebenden Komponisten ausklammert), kann kaum den Anspruch einer repräsentativen Stichprobe erheben. Und auch bei den nicht-kanonischen Komponisten ist, sofern keine Gesamtausgaben verfügbar sind, Vorsicht geboten, da deren Werke vielfach bereits vorselektiert wurden, und dabei das selten offengelegte Auswahlkriterium vor allem in den älteren Denkmälerausgaben darin lag, eine bestimmte historiographische Position zu untermauern und/oder die Progressivität eines bestimmten Komponisten aufzuzeigen (etwa die Viersätzigkeit der Sinfonien von Johann Stamitz in den von Hugo Riemann herausgegebenen *Denkmälern der Tonkunst in Bayern*).

Die Verfügbarkeit musikalischer Daten in symbolischer Form ist allerdings noch nicht hinreichend für Korpusforschung. Um von ›Korpusforschung‹ im engeren Sinn sprechen zu können, müssen die Daten in einem digitalen (maschinenlesbaren) Format vorliegen. Ein großes Desiderat liegt allerdings darin, dass in nur vergleichsweise wenigen Fällen brauchbare Daten vorhanden sind, und selbst bei diesen wenigen Fällen können systematisch verzerrende Störfaktoren wirksam sein. So warnt etwa David Huron generell mit Bezug auf linguistische Korpusstudien vor einem »cultural bias« und insbesondere einem »strong bias towards English-language material«<sup>37</sup>.

Zu den brauchbaren und häufig verwendeten Datensätzen zählt der von Helmut Schaffrath (1995) kompilierte Essener Volkslied-Korpus, der sich zur Analyse von Prinzipien der Melodiebildung eignet und zu diesem Zweck ausgiebig genutzt wurde.<sup>38</sup> Ein weiterer Korpus, der sich zur Analyse von Harmonik anbietet, ist der *Yale-Classical Archives Corpus* (YCAC) von Christopher White und Ian Quinn (2016), der auf der umfangreichen MIDI-File-Sammlung auf [classicalarchives.com](http://classicalarchives.com) basiert und ein breites Spektrum

34 Vgl. z. B. MacKay 2003.

35 Vgl. Rohrmeier/Graepel 2012; Rohrmeier 2013.

36 Vgl. dazu Neuwirth 2010.

37 Huron 2013, 5.

38 Vgl. etwa Huron 2006.

an Stilen westlicher Musik von der Renaissance bis ins späte 20. Jahrhundert umfasst (mehr als 14.000 Werke).<sup>39</sup> Durch das Übergewicht kanonischer (deutscher) Komponisten (Bach ist mit mehr als 2500 Einträgen vertreten, Mozart mit knapp 1500, Haydn und Beethoven mit je knapp 1000) ist dieser Korpus allerdings nicht in der Lage, einen allgemeinen europäischen Stil repräsentativ abzubilden. Weitere Einschränkungen der Brauchbarkeit des Korpus liegen darin, dass die Files eine nicht unerhebliche Fehlerrate (von 8%) in der Tonhöhenkodierung aufweisen und die Daten keine verlässlichen metrischen Informationen enthalten. Für musikanalytische Nutzer\*innen gestaltet sich der Vergleich der MIDI-Daten mit herkömmlichen Partituren schwierig.

### 3. Einige Fallbeispiele für musikalische Korpusforschung

Ungeachtet der beschriebenen methodischen Probleme, um deren Lösung sich die Musikforschung in den kommenden Jahren bemühen müssen, wird das enorme Potenzial, das der Umgang mit Korpora für die Musikforschung in sich birgt, an einer Reihe von Studien deutlich, die schwerpunktmäßig in den letzten Jahren erschienen sind.<sup>40</sup> Im Folgenden wird zunächst ein (notwendig verkürzter) Überblick über wenige repräsentative Studien zur korpusbasierten Analyse verschiedener musiktheoretischer Gegenstandsbereiche (Harmonik, Kontrapunkt, Melodik, Rhythmik/Metrik) gegeben. Ausgehend davon soll diskutiert werden, inwieweit diese Studien zur Präzisierung musiktheoretischer Annahmen beitragen können und worin die Herausforderungen zukünftiger korpusbasierter Forschung bestehen.

#### 3.1 Harmonik

Welche Akkorde folgen bevorzugt aufeinander, welche selten oder nie? Welche Akkorde stehen in einem eher asymmetrischen Verhältnis zueinander, welche in einem symmetrischen?<sup>41</sup> Wie ändert sich das Akkordverhalten in Abhängigkeit vom jeweils vorliegenden stilistischen Kontext bzw. umgekehrt, inwieweit wird ein Stil über das Akkordverhalten definiert? Welche Rolle spielt Stimmführung, insbesondere die Bass-Stufenprogression für das Akkordverhalten?

Eine der ersten Arbeiten, die versucht hat, eine genauere, wenn auch pädagogisch intendierte Spezifizierung von Akkordprogressionen (im Dur-Kontext) zu leisten, ist Walter Pistons Lehrbuch *Harmony* (1941). Die Kritikpunkte an der von Piston dargebotenen Tabelle (Tab. 1) liegen auf der Hand. Zum einen liefert Piston keine quantitativ präzisen Angaben über das Akkordverhalten, sondern lediglich tentative, qualitative Charakterisierungen. Zum anderen bleibt völlig unklar, von welchem Korpus und mittels welcher Methode diese Charakterisierungen abgeleitet wurden. Piston scheint sich hier dessen zu bedienen, was oben unter den Begriff der ›intuitiven Statistik‹ subsumiert wurde.

39 Vgl. White/Quinn 2016; White 2014.

40 Einen lesenswerten Überblick über die ältere Literatur bietet Nettheim 1997.

41 Symmetrische Akkordfolgen lassen sich leicht umkehren, asymmetrische dagegen nicht. Die Variable ›Symmetrie‹ lässt sich somit als geeignetes Maß der Direktionalität eines harmonischen Systems verstehen.

Table of usual root progressions			
	is followed by	sometimes by	less often by
I	IV or V,	VI,	II or III,
II	V,	IV, VI,	I, III,
III	VI,	IV,	I, II or V,
IV	V,	I or II,	III or VI,
V	I,	VI or IV,	III or II,
VI	II or V,	III, IV,	I
VII	III,	I	

Tabelle 1:  
Walter Piston, *Harmony*, »Table of usual root progressions«<sup>42</sup>

Nun ist diese Kritik nur bedingt angebracht, denn es handelt sich bei Pistons Übergangstabelle um etwas, das *en passant* im Kontext eines Lehrbuchs (also nicht in einer streng-wissenschaftlichen Arbeit) als heuristischer Ausgangspunkt gesetzt und im weiteren Verlauf des Buches weiter ausdifferenziert wird. Viel wichtiger sind demzufolge die Entwicklungen, die dadurch angeregt wurden. Im Anschluss an Piston haben einige Theoretiker und Kognitionsforscher versucht, die empirische Untersuchung harmonischer Tonalität auf eine fundierte Basis zu stellen.<sup>43</sup> Trotz dieser Initiativen gibt es allerdings gegenwärtig nur wenige handannotierte harmonische Analysen, die maschinenlesbar und quantitativ verwertbar sind.<sup>44</sup>

Der 2009 von David Temperley erstellte Kostka-Payne-Korpus, der auf dem bekannten Lehrbuch *Tonal Harmony* von Stephan Kostka und Dorothy Payne (1995) basiert, enthält lediglich Ausschnitte aus Werken von (u. a.) Bach, Haydn, Mozart, Beethoven, Schubert, Schumann, Chopin, Brahms und Tschaiakowsky.<sup>45</sup> Dieser Umstand nährt den Verdacht der Stichprobenverzerrung, da die Auswahl der Exzerpte nicht randomisiert erfolgt ist, sondern – im Kontext eines Lehrbuchs – möglicherweise durch präempirische theoretische Erwägungen bedingt war. Es ist also nicht auszuschließen, dass die Auswahl der Exzerpte in *Tonal Harmony* primär dadurch motiviert war, einen bestimmten theoretischen Standpunkt zu untermauern.

Trotz dieses Vorbehalts liefert der Kostka-Payne-Korpus einen diskussionswürdigen Modellfall für die Möglichkeiten der empirischen Erforschung von Harmonik. Berechnet werden die Anzahl der Akkordinstanzen (896) und der sowohl numerische als auch zeitlich-proportionale Anteil bestimmter Akkordtypen an der Gesamtmenge. Aus der Übergangstabelle (Tab. 2) können etwa die häufigsten Akkordpaare abgeleitet werden (V-I, I-V, ii-V, I-IV, I-ii, vi-ii, IV-I, vii°-I, I-vii°, I-vi; die Fortschreitung IV-V ist überraschenderweise nicht dabei, was der musiktheoretischen Intuition widersprechen dürfte). Darüber hinaus zeigen sich interessante Einzelbefunde, die weitestgehend den Beschreibungen tonaler Syntax in heutigen Lehrbüchern entsprechen. So schreitet etwa die V. Stufe häufiger zur I. Stufe als umgekehrt, die beiden Akkorde stehen also zueinander in einem asymmetrischen Verhältnis. Die V. Stufe wird weitaus häufiger durch die ii. Stufe vorbe-

42 Piston 1941, 17.

43 Vgl. Budge 1943. De Clercq/Temperley 2011 untersuchen Harmonik in Rock/Pop.

44 Von handannotierten Korpora ist die Rede, wenn menschliche Annotator\*innen nach bestimmten, vorab festgelegten Richtlinien den Datensatz erzeugen. Dies ist zu unterscheiden von automatischen Verfahren, bei denen der Datensatz durch ein Computerprogramm kriert wird.

45 Vgl. Temperley 2009 sowie Temperley 2011, 148 f.

reitet als durch die IV. Stufe.<sup>46</sup> Folgen die IV. und die ii. Stufe aufeinander, so deutlich häufiger in der Reihenfolge IV-ii als umgekehrt. Vermeintlich irreguläre Folgen wie V-IV treten deutlich seltener auf als IV-V, und dies vermutlich hauptsächlich im Kontext von Bass-Skalenharmonisierungen in Umkehrung (absteigend: I-V<sup>6</sup>-IV<sup>6</sup>-V<sup>7</sup> bzw. aufsteigend: V<sup>7</sup>-IV<sup>6</sup>-V<sup>6</sup>-I). Da Temperley in der statistischen Auswertung nur den Grundtonbezug der Akkorde berücksichtigt, den Bassbezug (bzw. die Akkordumkehrung) hingegen ausklammert, leidet, wie in letztgenanntem Fall, die Genauigkeit der Prädiktion von Akkordfolgen.<sup>47</sup> Stimmführungszusammenhänge werden sichtbarer, wenn das kontextuelle Zeitfenster von *bigrams* auf *tri-* oder *four-grams*<sup>48</sup> ausgedehnt wird (wobei damit zugleich die Anzahl wiederholter Akkordmuster sinkt).

Chromatic Root Transitions as Proportion of Count for Consequent Chord												
Cons Ant	I	♭II	II	♭III	III	IV	♯IV	V	♭VI	VI	♭VII	VII
I	0.000	0.027	0.121	0.004	0.016	0.176	0.008	0.453	0.043	0.066	0.012	0.074
♭II	0.200	0.000	0.533	0.000	0.000	0.000	0.067	0.133	0.000	0.000	0.000	0.067
II	0.222	0.030	0.000	0.010	0.040	0.010	0.071	0.455	0.020	0.081	0.000	0.061
♭III	0.100	0.100	0.000	0.000	0.000	0.000	0.000	0.400	0.400	0.000	0.000	0.000
III	0.053	0.000	0.105	0.000	0.000	0.368	0.000	0.053	0.000	0.368	0.000	0.053
IV	0.471	0.029	0.147	0.000	0.059	0.000	0.044	0.162	0.000	0.015	0.015	0.059
♯IV	0.438	0.000	0.000	0.000	0.000	0.000	0.000	0.562	0.000	0.000	0.000	0.000
V	0.848	0.000	0.041	0.005	0.010	0.020	0.000	0.000	0.036	0.030	0.000	0.010
♭VI	0.192	0.077	0.308	0.000	0.038	0.115	0.000	0.077	0.000	0.155	0.077	0.000
VI	0.093	0.047	0.651	0.000	0.023	0.093	0.047	0.023	0.000	0.000	0.000	0.023
♭VII	0.000	0.000	0.000	0.833	0.000	0.000	0.000	0.167	0.000	0.000	0.000	0.000
VII	0.818	0.000	0.000	0.000	0.091	0.000	0.030	0.030	0.030	0.000	0.000	0.000

Tabelle 2: David Temperley, »A Statistical Analysis of Tonal Harmony«, Übergangswahrscheinlichkeiten im Kostka-Payne-Korpus<sup>49</sup>

Die Außerachtlassung von Akkordumkehrungen ist nicht unproblematisch. Sie suggeriert, Akkordfolgen könnten unabhängig von Stimmführung untersucht werden (ganz im Sinne von Dahlhaus' Diktum »Eine Stufenfolge IV-V-I oder eine Funktionenfolge wie

46 Dies entspricht der musiktheoretischen Intuition, dass der ii<sup>6</sup>-Akkord wohl deutlich häufiger als der Akkord der IV. Stufe als Prädominante im kadenzierenden Kontext Verwendung findet.

47 Versteht man allerdings die grundtonbasierte Funktionalität nur als Approximation, die dem Sparsamkeitskriterium genügt, so gewinnt dieser Ansatz an Legitimität. In diese Richtung argumentiert David Temperley: »An immediate problem with this model is that its predictions do not always hold: sometimes dominants move to tonics, for example, as in a plagal cadence. Admittedly, such exceptions show that functional harmonic theory is imperfect as a model of tonal harmony; but they do not show that it is useless. A theory whose predictions hold true most of the time can still be of great value; we use such theories all the time in our daily lives. Imperfect though it may be, functional harmonic theory represents a powerful and valid generalization about tonal harmony, better than many conceivable alternatives – for example, a theory that posited that chords are chosen at random without regard for the previous chord, or that dominants move to tonics and dominants move to dominants. On this basis, I would argue, we are justified in positing functional harmony as part of the knowledge that common-practice composers brought to bear in their compositional process.« (Temperley 2011, 148) Vgl. außerdem Jacoby/Tishby/Tymoczko 2015, 2–9.

48 Zum Begriff des *n-gram* siehe unten, Abschnitt 3.2.

49 Temperley 2009.

S-D-T besagt nichts über die Stimmführung, die absurd sein kann, ohne die ‚harmonische Logik‘ aufzuheben<sup>50</sup>). Tatsächlich ist längst unumstritten, dass es die Einbettung von Klängen in ein Stimmführungsgefüge ist, die über die Funktionalität der Einzelklänge entscheidet:<sup>51</sup> IV<sup>6</sup> etwa kann ebenso wie vi oder vii<sup>06</sup>/V ein trugschlüssiger Akkord sein (ii<sup>6</sup>-V<sup>4</sup>-V<sup>7</sup>-IV<sup>6</sup>) oder unterschiedliche prolongierende Funktionen erfüllen (Tonikaprolongation: z. B. I-IV<sup>6</sup>-I<sup>6</sup>, z. B. mit der Oberstimme 3-4-5; oder Dominantprolongation: z. B. I-V<sup>6</sup>-IV<sup>6</sup>-V<sup>7</sup>). In den genannten Kontexten lässt sich IV<sup>6</sup> nicht ohne Beeinträchtigung des syntaktischen Zusammenhangs durch einen grundstelligen Akkord der IV. Stufe ersetzen. Die Ersetzung durch einen Akkord der vi. Stufe ist allerdings denkbar (etwa in einer I-V<sup>6</sup>-IV<sup>6</sup>(bzw. vi)-I<sup>6</sup>-Progression), und die Verwendung unterschiedlicher Akkordsymbole aufgrund unterschiedlicher Grundtöne maskiert diese funktionale Äquivalenz.

Die hier vorgebrachten Kritikpunkte dürfen nicht als Kritik gegen Korpusforschung per se verstanden werden. Tatsächlich lässt sich die Prädiktionskraft grundton- und bassonbezogener Theorien sogar quantitativ untersuchen.<sup>52</sup> Doch auch die Frage, inwieweit in die (harmonische) Annotation Kontextinformationen (wie z. B. Stimmführungsaspekte und metrische Informationen), musiktheoretische Expertise und musikalische Intuition einfließen, bleibt, sofern dies nur konsistent geschieht, jedem/r Forscher\*in selbst überlassen. (So ist durchaus denkbar, dass die harmonische Annotation Klänge nicht nach ihrem Grundtonbezug klassifiziert, sondern nach ihrer Stimmführungsfunktionalität, die wesentlich von der strukturellen Bassmelodik herrührt.) Dass dabei ein gewisses Vorverständnis in Anschlag gebracht wird, sollte auch nicht zum Vorwurf der Zirkularität Anlass geben, denn viele Aspekte bleiben davon unberührt, so etwa die Häufigkeit bestimmter harmonischer Konstellationen.

Vergleicht man die Ergebnisse der Auswertung des Kostka-Payne-Korpus mit dem Mozart-Klaversonaten-Korpus von Dmitri Tymoczko<sup>53</sup>, so zeigen sich viele Übereinstimmungen. Allerdings lassen sich die Unterschiede (z. B. das deutlich häufigere Auftreten der Progressionen ii-V und IV-V bei Tymoczko) zum einen dadurch erklären, dass Temperleys Korpus eine größere stilistische Bandbreite abdeckt, zugleich aber wesentlich kleiner ausfällt (896 vs. ca. 6.500 Akkordpaare). Zum anderen fehlen in der Übergangstabelle bei Tymoczko ca. 23 % der im Mozart-Korpus vorhandenen Akkordfolgen (vor allem Sequenzprogressionen).

Einer gewissen Beliebtheit erfreuen sich neben Mozarts Klaversonaten auch Bach-Choräle – nicht zufällig, handelt es sich doch hier um ein Repertoire, das gerne im Tonsatzunterricht verwendet wird und zu dem folglich unterschiedlichste Regelsysteme existieren, anhand derer Studierende Stilkopien anfertigen.<sup>54</sup> Auch hier stellt sich die Frage, ob die von menschlichen Analytiker\*innen entwickelten Regelsysteme einer

50 Dahlhaus 1968, 59.

51 Vgl. z. B. Salzer 1952; Gauldin 1988; Aldwell/Schachter 1989; Polth 2000b.

52 Vgl. Jacoby/Tishby/Tymoczko 2015.

53 Tymoczko 2010. Der Mozart-Sonaten-Korpus von Dmitri Tymoczko, an dem eine Vielzahl von US-amerikanischen Musiktheoretiker\*innen mitwirkte, wurde selbst bislang nicht offiziell publiziert, sondern lediglich eine Auswahl an Auswertungen.

54 Vgl. Salzer/Schachter 1969; Knipphals/Möller 1995; Daniel 2000; Kaiser 2002; Kühn 2006, 157–163.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	C.E.G	D.G.B	C.F.A	C.E.A	D.F.A	D.F.G.B	D.F.#.A	E.G.B	C.D.F.#.A	E.G.#.B	C.D.F.A	C.E.G.A	D.F.B	C.#.E.A	C.D.G
1	6,37	5,79	4,13	1,68	1,15	0,84	0,48	0,29	0,54	0,15	0,88	0,24	0,37	0,11	0,63
2	8,98	2,26	0,63	1,28	0,51	0,51	0,56	0,55	0,25	0,34	0,38	0,38	0,01	0,04	0,04
3	3,04	1,22	0,70	0,26	0,56	0,50	0,14	0,23	0,11	0,05	0,11	0,06	0,40	0,02	0,07
4	0,92	1,38	0,56	0,85	0,63	0,16	0,34	0,53	0,44	0,29	0,43	0,07	0,12	0,05	0,01
5	D.F.A	0,99	0,81	0,09	0,50	0,43	0,44	0,26	0,04	0,01	0,04	0,04	0,02	0,02	0,04
6	D.F.G.B	2,56	0,09	0,07	0,15	0,04	0,08	0,04	0,11	0,06	0,01	0,01	0,01	0,01	0,01
7	D.F.#.A	22	0,11	1,69	0,04	0,04	0,04	0,11	0,02	0,03	0,01	0,33	0,09	0,02	0,04
8	E.G.B	28	0,35	0,18	0,63	0,42	0,11	0,09	0,11	0,02	0,03	0,01	0,01	0,01	0,01
9	C.D.F.#.A	25	0,05	1,04	0,06	0,01	0,04	0,07	0,01	0,09	0,01	0,01	0,01	0,04	0,04
10	E.G.#.B	18	0,02	1,06	0,03	0,04	0,04	0,02	0,02	0,05	0,01	0,01	0,01	0,04	0,04
11	C.D.F.A	18	1,12	0,02	0,01	0,04	0,47	0,04	0,09	0,02	0,01	0,01	0,04	0,04	0,04
12	C.E.G.A	19	0,04	0,08	0,03	0,12	0,04	0,04	0,04	0,02	0,01	0,01	0,04	0,04	0,04
13	D.F.B	23	0,80	0,02	0,03	0,04	0,01	0,04	0,03	0,01	0,01	0,01	0,04	0,04	0,04
14	C.#.E.A	25	0,08	0,04	0,01	0,12	0,46	0,03	0,03	0,01	0,01	0,01	0,04	0,10	0,10
15	C.D.G	10	0,67	0,01	0,01	0,19	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01

Total numbers of (non-zero) transitions to all pc sets found for each column and row

statistischen Analyse größerer Datenmengen standhalten.<sup>55</sup> Von den zahlreichen quantitativen Studien, die inzwischen zu dieser Fragestellung vorgelegt wurden, soll hier nur die Studie von Martin Rohrmeier und Ian Cross (2008) kurz angesprochen werden (Tab. 3): Anders als im Kostka-Payne- und Mozart-Korpus ist die Wahrscheinlichkeit eines Übergangs von der ii. in die I. vs. V. Stufe ungefähr gleich wahrscheinlich. Darüber hinaus bestätigt die Auswertung für Übergänge zwischen Pitch-class-sets den Eindruck tonaler Direktionalität und macht ferner auf Unterschiede zwischen Dur und Moll aufmerksam: In Moll geht die iv. Stufe ca. sechsmal so häufig in die V. Stufe wie umgekehrt und die i. Stufe in die VI. Stufe dreimal so häufig wie umgekehrt.<sup>56</sup>

links: Tabelle 3:  
Martin Rohrmeier / Ian Cross, »Statistical Properties of Harmony in Bach's Chorales«, Übergangswahrscheinlichkeit von Akkorden (Pitch-class-sets) für Durtonarten in Bach-Chorälen<sup>57</sup>

### 3.2 Kontrapunkt

Im Bereich der korpusgestützten Erforschung von Kontrapunkt lassen sich gegenwärtig zwei einander ergänzende Initiativen ausmachen: der ELVIS-Ansatz von Peter Schubert und Julie Cumming (2015) sowie die schematheoretische Untersuchung von Satzmodellen.<sup>58</sup> An

55 Vgl. McHose 1947; Boyd 1999; Rohrmeier/Cross 2008; Fitsioris/Conklin 2008; Moraitis 2011. Weitere wichtige Gründe für die Bevorzugung von Bachs Chorälen für Korpusforschung liegen in ihrer digitalen Verfügbarkeit sowie in der klaren Stimmentrennung.  
56 Vgl. Rohrmeier/Cross 2008.  
57 Ebd., Appendix, Table 2.  
58 Zur Satzmodell-Forschung, insofern sie empirisch fundiert ist, vgl. u. a. Gjerdingen 1988 und 2007; Byros 2009.

dieser Stelle sei nur von erstgenanntem Projekt berichtet, das an der *McGill University* angesiedelt ist.

In der Auseinandersetzung mit Renaissance-Kontrapunkt hat sich ein spezifischer Ansatz herausgebildet, der auf das Erkennen von wiederholten kontrapunktischen Fügungen oder Klangschritten abzielt. Solche ›Module‹ sind als aufeinanderfolgende Simultanintervalle definiert und werden zuweilen mit sinntragendem sprachlichen Vokabular verglichen. Während kontrapunktische Module bislang ›per Hand‹ identifiziert und analysiert wurden<sup>59</sup>, ermöglicht das von Schubert und Cumming (und Mitarbeitern) entwickelte *Python*-gestützte<sup>60</sup> Framework ELVIS nicht nur die automatische Erkennung, sondern auch die grundsätzlich beliebige Ausdehnung der Anzahl untersuchter Simultanintervall-Folgen (*n-grams*), wobei auch die melodischen Schritte der tiefsten Stimme kodiert werden.<sup>61</sup> Doch obwohl der Berücksichtigung umfangreicher Intervallfolgen technisch keine Grenzen gesetzt sind, werden *trigrams* (bestehend aus drei Intervallen) aufgrund ihres optimalen Informationsgehalts bevorzugt; geht man darüber hinaus, so nimmt die Anzahl wiederholter Patterns deutlich ab. Eine Automatisierung des Analyseverfahrens ist hier besonders vorteilhaft, da die Intervallsequenzen in der Regel eine nur geringe kognitive Salienz aufweisen und demnach für menschliche Analytiker\*innen schwerer erkennbar sind.

In einem kürzlich publizierten Aufsatz untersuchten Schubert und Cumming zwölf Duette Orlando di Lassos mit Blick auf kontrapunktische Module. Von den insgesamt 1.036 *trigrams* treten nicht weniger als 322 nur ein einziges Mal auf. Wenig überraschend ist das stückübergreifend häufige Auftreten von Kadenzmodulen mit Sekund- bzw. Septimvorhalt wie  $\langle 7_1 6_2 8 \rangle$ .<sup>62</sup> Aussagekräftiger ist demgegenüber der Befund, dass weitere häufig vorkommende Module entweder stückübergreifend mit einer standardisierten Solmisation (wie etwa  $\langle 6_2 5_2 3 \rangle$ ) oder stückspezifisch mit bestimmten Modi (wie etwa  $\langle 5_1 3_3 5 \rangle$  mit dem 6. Modus oder  $\langle 8_3 3_1 5 \rangle$  mit dem 8. Modus) verknüpft sind.<sup>63</sup>

Geht man über dieses begrenzte Repertoire hinaus und unternimmt wie Christopher Antila und Julie Cumming den Versuch eines diachronen Stilvergleichs in einem Zeitraum von 1440–1585 (mit einer Lücke zwischen 1522–1539), so zeigen sich weitere interessante Befunde, die sich mit Hilfe von ELVIS gewinnen lassen.<sup>64</sup>

Die Autoren unterscheiden drei aufeinanderfolgende Perioden, die Ockeghem-Generation, die Josquin-Generation und die Palestrina-Generation, und wählen dabei für den jeweiligen Zeitstil repräsentative, großteils vierstimmige Werke verschiedener Komponisten (u. a. Ockeghem, Dufay, Isaac, Josquin, Palestrina, Rore) bzw. verschiedener

59 Die Musiktheorie blickt hier auf eine lange Geschichte zurück, die u. a. in Menke 2015 anschaulich zusammengefasst ist.

60 Zu *Python* und *Music21* siehe z. B. Cuthbert/Ariza 2010.

61 ELVIS steht für »Electronic Locator of Vertical Interval Successions«.

62 Groß geschriebene Zahlen bezeichnen Simultanintervalle; die tiefgestellten Zahlen denotieren die Bewegung der Unterstimme, die im vorliegenden Fall liegen bleibt und dann einen Ganzton nach unten schreitet.

63 Vgl. Schubert/Cumming 2015.

64 Vgl. Antila/Cumming 2014.



Gattungszugehörigkeit im Rahmen geistlicher und weltlicher Vokalmusik aus. In Übereinstimmung mit der oben formulierten Maßgabe wird konsequent zwischen Trainings- und Testset unterschieden.

Die musiktheoretische Erfahrung wird in dieser Studie insofern bestätigt, als gezeigt werden konnte, dass bestimmte Kadenzmodule mit Synkopensdissonanz stilübergreifend häufig auftreten (z. B.  $\langle 7_1 6_2 8 \rangle$ ) und dementsprechend auch in den Lasso-Duetten zu finden waren. Besonders informativ sind allerdings die Unterschiede zwischen den drei Korpora. Die Anzahl distinkter *trigrams* ist im Josquin-Korpus höher (17) als im Ockeghem- oder Palestrina-Korpus (jeweils elf). Allerdings ist der prozentuale Anteil wiederholter *trigrams* bei  $\langle$ Josquin $\rangle$  und  $\langle$ Palestrina $\rangle$  höher als bei  $\langle$ Ockeghem $\rangle$ . Zugleich nimmt die Anzahl stildifferenzierender *trigrams* von  $\langle$ Ockeghem $\rangle$  zu  $\langle$ Josquin $\rangle$  minimal zu (von fünf auf sechs), bei  $\langle$ Palestrina $\rangle$  aber stark ab (von sechs auf eins): Das Modul  $\langle 5_2 6_2 6 \rangle$  findet sich nur im Palestrina-Korpus. Folglich erweisen sich manche Module als generationenübergreifend gültig, andere wiederum sind eher zeitraumspezifisch. Die Unterschiede in der Verwendung *unterschiedlich vieler* Module bietet interessante Aufschlüsse über stilistische Differenzen von Komponisten und Zeiträumen.<sup>65</sup>

### 3.3 Melodik

Im Bereich der Melodik wurden schwerpunktmäßig die folgenden Aspekte in den Blick genommen: (1) Das sogenannte  $\langle$ post-skip reversal $\rangle$ ; (2) die bogenförmige Kontur<sup>66</sup>; (3) die statistischen Übergangswahrscheinlichkeiten zwischen melodischen Skalenstufen<sup>67</sup>. Im Folgenden soll exemplarisch der erstgenannte Themenkomplex diskutiert werden.

Einer der kaum hinterfragten musiktheoretischen Grundsätze im Bereich der Melodiebildung betrifft die als  $\langle$ post-skip reversal $\rangle$  bezeichnete Annahme, dass einem melodischen Sprung eine (schrittweise) Bewegung in Gegenrichtung folgt. In seinem *Lehrbuch der klassischen Vokalpolyphonie* (1935) hat etwa Knud Jeppesen unter Bezugnahme auf Palestrina zahlreiche Regeln für die Melodiebildung formuliert, darunter auch die Regel, dass »ein größerer Sprung abwärts häufig durch einen zweiten Sprung (selbstverständlich in entgegengesetzter Richtung) beantwortet [wird]; einem großen Sprung nach oben folgt dagegen am gebräuchlichsten eine stufenweise Bewegung abwärts.«<sup>68</sup> Leonard B. Meyer erklärt dieses Phänomen durch das gestaltpsychologisch fundierte Konzept des  $\langle$ gap-fill $\rangle$ : Eine durch den Sprung entstandene Lücke erfordert eine Auffüllung durch eine Umkehr der melodischen Bewegungsrichtung.<sup>69</sup>

Vor mehr als 15 Jahren hat sich das Forschungsteam um David Huron dieser Hypothese angenommen und sie einer empirischen Prüfung unterzogen.<sup>70</sup> Dabei wurde ne-

65 Generell ist anzumerken, dass die Beschränkung von ELVIS auf Musik der Renaissance nicht zwingend ist; denkbar wäre eine Nutzung zur präzisen Beschreibung von Stil und Stilwandel von Musik im Allgemeinen (sofern sie in symbolischem Datenformat vorliegt).

66 Vgl. z. B. Huron 1996. Eitan 1997 befasst sich mit der Gestaltung melodischer Höhepunkte.

67 Vgl. z. B. Huron 2006, 158f.; Temperley/De Clercq 2013.

68 Jeppesen 1935, 71.

69 Vgl. Meyer 1973, 145–157.

70 Für eine frühe Studie vgl. Watt 1924.

ben dem oben genannten Essener Volkslied-Korpus auch chinesische, südafrikanische und indianische Musik untersucht. Paul von Hippel wies anhand von Stichproben aus diesen diversen Stilen nach, dass das ›post-skip reversal‹ nicht eine Folge des ›gap-fill‹ ist, wie die musiktheoretische Intuition suggeriert, sondern sich besser durch das rein statistische Konzept ›Tendenz zur Mitte‹ (›regression to the mean‹) erklären lässt: Die großen Sprünge führen in der Regel an die Grenzen der verfügbaren Tessitura, weswegen eine nachfolgende Annäherung an den mittleren Bereich der Tessitura allein aus statistischen Gründen wahrscheinlich ist.<sup>71</sup> Ein statistisches Artefakt wird also gegenüber einer inhaltlichen Kausalerklärung bevorzugt.<sup>72</sup> Auf die Extremwerte der Verteilung (den Grenzbereich des Ambitus) folgen in der Regel Werte um den statistischen Durchschnitt (den mittleren Bereich des Ambitus) herum. Daraus folgt zweierlei: Erstens, eine Umkehrung der Bewegungsrichtung tritt auch dann ein, wenn die Grenze des verfügbaren Tonraums schrittweise erreicht wird; zweitens, unterhalb der Tonraumschwelle ist eine Umkehrung der Bewegungsrichtung nach einem Sprung deutlich weniger wahrscheinlich. Entscheidend ist folglich nicht die Größe des Intervalls (Schritt vs. Sprung), sondern seine Positionierung im Tonraum. Diese wichtige Nuance blieb in der Musikforschung traditioneller Prägung bislang unberücksichtigt und verweist nachdrücklich auf den Mehrwert von Korpusforschung.

### 3.4 Rhythmik/Metrik

Neben Harmonik, Melodik und Kontrapunkt gibt es, wenn auch in geringerem Maße, beachtenswerte Studien zur Rhythmik und Metrik in verschiedenen historischen Stilen. Dazu zählt etwa die Studie von John Ito (2014) zur metrischen Platzierung von Kadenzten, die Aussagen Heinrich Christoph Kochs auf den Prüfstand stellt, oder auch diverse Studien zu rhythmischen Aspekten populärer Musik: Die Studie von Anja Volk und Bas de Haas (2013) zu Aspekten der Synkopierung in Joplins Ragtimes, Nicole Biamontes Arbeit zu Formfunktionen metrischer Dissonanz in Rockmusik (2014) oder Hurons und Ommens Arbeit zur Synkopierung in amerikanischer Populärmusik zwischen 1890 und 1939 (2006). Generell ist in der gegenwärtigen musikalischen Korpusforschung ein starker Trend zu beobachten, Repertoires außerhalb der westlichen Kunstmusik in den Blick zu nehmen, der nicht zuletzt durch die Arbeiten von Temperley und De Clercq (2013) oder das *McGill University Billboard Project*<sup>73</sup> vorangetrieben wird.<sup>74</sup>

71 Von Hippel/Huron 2000.

72 Zum zu Grunde liegenden Problem der Entscheidungspsychologie allgemein vgl. Kahneman 2011, 166–184.

73 Vgl. McGill University, <http://ddmal.music.mcgill.ca/research/billboard> (30.6.2017).

74 Vgl. auch Burgoyne/Wild/Fujinaga 2013.

#### 4. Ausblick: Perspektiven und Desiderate

Trotz der nach wie vor bestehenden methodischen Herausforderungen gegenwärtiger Korpusforschung verspricht dieser Ansatz für die Musiktheorie eine Vielzahl an Vorteilen: Zunächst erlaubt er die quantitativ präzise Evaluation generalisierender Aussagen über musikalische Strukturen und mag dabei als Korrektiv und Komplement unserer analytischen Intuitionen fungieren, etwa indem bislang übersehene Muster und Regularitäten aufgezeigt werden. Korpusforschung zwingt Musiktheoretiker\*innen geradezu, ein höheres Maß an Präzision hinsichtlich ihrer Aussagen über musikalische Strukturen an den Tag zu legen, etwa schon aufgrund der Notwendigkeit, ein funktionsfähiges Annotationssystem zu entwickeln und die zu untersuchenden Variablen zu operationalisieren, d. h. so zu definieren, dass sie in Beobachtungssätze übersetzt werden können.

Außerdem führt Korpusforschung deutlich vor Augen, dass die Partiturform, die Musikanalytiker\*innen zumeist wie selbstverständlich zu Grunde legen, nur eine von vielen symbolischen Repräsentationsformen von Musik ist, die von *piano roll* über MIDI und *\*\*kern* reichen und jeweils unterschiedliche Analysemethoden erfordern. Doch auch Audiodaten (etwa von *performances*) können zum Analysegegenstand von digitaler Forschung gemacht werden.<sup>75</sup>

In den letzten Jahren ist ein steigender Trend zu beobachten, dass auch herkömmlich ausgebildete Musiktheoretiker\*innen vor allem in den USA Elemente von Korpusforschung in ihre musiktheoretische Arbeit integrieren.<sup>76</sup> Doch solange quantitative Methoden nicht genuiner Bestandteil der musikwissenschaftlichen bzw. -theoretischen Ausbildung sind, ist auch mit Blick auf eine sich zunehmend ausdifferenzierende und plurale Methodik eine Kooperation mit Expert\*innen im Bereich Statistik und Computerwissenschaft unumgänglich. Es ist zu erwarten, dass sich in den kommenden Jahren aus dem gegenwärtigen Methodenpluralismus ein fester Kern an empirischen Standardverfahren herauskristallisieren wird. Die methodischen Fortschritte müssen allerdings zwingend durch Reflexionen über fruchtbare Weisen der Interaktion von quantitativen und qualitativ-hermeneutischen Ansätzen ergänzt werden. In diesem Sinne schließt der Beitrag mit der begründeten Hoffnung, dass auch die deutschsprachige Musiktheorie sich dem reichhaltigen Potenzial der Korpusforschung nicht verschließen und das Momentum dieser jungen Bewegung gewinnbringend aufgreifen wird.

75 Einen informativen Überblick über Repräsentationsformen von Musik bietet Müller 2015.

76 Vgl. z. B. Biamonte 2014; Richards 2016.

## Literatur

- Aldwell, Edward / Carl Schachter (1989), *Harmony and Voice Leading*, 2. Auflage, Toronto: Jovanovich.
- Antila, Christopher / Julie Cumming (2014), »The VIS Framework. Analyzing Counterpoint in Large Datasets«, in: *Proceedings of the 15<sup>th</sup> International Society for Music Information Retrieval Conference, ISMIR 2014, Taipei, Taiwan, October 27–31*, hg. von Hsin-Min Wang, Yi-Hsuan Yang und Jin Ha Lee, 71–76. [http://www.terasoft.com.tw/conf/ismir2014/proceedings/T014\\_162\\_Paper.pdf](http://www.terasoft.com.tw/conf/ismir2014/proceedings/T014_162_Paper.pdf) (30.6.2017)
- Beran, Jan (2004), *Statistics in Musicology*, New York: Chapman & Hall.
- Biamonte, Nicole (2014), »Formal Functions of Metric Dissonance in Rock Music«, *Music Theory Online* 20/2. <http://www.mtosmt.org/issues/mto.14.20.2/mto.14.20.2.biamonte.html> (30.6.2017)
- Boyd, Malcolm (1999), *Bach. Chorale Harmonization and Instrumental Counterpoint*, London: Kahn & Averill.
- Brown, Matthew / Douglas J. Dempster (1989), »The Scientific Image of Music Theory«, *Journal of Music Theory* 33/1, 65–106.
- Budge, Helen (1943), *A Study of Chord Frequencies Based on the Music of Representative Composers of the Eighteenth and Nineteenth Centuries*, Ph. D., Columbia University.
- Burgoyne, John A. / Jonathan Wild / Ichiro Fujinaga (2013), »Compositional Data Analysis of Harmonic Structures in Popular Music«, in: *Mathematics and Computation in Music. 4<sup>th</sup> International Conference, MCM 2013, Montreal, QC, Canada, June 12–14, Proceedings*, hg. von Jason Yust, Jonathan Wild und John A. Burgoyne, Berlin: Springer, 52–63.
- Byros, Vasili (2009), »Towards an ›Archaeology‹ of Hearing. Schemata and Eighteenth-Century Consciousness«, *Musica Humana* 1/2, 235–306.
- Caplin, William E. (1998), *Classical Form. A Theory of Formal Functions for the Instrumental Music of Haydn, Mozart, and Beethoven*, New York: Oxford University Press.
- Clarke, Eric / Nicholas Cook (Hg.) (2004), *Empirical Musicology. Aims, Methods, Prospects*, New York: Oxford University Press.
- Cuthbert, Michael S. / Christopher Ariza (2010), »Music21. A Toolkit for Computer-Aided Musicology and Symbolic Music Data«, in: *Proceedings of the 11<sup>th</sup> International Society for Music Information Retrieval Conference, ISMIR 2010, Utrecht, Netherlands, August 9–13*, hg. von J. Stephen Downie und Remco C. Veltkamp, 637–642. <http://ismir2010.ismir.net/proceedings/ismir2010-108.pdf> (30.6.2017)
- Dahlhaus, Carl (1968), *Untersuchungen über die Entstehung der harmonischen Tonalität*, Kassel: Bärenreiter.
- (1978), »Der rhetorische Formbegriff H. Chr. Kochs und die Theorie der Sonatenform«, *Archiv für Musikwissenschaft* 35/3, 155–177.
- Daniel, Thomas (2000), *Der Choralatz bei Bach und seinen Zeitgenossen. Eine historische Satzlehre*, Köln-Rheinkassel: Dohr.

- De Clercq, Trevor / David Temperley (2011), »A Corpus Analysis of Rock Harmony«, *Popular Music* 30/1, 47–70.
- Eberlein, Roland (1994), *Die Entstehung der tonalen Klangsyntax*, Frankfurt a. M.: Lang.
- Eitan, Zohar (1997), *Highpoints. A Study of Melodic Peaks*, Philadelphia: University of Pennsylvania Press.
- Fischer, David H. (1970), *Historians' Fallacies. Toward a Logic of Historical Thought*, New York: Harper & Row.
- Fitsioris, George / Darrell Conklin (2008), »Parallel Successions of Perfect Fifths in the Bach Chorales«, in: *Proceedings of the Fourth Conference on Interdisciplinary Musicology (CIM08), Thessaloniki, Greece, 3–6 July*.
- Gauldin, Robert (1988), *A Practical Approach to Eighteenth-Century Counterpoint*, Englewood Cliffs (NJ): Prentice Hall.
- Gjerdingen, Robert O. (1988), *A Classic Turn of Phrase. Music and the Psychology of Convention*, Philadelphia: University of Pennsylvania Press.
- (1991), »Defining a Prototypical Utterance«, *Psychomusicology* 10/2, 127–139.
- (2007), *Music in the Galant Style*, New York: Oxford University Press.
- (2014), »Historically Informed: Corpus Studies«, *Music Perception* 31/3, 192–204.
- Gossett, Philip (1989), »Carl Dahlhaus and the ›Ideal Type‹«, *19<sup>th</sup>-Century Music* 13/1, 49–56.
- Haimo, Ethan T. (1988), »Haydn's Altered Reprise«, *Journal of Music Theory* 32/2, 335–351.
- (1995), *Haydn's Symphonic Forms. Essays in Compositional Logic*, Oxford: Clarendon Press.
- Hepokoski, James / Warren Darcy (2006), *Elements of Sonata Theory. Norms, Types, and Deformations in the Late-Eighteenth-Century Sonata*, New York: Oxford University Press.
- Hippel, Paul von / David Huron (2000), »Why do Skips Precede Reversals? The Effect of Tessitura on Melodic Structure«, *Music Perception* 18/1, 59–85.
- Holtmeier, Ludwig (1997), »Nicht Kunst, nicht Wissenschaft. Zur Lage der Musiktheorie«, *Musik & Ästhetik* 1/2, 119–146.
- Huron, David (1996), »The Melodic Arch in Western Folksongs«, *Computing in Musicology* 10, 3–23.
- (2006), *Sweet Anticipation. Music and the Psychology of Expectation*, Cambridge (MA): MIT Press.
- (2013), »On the Virtuous and the Vexatious in an Age of Big Data«, *Music Perception* 31/1, 4–9.
- Huron, David / Ann Ommen (2006), »An Empirical Study of Syncopation in American Popular Music, 1890–1939«, *Music Theory Spectrum* 28/2, 211–231.
- Ito, John P. (2014), »Koch's Metrical Theory and Mozart's Music. A Corpus Study«, *Music Perception* 31/3, 205–222.

- Jacoby, Nori / Naftali Tishby / Dmitri Tymoczko (2015), »An Information Theoretic Approach to Chord Categorization and Functional Harmony«, *Journal of New Music Research* 44/3, 219–244.
- Jannidis, Fotis / Hubertus Kohle / Malte Rehbein (Hg.) (2017), *Digital Humanities. Eine Einführung*, Stuttgart: Metzler.
- Jeppesen, Knud (1935), *Kontrapunkt. Lehrbuch der klassischen Vokalpolyphonie*, Leipzig: Breitkopf & Härtel.
- Kahneman, Daniel (2011), *Thinking, Fast and Slow*, New York: Farrar, Straus and Giroux.
- Kaiser, Ulrich (2002), *Der vierstimmige Satz. Kantionalsatz und Choralatz*, Kassel: Bärenreiter.
- Knippahls, Hans-Jürgen / Dirk Möller (1995), *Johann Sebastian Bach – Der Choralatz. Ein Lehrwerk*, Wolfenbüttel: Mösele.
- Kostka, Stefan / Dorothy Payne (1995), *Tonal Harmony. With an Introduction to Twentieth-Century Music*, 3. Auflage, New York: McGraw-Hill.
- Kühn, Clemens (2006), *Musiktheorie unterrichten – Musik vermitteln. Erfahrungen – Ideen – Methoden. Ein Handbuch*, Kassel: Bärenreiter.
- MacKay, David (2003), *Information Theory, Inference, and Learning Algorithms*, Cambridge: Cambridge University Press.
- Manning, Christopher D. / Hinrich Schütze (1999), *Foundations of Statistical Natural Language Processing*, Cambridge (MA): MIT Press.
- McGill University, *The McGill Billboard Project*. <http://ddmal.music.mcgill.ca/research/billboard> (30.6.2017)
- McHose, Allen I. (1947), *The Contrapuntal Harmonic Technique of the 18<sup>th</sup> Century*, New York: Appleton – Century – Crofts.
- Menke, Johannes (2015), *Kontrapunkt I. Die Musik der Renaissance*, Laaber: Laaber.
- Meyer, Leonard B. (1973), *Explaining Music. Essays and Explorations*, Berkeley (CA): University of California Press.
- (1989), *Style and Music. Theory, History, and Ideology*, Philadelphia: University of Pennsylvania Press.
- Moraitis, Andreas (2011), »Konvention, Intention und Konstruktion. Die Stimmführungsparallelen in den Choralätzen Johann Sebastian Bachs und Georg Philipp Telemanns«, *ZGMTH* 8/3, 427–464. <http://storage.gmth.de/zgmth/pdf/652> (30.6.2017)
- Moretti, Franco (2013), *Distant Reading*, London: Verso.
- Müller, Meinard (2015), *Fundamentals of Music Processing. Audio, Analysis, Algorithms, Applications*, Berlin: Springer.
- Nettheim, Nigel (1997), »A Bibliography of Statistical Applications in Musicology«, *Musiology Australia* 20/1, 94–106.
- Neuwirth, Markus (2010), »Does a ›Monothematic‹ Expository Design Have Tautological Implications for the Recapitulation? An Alternative Approach to ›Altered Recapitulations‹ in Haydn«, *Studia Musicologica* 51/3–4, 369–385.

- (2011), »Joseph Haydn's ›Witty‹ Play on Hepokoski and Darcy's *Elements of Sonata Theory* [...]«, *ZGMTH* 8/1, 199–220. <http://storage.gmth.de/zgmth/pdf/586> (30.6.2017)
- Piston, Walter (1941), *Harmony*, New York: Norton.
- Polth, Michael (2000a), *Sinfonieexpositionen des 18. Jahrhunderts. Formbildung und Ästhetik*, Kassel: Bärenreiter.
- (2000b), »Wie wissenschaftlich kann Analyse sein?«, in: *Klang–Struktur–Metapher. Musikalische Analyse zwischen Phänomen und Begriff*, hg. von Michael Polth, Oliver Schwab-Felisch und Christian Thorau, Stuttgart: Metzler, 63–74.
- Quinn, Ian (2014), »Hidden Structure. Music Analysis Using Computers and Music21. A Toolkit for Computer-Aided Musicology« [Review], *Journal of the American Musicological Society* 67/1, 295–307.
- Ratz, Erwin (1973), *Einführung in die musikalische Formenlehre. Über Formprinzipien in den Inventionen und Fugen J. S. Bachs und ihre Bedeutung für die Kompositionstechnik Beethovens*, 3. Auflage, Wien: Universal Edition.
- Richards, Mark (2016), »Film Music Themes. Analysis and Corpus Study«, *Music Theory Online* 22/1. <http://www.mtosmt.org/issues/mto.16.22.1/mto.16.22.1.richards.html> (30.6.2017)
- Rohrmeier, Martin (2013), »Musical Expectancy – Bridging Music Theory, Cognitive and Computational Approaches«, *ZGMTH* 10/2, 343–371. <http://storage.gmth.de/zgmth/pdf/724> (30.6.2017)
- Rohrmeier, Martin / Ian Cross (2008), »Statistical Properties of Harmony in Bach's Chorales«, in: *Proceedings of the 10<sup>th</sup> International Conference on Music Perception and Cognition, Hokkaido University, Sapporo (Japan), August 25–29 2008*, hg. von Ken-ichi Miyazaki, Mayumi Adachi, Yuzuru Hiraga, Yoshitaka Nakajima und Minoru Tsuzaki, 619–627.
- Rohrmeier, Martin / Thore Graepel (2012), »Comparing Feature-Based Models of Harmony«, in: *From Sounds to Music and Emotions – 9<sup>th</sup> International Symposium on Computer Music Modeling and Retrieval, CMMR 2012, London, UK, June 19–22, 2012, Revised Selected Papers*, hg. von Mitsuko Aramaki, Mathieu Barthelet, Richard Kronland-Martinet und Sølvi Ystad, Berlin: Springer, 357–370.
- Rohrmeier, Martin / Patrick Rebuschat (2012), »Implicit Learning and Acquisition of Music«, *Topics in Cognitive Science* 4/4, 525–553.
- Rosen, Charles (1988), *Sonata Forms*, 2. Auflage, New York: Norton.
- Salzer, Felix / Carl Schachter (1969), *Counterpoint in Composition. The Study of Voice Leading*, New York: McGraw-Hill.
- Saysr, Elizabeth / Gregory Proctor (2008), »Playing the ›Science Card‹. Science as Metaphor in the Practice of Music Theory«, in: *What Kind of Theory is Music Theory? Epistemological Exercises in Music Theory and Analysis*, hg. von Per F. Broman und Nora A. Engebretsen, Stockholm University, 35–62.
- Schaffrath, Helmut (1995), *The Essen Folksong Collection in the Humdrum Kern Format*, hg. von David Huron, Menlo Park (CA): Center for Computer Assisted Research in the Humanities.

- Schubert, Peter / Julie Cumming (2015), »Another Lesson from Lasso. Using Computers to Analyse Counterpoint«, *Early Music* 43/4, 577–586.
- Temperley, David (2001a), *The Cognition of Basic Musical Structures*, Cambridge (MA): MIT Press.
- (2001b), »The Question of Purpose in Music Theory. Description, Suggestion, and Explanation«, *Current Musicology* 66, 66–85.
- (2009), »A Statistical Analysis of Tonal Harmony«. <http://davidtemperley.com/kp-stats> (30.6.2017).
- (2011), »Composition, Perception, and Schenkerian Theory«, *Music Theory Spectrum* 33/2, 146–168.
- Temperley, David / Trevor de Clercq (2013), »Statistical Analysis of Harmony and Melody in Rock Music«, *Journal of New Music Research* 42/3, 187–204.
- Temperley, David / Leigh VanHandel (2013), »Introduction to the Special Issues on Corpus Methods«, *Music Perception* 31/1, 1–3.
- Tymoczko, Dmitri (2003), »Progressions fondamentales, fonctions, degrés: une grammaire de l'harmonie tonale élémentaire«, *Musurgia* 10/3–4, 35–64.
- (2010), »Local Harmonic Grammar in Western Classical Music«. <http://dmitri.mycpanel.princeton.edu/mozart.pdf> (30.6.2017).
- Volk, Anja / W. Bas de Haas (2013), »A Corpus-Based Study on Ragtime Syncopation«, in: *Proceedings of the 14<sup>th</sup> International Society for Music Information Retrieval Conference, ISMIR 2013, Curitiba, Brazil, November 4–8*, hg. von Alceu de Souza Britto, Fabien Gouyon und Simon Dixon, 163–168. [http://www.ppgia.pucpr.br/ismir2013/wp-content/uploads/2013/09/10\\_Paper.pdf](http://www.ppgia.pucpr.br/ismir2013/wp-content/uploads/2013/09/10_Paper.pdf) (30.6.2017)
- Watt, Henry J. (1924), »Functions of the Size of Interval in the Songs of Schubert and of the Chippewa and Teton Sioux Indians«, *British Journal of Psychology* 14/4, 370–386.
- Weih, Claus / Dietmar Jannach / Igor Vatolkin / Günter Rudolph (2016), *Music Data Analysis. Foundations and Applications*, Boca Raton (FL): CRC Press.
- White, Christopher (2014), »Changing Styles, Changing Corpora, Changing Tonal Models«, *Music Perception* 31/3, 244–253.
- White, Christopher W. / Ian Quinn (2016), »The Yale-Classical Archives Corpus«, *Empirical Musicology Review* 11/1, 50–58. <http://emusicology.org/article/view/4958> (30.6.2017)
- Wingfield, Paul (2008), »Beyond »Norms and Deformations«. Towards a Theory of Sonata Form as Reception History«, *Music Analysis* 27/1, 137–177.